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

Band 28

2019

DGPF





Beiträge

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

20. – 22. Februar 2019 in Wien

Dreiländertagung OVG – DGPF – SGPF Photogrammetrie - Fernerkundung - Geoinformation - 2019

Thomas P. Kersten (Hrsg.)

ISSN 0942-2870

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

Band 28



2019

Beiträge

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

20. – 22. Februar 2019 in Wien

Dreiländertagung OVG – DGPF – SGPF Photogrammetrie - Fernerkundung - Geoinformation - 2019

ISSN 0942-2870

Thomas P. Kersten (Hrsg.)

ISSN 0942-2870

Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF) e.V. Band 28, 725 S., Hamburg 2019 Hrsg.: Thomas P. Kersten

© Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF) e.V. München 2019

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: <u>geschaeftsstelle@dqpf.de</u>

Redaktion:

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

VORWORT

Die Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF), die Österreichische Gesellschaft für Vermessung und Geoinformation (OVG) und die Schweizerische Gesellschaft für Photogrammetrie und Fernerkundung (SGPF) veranstalten die gemeinsame

Dreiländertagung 2019 "Photogrammetrie, Fernerkundung und Geoinformation"

vom 20. bis 22. Februar 2019 in Wien. Die Veranstaltung schließt sich alle drei Jahre der Wissenschaftlich-Technischen Jahrestagung der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation an und wird abwechselnd von den drei Ländern Deutschland, Schweiz und Österreich organisiert.

Nach den vergangenen Veranstaltungen 2007 in Muttenz, 2010 in Wien, 2013 in Freiburg i. Br. und 2016 in Bern findet die diesjährige Tagung an der Universität für Bodenkultur Wien (BOKU) statt. Sie wird vor Ort durch die OVG und das Institut für Vermessung, Fernerkundung und Landinformation (IVFL) organisiert. Die BOKU versteht sich als Lehr- und Forschungsstätte für erneuerbare Ressourcen, in welcher das IVFL jene technischen Disziplinen abdeckt, welche sich mit raumbezogenen Daten beschäftigen.

Die Dreiländertagung wird am Mittwochnachmittag den 20.02.2019, durch die Präsidenten der drei Gesellschaften eröffnet, die zu den nachfolgenden Plenarsitzungen einladen. Es folgen Kurzvorträge und Poster zum Karl-Kraus-Nachwuchsförderpreis – der erste Abend schließt mit einem "Come Together". Am Donnerstag finden Fachvorträge, Posterpräsentationen und die bewährten Anwenderforen statt, der Abschluss am Abend ist eine festliche Veranstaltung im Wiener Rathaus mit der Verleihung des Karl-Kraus-Nachwuchsförderpreises. Am Freitag werden die Sitzungen mit den Fachvorträgen und die Anwenderforen weitergeführt. Mit der Schlussveranstaltung gegen Freitagmittag wird die Tagung abgeschlossen.

Wissenschaftler/innen, Entwickler/innen und Anwender/innen aus den Bereichen der Photogrammetrie, Fernerkundung und Geoinformatik haben 90 wissenschaftliche Beiträge angemeldet, von denen 62 im Tagungsband als Publikationen enthalten sind. Insgesamt haben 121 Autoren an den Beiträgen mitgearbeitet, wobei einige Institutionen aus ihren Arbeits- und Forschungsgruppen mehrere Beiträge eingereicht haben. Ein besonderer Dank gilt den Reviewern der eingereichten Beiträge sowie dem Editor für die Zusammenstellung des Tagungsbandes.

Durch eine Fachfirmenausstellung wird die wissenschaftliche Tagung abgerundet. Institutionen aus Wirtschaft, Forschung und Verwaltungen stellen dabei neueste Technologien, Produktentwicklungen, Informationen zu Aus- und Weiterbildung, Studiengangsinformationen sowie andere Aktivitäten oder Produkte vor.

Wir danken allen Autor/innen, Teilnehmer/innen und Organisationskräften, ohne deren Beteiligung und engagierten Einsatz die Tagung nicht möglich wäre.

Prof. Dr.-Ing. Uwe Stilla Präsident DGPF



Dipl.-Ing. Julius Ernst Präsident OVG



Dipl.-Ing. Stephane Bovet Präsident SGPF



INHALTSVERZEICHNIS

ANWENDUNG 3D STADTMODELLE

J. WILKENING, A. KAPAJ & J. CRON: Creating a 3D Campus Routing Information System with ArcGIS Indoors	. 1
J. FALKNER & L. EYSN: Kappazunder – Ein Mobile Mapping Tool als Lösung für Fragestellungen der Verwaltung im öffentlichen Raum!.	15
OPTISCHE 3D-MESSTECHNIK	
T. LUHMANN, M. CHIZHOVA, D. GORKOVCHUK, A. JEPPING & HEIDI HASTEDT: Kombination von terrestrischem Laserscanning, UAV- und Nahbereichs- photogrammetrie zur Erfassung komplexer Kirchenbauwerke in Georgien	27
P. HÜBNER, S. LANDGRAF, M. WEINMANN & S. WURSTHORN: Evaluation of the Microsoft HoloLens for the Mapping of Indoor Building Environments.	44
A. HANEL, P. SUDI, S. PFENNINGER, E. STEINBACH & U. STILLA: Filter-based Pose Estimation for Electric Vehicles Relative to a Ground-based Charging Platform Using On-board Camera Images	54
BILDANALYSE - COMPUTER VISION I	
T. PETERS & C. BRENNER: Automatic Generation of Large Point Cloud Training Datasets Using Label Transfer.	68
GEOINFORMATIK - METHODIK	
T. MEYER & A. BRUNN: 3D-Punktwolken-Management mit PostgreSQL/PostGIS für praktische Anwendungen in Geoinformationssystemen.	83

L. S. TESSEMA, R. JÄGER & U. STILLA:	
Extraction of an IndoorGML Model from an Occupancy Grid Map	
Constructed Using 2D LiDAR.	97

AUSWERTUNG VON FERNERKUNDUNGSDATEN

S. SCHMOHL & U. SÖRGEL: ALS Klassifizierung Mit Submanifold Sparse Convolutional Networks	111
AM. LOGHIN, J. OTEPKA, W. KAREL, M. PÖCHTRAGER & N. PFEIFER: Accuracy Analysis of Digital Elevation Models from very High Resolution Satellite Imagery.	123
M. WEINMANN & U. WEIDNER: Relevance Assessment of Spectral Bands for Land Cover and Land Use Classification: A Case Study Involving Multispectral Sentinel-2-like and Hyperspectral Data.	138
J. WOLF, S. DISCHER & J. DÖLLNER: Techniken zur kombinierten Darstellung von 2D-Bodenradar und 3D-Punktwolken zur Analyse des Straßenraums.	154

MOBILE MAPPING SYSTEME UND UAV I

S. KRAUSE, M. STRER, JP. MUND & T. SANDERS:	
UAV Remote Sensing Data Handling: A Transition from Testing	
to Long-Term Data Acquisition for Intensive Forest Monitoring.	167
R. LADSTÄDTER, P. LULEY, S. LADSTÄTTER & H. MAYER:	
UHD Mapping von Teststrecken für automatisiertes Fahren	175

FERNERKUNDUNG - MULTI- UND HYPERSPEKTRAL

J. WOLFE, T. BAHR & N. HOLZER:	
Creating Models of Custom Image Classification Workflows Using Machine Learning Techniques.	188
A. JUNG, M. VOHLAND, M. MAGYAR, L. KOVÁCS, T. JUNG, N. PÉTERFALVI, B. KELLER, F. SILLINGER, R. RÁK & K. SZALAY: Snapshot Hyperspectral Imaging for Field Data Acquisition in Agriculture	
(in Raspberry Plantation).	196

FERNERKUNDUNG - ANWENDUNG AGRAR I

M. Forkel, W. Dorigo, N. Carvalhais & K. Thonicke:	
Fernerkundung für Globale Klima-Vegetationsmodelle.	203
L. GRAF, I. KAUSCH, H. BACH & H. TOBIAS: Using Harmonic Analysis of Green LAI Time Series Obtained from	
Sentinel-2 Imagery for Daily Representation of Crop Growth in a	214
	- • •

FERNERKUNDUNG – HYDROLOGIE I

T. PIECHL & H. DÖLLER: Integrales Seevermessungsprojekt Wörthersee als Beispiel für Sensorfusion	226
I. TRINKS, W. NEUBAUER, T. TAYLOR, M. WALLNER, K. LÖCKER & J. LESKOVAR: Hochauflösende unterwasserarchäologische Prospektion oberösterreichischer Pfahlbauten und Seen mit Fächerecholot und Sediment-Sonar. 2	234
AUS- UND WEITERBILDUNG	
 C. LINDNER, C. MÜLLER, H. HODAM, C. JÜRGENS, A. ORTWEIN, J. SCHULTZ, F. SELG & A. RIENOW: Das Klassenzimmer im Weltraum – Anwendungen zur Erdbeobachtung zwischen Realität und Virtualität. 	241
R. MANSBERGER, T. BAUER, M. IMMITZER & F. VUOLO: Bildungsprojekte im Fachbereich Photogrammetrie, Fernerkundung und GIS: Herausforderungen, Chancen und Grenzen	256
V. WICH & A. BRUNN: Methodik des Blended Learnings in der Aus- und Weiterbildung der mobilen 3D Vermessung	270
M. TEUCHER & C. GLÄßER: E-Learning und Learning Analytics in der universitären Ausbildung der Geowissenschaften - Blended Learning im Masterstudium	282

FERNERKUNDUNG - ANWENDUNG AGRAR II

F. VUOLO, L. ESSL, T. SANDEN & A. SPIEGEL:	
Multidisziplinäre Überlegungen zur nachhaltigen N-Düngung unter	
Berücksichtigung der Möglichkeiten der satellitengestützten	
Präzisionslandwirtschaft.	288
A. KLINGLER, A. SCHAUMBERGER, F. VUOLO & E. M. PÖTSCH:	
Vergleich unterschiedlicher Methoden zur Schätzung	
von Blattflächenindex in Dauergrünlandbeständen.	301

FERNERKUNDUNG – HYDROLOGIE II

 G. PAULUS, KH. ANDERS, M. ERLACHER, P. MAYR & R. SCHNEEBERGER: Wörthersee 3D – Empirischer Vergleich von Flachwasser- bathymetriedaten: UAS Photogrammetrie versus Single Beam Echolot 	310
C. MULSOW, G. MANDLBURGER, C. RESSL & HG. MAAS: Vergleich von Bathymetriedaten aus luftgestützter Laserscanner- und Kameraerfassung.	318
K. RICHTER, D. MADER, P. WESTFELD & HG. MAAS: Analyse der Auswirkung von Wellen auf konventionelle Refraktions- korrekturmethoden in der Laserbathymetrie.	334
D. INNERHOFER, G. MANDLBURGER, F. STEINBACHER & M. AUFLEGER: Vermessung hochalpiner Seen mittels moderner Technologien.	347

BILDANALYSE - COMPUTER VISION II

J. KANG, L. CHEN, F. DENG & C. HEIPKE: Encoder-Decoder Network With Dilated Convolution For Local Structure Preserving Stereo Matching.	363
L. CHEN, D. ZHAO & C. HEIPKE: Complementary Feature Learning from RGB and Depth Information for Semantic Image Segmentation.	375
M. DOROZYNSKI, D. WITTICH & F. ROTTENSTEINER: Deep Learning zurAnalyse von Bildern von Seidenstoffen für Anwendungen im Kontext der Bewahrung des kulturellen Erbes.	387

FERNERKUNDUNG - ANWENDUNG FORST II

C. GINZLER, M. MARTY & L. T. WASER:	
Landesweite digitale Vegetationshöhenmodelle aus historischen	
SW-Stereoluftbildern – ein Schatz in den Luftbildarchiven.	400

FERNERKUNDUNG LIDAR - METHODIK

D. MADER, K. RICHTER, P. WESTFELD, R. WEIß & HG. MAAS: Detektion und Extraktion von Gewässersohlenpunkten aus ALB-Daten mittels Full-Waveform Stacking.	407
W. DOBLER, F. STEINBACHER, R. BARAN, W. BENGER, M. RITTER & W. LEIMER: Die Möglichkeiten der Full Waveform für bathymetrische Airborne LiDAR Daten.	418
R. SCHWARZ & M. PFENNIGBAUER: Detaillierte Analyse komplexer Wellenformen eines UAV Laser Bathymetrie Profilmessgerätes.	428
G. MANDLBURGER & H. LEHNER: Single Photon LiDAR – Grundlagen Und Erste Evaluierungsergebnisse	443

MOBILE MAPPING SYSTEME UND UAV II

S. BLASER, S. NEBIKER & D. WISLER: Portables bildbasiertes Mobile Mapping System im Einsatz unter Tage – Systemaufbau, Georeferenzierungs- und Genauigkeitsuntersuchungen 458
M. CRAMER, G. MANDLBURGER, D. LAUPHEIMER, N. HAALA & P. HAVEL: Potenzial ultrahoch-auflösender und -genauer UAV-basierter 3D-Datenerfassung
HJ. PRZYBILLA: Untersuchungen zur Leistungsfähigkeit aktueller bildgebender UAV-Sensoren
FERNERKUNDUNG - ANWENDUNG FORST III
M. BRUGGISSER, M. HOLLAUS, D. WANG & N. PFEIFER: Abgrenzung homogener Waldbestände aus ALS-Punktwolken
S. GANZ, Y. KÄBER & P. ADLER: Waldinventur aus Luftbildern und LiDAR-Daten - Mit welcher Genauigkeit und Präzision lassen sich Baumhöhe, Kronenradius und Kronenansatz von Douglasien ableiten?
S. MUSTAFIĆ & M. SCHARDT: Deep Learning-basierte Baumartenklassifizierung auf Basis von ALS-Daten
Fernerkundung – Hydrologie III

K. BAUMGARTNER, R. KLAR & M. AUFLEGER: Evaluierung der Airborne LiDAR Bathymetrie in der Fließgewässer- Modellierung.	. 537
R. BOERNER, L. HOEGNER & U. STILLA: Voxelbasierte Änderungsdetektion des Gewässerbodens in ALB Daten unter Verwendung von semantischen Referenzdaten.	. 548
C. ULRICH, L. BANNEHR, M. HUPFER & M. SCHULZE: Entwicklung von Indizes zur Ableitung der Eisenkonzentration verockerter Fließgewässer mittels hyperspektraler VNIR- und SWIR-Daten.	. 558

POSTER

V. KAUFMANN & W. SULZER: Dokumentation des Gletscherrückgangs am Gössnitzkees für den Zeitraum 1982-2018 – eine Gletschergeschichte mit Ablaufdatum	572
G. SEIER, M. WECHT & W. SULZER: Erfassung von Veränderungen der Sedimentkörper eines Wildbaches (Krumeggerbach, Wölzer Tauern) mittels unbemanntem Luftfahrzeug	587
T. MIKL & V. KAUFMANN: Kostengünstige 3D-Dokumentation von Skulpturen des Barockbildhauers Philipp Jakob Straub	597
I. PFEIL, M. VREUGDENHIL, P. DORNINGER & W. WAGNER: Räumlicher Zusammenhang zwischen Sentinel-1 Polarisationsdaten und Vegetationshöhe in einem landwirtschaftlichen Gebiet in Niederösterreich.	606
R. BECKER: Multi-Temporal Aerial Imagery for Automatic Mapping and Trend Analysis.	616
M. KAMPEN, S. LEDERBAUER, JP. MUND & M. IMMITZER: UAV-Based Multispectral Data for Tree Species Classification and Tree Vitality Analysis.	623

BEITRÄGE KARL-KRAUS-NACHWUCHSPREIS

M. WEIRATHER: Automatisierte Detektion von Waldbrandflächen mit Sentinel-2 Satellitendaten.	640
A. F. MEYER, N. LACK & D. JORDAN: Wildtier-Monitoring mit UAVs - Künstliche Intelligenz zur automatisierten Detektion von Infrarot-Signaturen.	647
J. KNÖTTNER, D. ROSENBAUM & A. BRUNN: Trennung von parkenden und am Verkehr teilnehmenden Fahrzeugen basierend auf einer automatischen Verkehrserfassung aus Luftbildern	654
L. WINIWARTER & G. MANDLBURGER: Classification of 3D Point Clouds using Deep Neural Networks.	663
S. KÖNIG, J. A. SCHULTZ, A. SCHOCH, J. BLÖTHE, L. SCHROTT & F. THONFELD: Mountain Permafrost Distribution Modeling – A Geomor-phometry-Remote Sensing Approach for the Hohe Tauern National Park, Austria.	675
M. SPERLICH & H. WEINACKER: Development of large scale 3D point cloud processing modules – minimizing the number of input parameters through statistical modelling and optimization.	686
M. KÖLLE, D. LAUPHEIMER & N. HAALA: Klassifikation hochaufgelöster LiDAR- und MVS-Punktwolken zu Monitoringzwecken	692

H. AUGUSTIN, M. SUDMANNS, D. TIEDE & A. BARALDI: Semantic Data Cubes Utilising Free and Open-Access EO-Data to Generate Spatially-Explicit Evidence for Environmental Monitoring: Applied Use-Case in Syria Based on Sentinel-2 Data.	702
S. NIETIEDT, O. KAHMEN & T. LUHMANN: Untersuchung von Orientierungs- und Matchingverfahren für die hochgenaue 3D-Oberflächenerfassung von Schweißnähten mit einem mobilen Kamerasystem.	711
AUTORENVERZEICHNIS	720
PROGRAMMKOMITEE	725

Creating a 3D Campus Routing Information System with ArcGIS Indoors

JAN WILKENING¹, ARMAND KAPAJ² & JULIANE CRON²

Abstract: Humans spend an increasing amount of time indoors. While it is relatively easy to find the best route between buildings, it is more difficult to find the best route within a building. This is a particular problem for large, complex buildings with a high fluctuation of visitors like in university buildings. In this contribution, we show how an interactive web-based 3D campus routing and information system (CRIS) can be designed using ArcGIS Indoors and the ArcGIS platform. From a scientific point of view, we focus on the question how this system needs to be designed in order to facilitate effective and efficient indoor routing. The most important design elements of the CRIS were identified by conducting a user study with experts in Cartography. This study showed that the system can help users without prior knowledge in solving routing tasks.

1 Introduction and problem statement

The amount of time that humans spend in indoor spaces is significant. A series of studies carried out on different continents at different times indicates that the average person spends approximately 90% of their time in indoor environments (home, work, and other indoor spaces) (JENKINS et al. 1992; BRASCHE & BISCHOF 2005; SCHWEIZER et al. 2006). In the last decade, indoor navigation began to become a hot topic for research (FALLAH et al. 2013; CHO & CHOI 2015, DUDAS; GHAFOURIAN & KARIMI 2009). Furthermore, according to CHO & CHOI (2015) these days at least 170 companies work on topics like indoor location, indoor maps, building tracking and indoor navigation.

Humans are familiar to navigate outdoors and can use a wide variety of devices to orientate themselves (e.g. maps, navigation systems, street signs etc.). However, navigation within building premises is not an easy task to complete, especially due to the lack of familiarity with the premises, inadequate indoor maps and missing signage. Therefore, users often end up getting lost. According to CARLSON et al. (2010) this happens due to three main factors: (1) the spatial structure of the buildings, (2) the cognitive maps that users construct while navigating, and (3) the strategies and spatial abilities of the users.

Finding the way around and inside university campuses is a particular challenge, due to the special spatial structures. University campuses often consist of several buildings with various connection routes, entrance points, floor levels, dimensions, confusing numbering and naming systems, limited lines of sight, lack of local cues, missing signage, and a lack of well-designed and user-friendly indoor maps. These factors impair the indoor navigation process of various user groups on and between campuses. Moreover, university campuses deal with a high fluctuation of various user

¹ Esri Deutschland GmbH, Ringstraße 7, D-85402 Kranzberg, E-Mail: j.wilkening@esri.de

² Technical University of Munich, Chair of Cartography, Arcisstraße 21, D-80333 Munich, E-Mail: armand.kapaj@hotmail.com, juliane.cron@tum.de

groups such as students (current and new students), staff members, and visitors. The process of indoor navigation in university buildings is particularly difficult for first-time students and visitors, who are not familiar with the architectural structure of the building.

To tackle these problems, we implemented a web-based 3D Campus Routing Information System that is designed to support users in finding their way on university campuses based on ArcGIS Indoors. A campus routing information system can be defined as a system designed and build to help various user groups to navigate in complex indoor spaces, such as university campuses, airports, shopping malls, etc. While this paper focuses on university campuses, the methodology can be transferred to several other campuses including shopping malls, airports, etc. The several steps for creating and evaluating the campus routing system will be discussed in the remainder of this article.

2 Creation of a 3D Campus Routing Information System

Building university Campus Routing Information Systems (CRIS) requires different steps to transform the raw data (e.g. building data CAD files) into an effective and efficient campus routing application with the help of different software, tools, and methods. These steps can be categorized in four main groups: (1) data pre-processing, (2) data processing, (3) web application, and (4) campus routing evaluation, as shown in Figure 1 below.



Fig. 1: Steps to design a Campus Routing Information System (CRIS)

For the implementation of our 3D CRIS, we chose Esri's ArcGIS platform. This platform solves the tasks of data pre-processing, processing and publishing very efficiently, because it consists of several components (ArcMap, ArcGIS Pro, ArcGIS Online, ArcGIS Server, ArcGIS API for JavaScript, Survey123) that can be seamlessly integrated with each other: CAD data can be transferred to GIS data with built-in tasks for ArcGIS Pro, spatial data from Desktop GIS (ArcMap and ArcGIS Pro) can be uploaded and hosted as web services with REST endpoint in ArcGIS Online,

and these REST endpoints can be used in several web mapping applications, for instance with the ArcGIS API for JavaScript. The application itself can be evaluated with Survey123 for ArcGIS. Table 1 represents the different steps of our workflow, including the concrete process and the software component that was used.

Workflow	Process	Software	
	Generalization and inspection of CAD files		
Data pre-pro- cessing	Generating new CAD files	ArcGIS Desktop 10.6	
	Projection	ArcGIS Pro	
Data processing	Create and publish campus basemap	ArcGIS Pro and ArcGIS Online	
	Create building interiors	ArcGIS Pro: ArcGIS In- doors and ArcGIS Server	
	Create campus scene		
	Create and publish campus network		
	Publish campus scene and locator layers	ArcGIS Online	
Web application	Setting up and configuring the 3D campus ap-	ArcGIS API for JavaScript	
	plication	4.4	
	First evaluation: design and visualization	Sunvoy122 for ArcCIS	
USEI SIUUIES	Second evaluation: usability and utility	Survey 123 IOLAICOIS	

Tab 1: Workflow,	processes and softw	vare components nece	ssary to create the CRIS
------------------	---------------------	----------------------	--------------------------

A similar approach for designing a CRIS was taken by WILKENING et al. (2018) for a campus routing system at the University of Applied Sciences in Würzburg³. However, the CRIS for Würzburg was built from scratch, without making any use of templates. For the scope of this contribution, we used ArcGIS Indoors, which is a novel indoor mapping product provided by Esri⁴. ArcGIS Indoors is a complete mapping system for assembling, managing and sharing building and campus information. It is used for location discovery and wayfinding, asset management, operational data analysis, and crowdsource reporting to keep the indoor environment functional. It constitutes a "one stop-shop" application to complete all the necessary steps to generate a 3D CRIS.

At the time of our study, a beta version of ArcGIS Indoors was available, which consisted of two major parts: (1) an ArcGIS Pro project including built-in tasks for all the necessary steps to complete the data processing steps, and (2) a web application folder, which contains a set of folders used to create a web application template.

ArcGIS Indoors covers the entire workflow for adding floorplan information to the ArcGIS platform. Once the indoor CAD data are converted to GIS data in the ArcGIS platform, they can be styled for 2D and 3D maps and scenes, published as web services, and consumed in a wide range of applications. ArcGIS Indoors' web application folder is based on the ArcGIS API for JavaScript 4.4, which is used for the map and interface visualization process and interaction, and contains several HTML, CSS and JavaScript files.

³ <u>https://gis.fhws.de/campus/campus_roeri_3D.html</u>

⁴ https://www.esri.com/en-us/landing-page/product/2018/arcgis-indoors

2.1 Case study: The Technical University of Munich (TUM) Campus

Our CRIS was implemented as a Master thesis project at the Technical University of Munich (TUM). While the TUM facilities are located at several sites in the Munich area, we focus on the main building of the Munich downtown campus at Arcisstraße as a case study for the implementation of the Campus Routing Information System.

This campus represents an ideal example where the indoor navigation process is not an easy task to handle, due to several reasons: presence of several buildings, many entrance points, different numbering and naming systems for each individual building, lack of signage, continuous construction sites, several connection routes in several floor levels, connections between the buildings, mezzanine floor levels, restricted access to several points, etc.

TUM provides different means to facilitate the process of indoor navigation: (1) floorplans at the entrance of each building, and (2) a web application service "Roomfinder"⁵. However, neither of these options solve the indoor navigation problem to a satisfying degree, as they do not provide room information, interactivity, 3D representations or options for route planning and communication.

LORENZ et al. (2013) state that existing indoor navigation systems make use of building blueprints to represent interior spaces, but they do not fulfill cartographic requirements for map design and visualization techniques. Moreover, WINTER et al. (2017) state that floorplan maps are notorious for their difficult reading, require advanced mental rotation and orientation skills. These factors impair efficient route finding also within the TUM campus to many users. Therefore, it is a crucial task to develop a more suitable application for facilitating the indoor navigation process within the TUM main campus.

3 Implementation steps: From CAD data to CRIS

In this section, we describe the workflow that we used to implement an interactive 3D web application with routing functionality based on CAD files as input data.

3.1 Data pre-processing

ZLATANOVA et al. (2013) state that the software tools used for indoor modeling are largely generic CAD or computer graphics tools. In addition, GOTLIB & MARCINIAK (2012) mention that the conventional way of representing indoor spaces cannot be used in indoor navigation systems due to the high level of detail, which will result in a cognitive load to the users. Therefore, to avoid confusion coming from the highly detailed CAD files, a pre-processing step is highly required to create a CRIS.

The CAD files for the TUM main campus were provided by the Chair of Cartography at TUM. These files are in DWG file format, which is a proprietary binary file format for storing two- and three-dimensional CAD data. Before converting these files to a GIS file format, a prior inspection should be performed to check the validity of CAD data.

⁵ <u>https://portal.mytum.de/campus/roomfinder</u>

Our pre-processing considered the following aspects of the CAD files:

- *Generalization*, to identify if the CAD files depict relevant information for facilitating users' indoor navigation process.
- *Categorization of floorplan lines* based on their type (wall, door, window, stair, etc.), and applying the same naming scheme across the building (i.e., if the interior wall is labeled as "interior_wall" in one floor, it should have the same name in all the floors).
- *Categorization of interior spaces* based on their type (lecture hall, seminar rooms, office, hallway, etc.), with the help of annotations. Annotations enhance the CAD files with textual information related to the space function and identification number.
- *Projection* of CAD files to a real-world coordinate system, to align the input data to the right position in the map.

The data pre-processing constitutes a very important step, as it will provide the basis for all the following steps. In order to create an effective and efficient CRIS, a lot of manual work might be needed, whose benefits will be evident in the later steps.

3.2 Data processing with ArcGIS Indoors

Data processing makes use of the generated CAD files to create a customizable Indoor GIS and a point-to-point routing network that connects all the indoor spaces into a single network. An Indoor GIS does not only contain basemaps for visualization purposes, but also attributive data for the features such as buildings, floors, floorplan lines (wall, stairs, windows, etc), and indoor spaces (numbers, persons, room categories etc).

The data processing workflow includes the following steps: (1) creating building interiors, (2) visualizing and designing a campus scene, (3) creating a campus network, and (4) creating the web application. These steps are described in detail below.

3.2.1 Creating building interiors (CAD to GIS)

KRISP et al. (2014) point out that the functions of an indoor navigation system based on CAD file format are very limited, as indoor knowledge also includes topological relationships between indoor objects, which is not easy to identify. Indoor knowledge includes topological relationships between indoor objects. Converting CAD files to a GIS file format solves this issue. Therefore, this step constitutes the starting point in creating an Indoor GIS.

ArcGIS Indoors offers the possibility to add, view and manipulate CAD files. In addition, the building attributes stored in the CAD files are automatically converted into an Indoor GIS, using a configurable Excel file that is available in the ArcGIS Indoors download folder. This Excel file contains building properties information, CAD layer mapping and file information, which can be edited and modified.

The result of this CAD-GIS conversion step is an interactive 3D scene populated with building floorplan lines, interior spaces, floor and building footprint (see Figure 2). These features constitute the basis for our Indoor GIS and will be used further to visualize and design the campus scene to create a cartographically appealing CRIS.





3.2.2 Visualizing and designing a campus scene

The Indoor GIS layers can easily be styled and designed following cartographic design principles and allow a 3D visualization for a more realistic perspective, based on their enabled Z-values. This step was used to symbolize interior lines and spaces according to their type attributes and intended usage, as well as to symbolize the building façade.

The process of symbolizing interior lines was used to select non-traversable wall lines that are used to create the indoor network. Interior lines (walls and doors) are extruded to one-meter height to create partial-height lines, which are adequate to see the floorplan layout in 3D and not obscure the indoor route. The symbolization process of interior spaces was used to classify these spaces based on their intended usage (i.e., office, lecture halls, library, toilets, etc.). To symbolize these spaces, we applied a color-coded scheme. Furthermore, we used interior space polygons to generate interior space points (locator layers), which are used for network generation as "address" attribute for locating and routing between the interior spaces in the TUM CRIS. The designed campus scene and the locator layers were published as web services with REST endpoint in the in the cloud-based portal ArcGIS Online. These shared services formed the basis for the web application (see 3.2.4).

3.2.3 Campus network

Creating a campus network constitutes the most important step of the campus routing information system, as it enables a point-to-point routing in the campus routing web app. The campus network is based on the shortest-route algorithm. This algorithm will create a lattice-based (fish-net) network that covers every walk-able area of the building (see Figure 3).



Fig. 3: Lattice-based network (detailed view in inset map on lower right corner)

The route is computed based on a more realistic approach of human walking behavior in open spaces (walking diagonally instead of walking parallel to the walls). In addition to creating an indoor campus network, ArcGIS Indoors offers the possibility to create and connect outdoor pathways around buildings in case the campus consists of several buildings. The developed CRIS for the TUM does not yet offer the possibility of localization to the users. Therefore, the users need to rely on visualized spatial structures such as doors, walls, windows, stairs, stair handlers, etc. to identify their position and to navigate within the TUM building. As networks cannot be shared as services to ArcGIS Online yet, we had to use our ArcGIS Server site to share the network as a web service. Thus, the network could be used within the web application as the routing service for the CRIS.

3.2.4 Web application

As described above, the ArcGIS Indoors ZIP archive contains a web application template based on the ArcGIS API for JavaScript. By default, the web application displayed the Esri Campus in Redlands, California. In order to display the TUM Campus data and the web services we created in the previous steps, we had to modify the "config.js" file with the URL of the REST endpoint and the ArcGIS Online portal IDs of the published services (campus scene, locator layers, and campus network). The JavaScript file is highly customizable and allows several changes on design and parameters.

The main view of the TUM CRIS is composed of a campus web scene in 3D view and various functionalities like search panel, information panel and navigation buttons (see Figure 4).



Fig. 4: The main view of the TUM CRIS upon launching the web application

These functionalities are designed to provide information and help users to interact with the system, and presented in the form of toggle buttons. Each of these buttons contains information upon hover. These functionalities are provided as part of the web application template.

Figure 5 demonstrates an example of the route communicated to the user via the TUM CRIS. The dark-yellow 3D polyline displays the default route visualization that follows the stairs. An office on the ground floor serves as the starting point of the route in the case study, while the destination point is an office on the 5th floor. In addition, the side panel provides users with the time and distance to reach the destination. If users toggle the "*Need an elevator*?" button, the route between the two spaces is re-calculated and uses the closest elevator to generate a route to the destination. This route is also represented in 3D and is visualized using a blue color.



Fig. 5: Visualizing the default route in the web application

4 Evaluation

After implementing our CRIS for TUM, we conducted two user studies. The first study focused on map design and visualization, while the scope of the second study was the usability and the utility of the CRIS. The methodologies adopted for these studies varied from quantitative to qualitative, as well as mixed methods adopted for this purpose.

4.1 Map design and visualization

The first user study aimed to identify to which extent the applied designed principles for the modeling of indoor spaces were appropriate, and how our CRIS can be further improved. Based on the methods proposed by ROTH et al. (2017), the first evaluation is based on the expert-based method. Expert-based methods solicit input and feedback about an interactive map from consultants with training and experience in map and interface design and evaluation. It is important that an expert is a person from outside the project team, as it is necessary that he or she has little or no prior knowledge about the interactive map and the interface. Therefore, the selection of participants for this user study was chosen based on their background.

Forty-one participants from different countries participated in this evaluation, of which 24 participants were female and 17 male. The dominant age group of participants was 25 to 34 years. Most of them had a cartography (18 participants) or geography (12 participants) background. The educational level of participants was distributed as follow: Master (23), Bachelor (10), PhD (5), and higher than PhD (3).

The participants filled an online questionnaire designed with Survey123 for ArcGIS. In this questionnaire, participants were asked about their opinion regarding the applied design principles to visualize several indoor features. Most of the questions for this section were created in the form of a comparative evaluation, where the users were provided with various design methods, illustrated with screenshots taken from the CRIS. Afterwards, a statement was formulated to gather users' opinion. The answers were collected on a Likert scale and the results gathered from them are illustrated in Figure 6 below.



Fig. 6: Users' opinion regarding map design questions

In general, most of the participants (N=35) considered it easy or extremely easy (see Figure 7) to follow the route and reach the destination with our CRIS. 31 participants were satisfied or very satisfied with the 3D route representation (see Figure 8). Regarding advantages of a 3D map representation for indoor navigation systems, about half of the participants named a realistic representation as the biggest advantage, followed by the representation of vertical structures like walls, doors and stairs.



Dreiländertagung der DGPF, der OVG und der SGPF in Wien, Österreich – Publikationen der DGPF, Band 28, 2019

Fig. 7: Results of the question "How convenient is it for you to follow the route and reach the destination?"



Fig. 8: Results of the question "How satisfied are you with the 3D route representation?"

After gathering feedback from this first user study regarding map design and map usage of the TUM CRIS, we developed the final version of the TUM CRIS based on the following design principles:

• Doors are visualized with a level of transparency, where the route is always visible.

- Interior features are extruded to 1m height to allow a visualization of the whole route and visible floorplan.
- Stairs and stair handlers are visualized in a 3D representation to facilitate users' indoor navigation among floor levels.
- The textured model of the building is used as the main view of the TUM CRIS, and when the users generate a route, the exterior walls of the building will be transparent.
- Two different colors are used to communicate the route that follows the stairs and elevators.

To test the effectiveness and efficiency of the final version of the TUM CRIS, a second user study was conducted.

4.2 Usability and utility

The purpose of the second user study was to evaluate the usability and utility of the developed CRIS for the TUM main building. The usability and utility were tested using a scenario-based design, which implies the test persons were asked to navigate from their starting point to a destination point with the help of the CRIS. The evaluation process is built on the user-based methods (ROTH et al. 2017), which solicits input and feedback about an interactive map and interface from a representative set of target users. First-time visitors were chosen as the target group to test the developed CRS because of their lack of familiarity with the TUM building. Their prior knowledge of the building spatial structure, naming and numbering system, presence of landmarks, etc. will not interfere during the interaction with the TUM CRIS. Additionally, they were chosen, since they are an unacquainted user group. Hence, if a visitor is able to use and navigate with the TUM CRIS, it is assumed that the CRS is effective and efficient also for everyday users of the TUM premises (students and staff members).

A mixed method of quantitative and qualitative data was adopted for this evaluation. The user study consisted of two main parts: (1) a user experiment conducted within the premises of the TUM main campus building, (2) a questionnaire to gather feedback provided after finishing the experiment. The study was conducted with five participants (3 males, 2 females).

Within the task, our participants were asked to identify their position in the building by using structural features as orientation points, as our system did not feature any indoor positioning functionality. All the five participants were able to identify their position, which was used later as starting point or origin of the route. The next step was to locate a certain room with a certain number. To solve this task, our participants had to make use of the search panel on the developed CRIS. As part of the thinking aloud adopted method, the participants were asked to describe what direction they had to follow to reach the destination. In addition to thinking aloud, users were asked to determine their position related to other structural features (stairs, doors, and stair handlers) as they moved along the route. Using these proposed landmarks, all participants used the system every time they reached a decision point that required going up or down, or turning left or right. To determine their next move, they used the zooming, panning and rotation functions of the system. All five participants mentioned that they would like to use a similar indoor navigation system for other indoor spaces in the future and agreed on the usability and utility of our system.

Based on the feedback form, the following conclusions can be drawn for the interface design of the TUM CRIS:

- The general information provided on how to use the TUM CRIS is relevant and useful.
- The search box to find an indoor space is easy to use.
- The process to generate a route between indoor spaces is easy to handle.
- The buttons and tooltips are needed and helpful to use the CRIS.

5 Summary and outlook

In this article, we have shown how to create and design a web-based 3D Campus Routing and Information system (CRIS) that facilitates orientation on university campuses. This method can easily be transferred to any other building complex and is not limited to university campuses. Many further developments are expected for ArcGIS Indoors that were not yet part of the product when conducting our study, such as a native mobile application or indoor positioning. Since most universities around the world can use the full stack of ArcGIS technology including a wide range of apps, ArcGIS Server etc. within site licenses, these developments offer a wide range of possibilities, both for developers and for researchers. Regarding future research questions, it is of particular interest whether and to which extent CRIS actually help users in finding their way more efficiently on campuses, and which elements of the CRIS contribute to this increase in efficiency. Regarding the design of campus applications, our evaluation suggests that indoor structures should be visualized with a transparent design, so that the route is always visible. This seems to be more important than a realistic opaque representation, which would imply that the thematically relevant information (the route) was less salient.

Based on the results of the evaluation tests it can be concluded that the TUM CRIS is an indoor navigation web based application, which makes use of structural features to plan and convey the route to various users in an effective and efficient way.

While our user studies have shown that our participants enjoyed using the CRIS, this does not automatically imply that a wide range of users would actually use this system on-campus: Most people (still) tend to navigate indoors without maps. Thus, it remains an open question for future research how a CRIS should be designed that users actually use it, and whether a CRIS actually can improve human indoor wayfinding.

6 Literature

- BRASCHE, S. & BISCHOF, W., 2005: Daily time spent indoors in German homes Baseline data for the assessment of indoor exposure of German occupants. International Journal of Hygiene and Environmental Health, 208, 247-53.
- CARLSON, L.A., HÖLSCHER, C., SHIPLEY, T.F. & DALTON, R.C., 2010: Getting Lost in Buildings, Current Directions in Psychological Science, **19**, 284-89.
- CHO, Y. C. & CHOI, J.F., 2015: Spatial Information-Based 3D GIS for Indoor & Outdoor Integrated Platform Development from CRETA Platform, International Journal of Computer and Communication Engineering, 4, 397.

- DUDAS, P.M., GHAFOURIAN, M. & KARIMI, M.H., 2009: ONALIN: Ontology and Algorithm for Indoor Routing. Tenth International Conference on Mobile Data Management: Systems, Services and Middleware, 720-725.
- FALLAH, N., APOSTOLOPOULOS, I., BEKRIS, K. & FOLMER, E., 2013: Indoor Human Navigation Systems: A Survey, Interacting with Computers, **25**, 21-33.
- GOTLIB, D. & MARCINIAK, J., 2012: Cartographical Aspects in the Design of Indoor Navigation Systems, Annual of Navigation, **19**, 35-48.
- JENKINS, P.L., PHILLIPS, T.J, MULBERG, E.J. & HUI, S.P., 1992: Activity patterns of Californians: Use of and proximity to indoor pollutant sources, Atmospheric Environment. Part A. General Topics, **26**, 2141-48.
- KRISP, J., JAHNKE, M., LYU, H., & FACKLER, F., 2014: Visualization and Communication of Indoor Routing Information. Progress in Location-Based Services 2014, Gartner, G. & Huang, H. (eds.), Springer, 33-44.
- LORENZ, A., THIERBACH, C., BAUR, N. & KOLBE, T.H., 2013: Map design aspects, route complexity, or social background? Factors influencing user satisfaction with indoor navigation maps', Cartography and Geographic Information Science, **40**, 201-09.
- ROTH, R. E., ÇÖLTEKIN, A., DELAZARI, L., FILHO, H. F., GRIFFIN, A., HALL, A., KORPI, J., LOKKA, I., MENDONÇA, A. & OOMS, K., 2017: User studies in cartography: opportunities for empirical research on interactive maps and visualizations. International Journal of Cartography, 3, 61-89.
- SCHWEIZER, C., EDWARDS, R. D., BAYER-OGLESBY, L., GAUDERMAN, W.J., ILACQUA, V., JUHANI JANTUNEN, M., LAI, H.K., NIEUWENHUIJSEN, M. & KÜNZLI, N., 2006: Indoor timemicroenvironment-activity patterns in seven regions of Europe, Journal Of Exposure Science And Environmental Epidemiology, 17, 170.
- WILKENING, J., SCHÄFFNER, R. & STAUB, T., 2018: Interactive 3D Route Planner for the Campus Röntgenring in Würzburg. AGIT Journal für Angewandte Geoinformatik, 35-41.
- WINTER, S., TOMKO, M., VASARDANI, M., RICHTER, K.F., & KHOSHELHAM, K., 2017. Indoor localization and navigation independent of sensor based technologies, SIGSPATIAL Special, 9, 19-26.
- ZLATANOVA, S., SITHOLE, G., NAKAGAWA, M. & ZHU, Q., 2013. Problems in indoor mapping and modelling. Acquisition and Modelling of Indoor and Enclosed Environments, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, 40(4/W4), 63-68.

Kappazunder – Ein Mobile Mapping Tool als Lösung für Fragestellungen der Verwaltung im öffentlichen Raum!

JOHANNES FALKNER¹ & LOTHAR EYSN¹

Zusammenfassung: Die Digitalisierung hält in der Millionenstadt Wien Einzug. Dadurch entstehen neue Bedürfnisse und Möglichkeiten für eine effiziente Verwaltung. Im Programm "Wien gibt Raum" wird aktuell die Zuständigkeit für die Genehmigung und Verwaltung von Objekten im öffentlichen Raum neu organisiert. Eine großangelegte Mobile Mapping Kampagne im gesamten Wiener Stadtgebiet stellt dafür eine neuartige Datengrundlage dar. In einem digitalen Bilddienst namens "Kappazunder" können Objekte des öffentlichen Raumes mit guter geometrischer Genauigkeit erfasst und klassifiziert werden. In zwei Pilotprojekten werden zeitkritische Fragestellungen im öffentlichen Raum gelöst, wodurch das Potential dieser neuen Methode erstmals sichtbar wird. In diesem Beitrag wird ein praktischer Einblick in den Aufbau eines digitalen Bilddienstes sowie die Methodik zur großflächigen Inventarisierung von Werbeelementen und Verkehrszeichen aus Mobile Mapping Daten gegeben.

1 Programm "Wien gibt Raum"

Im Zuge des Programmes "Wien gibt Raum" organisiert der Magistrat der Stadt Wien die Zuständigkeiten für die Genehmigung und Verwaltung von Objekten und Aktivitäten im öffentlichen Raum neu. Dabei werden interne Prozesse der Stadtverwaltung hinter einer neu gestalteten digitalen Schnittstelle für Kundinnen und Kunden aufgestellt und durch speziell angepasste Softwaretools unterstützt (näheres in STRONDL et al. 2018). Eine wichtige Grundlage für diese Neuorganisation ist eine großflächige Inventarisierung von bestehenden Objekten im öffentlichen Raum. Neben einem Überblick über die tatsächlich im öffentlichen Raum befindlichen Objekte ist die Verortung mittels Raumbezug ein wichtiges Ziel. Diese Objekte, auch Stadtmöblierung genannt, sollen möglichst vollständig und mit guter geometrischer Genauigkeit erfasst werden, um in einem Fachinformationssystem für weiterführende Aufgaben verfügbar zu sein.

2 Mobile Mapping Kampagne

Für die Inventarisierung der Objekte haben sich in einer Vorstudie klassische Vermessungsmethoden als unwirtschaftlich und kaum realisierbar erwiesen, wodurch im Projekt Wien gibt Raum eine innovative Erfassungsmethode gewählt wurde. Als Datengrundlage für die Inventarisierung wurde eine bildgebende Mobile Mapping Kampagne für das Wiener Stadtgebiet realisiert. Weiterführende Informationen finden sich in STRONDL et al. (2018).

¹ Magistrat der Stadt Wien, MA 41 – Stadtvermessung, Muthgasse 62, A-1190 Wien, E-Mail: [johannes.falkner, lothar.eysn]@wien.gv.at

2.1 Eckdaten

Das Ziel der Kampagne war, ein qualitativ möglichst hochwertiges und leicht zugängliches Bildmaterial für die Nutzung innerhalb des Magistrats zur Verfügung zu stellen. Aus diesem Grund wurden hohe Anforderungen an die Georeferenzierung, Vollständigkeit des Straßennetzes, Datenschutz und Webviewer definiert. Die technischen Rahmenbedingungen der Kampagne in der 415 km² umfassenden Millionenstadt umfassen unter anderem ein Straßennetz von 3000 km Länge (Abbildung 1 links), hochauflösende Bilddaten aus bildgebender Sensorik und eine räumliche Abdeckung mit 3 m Bildabstand entlang der Fahrtrichtung. Aus Sicht des Datenschutzes wurde die Unkenntlichmachung von Personen und Kfz-Kennzeichen realisiert. Für die absolute Georeferenzierung der Daten war eine Genauigkeit von 10 cm gefordert. Die dazu benötigten Passpunkte wurden von der MA 41 – Stadtvermessung bereitgestellt. Weiterführend wurde eine geeignete Softwarelösung für die Nutzung der aufgenommenen Bilddaten beauftragt.

Die Ausschreibung und die Erfassung wurde operativ von der MA 41 – Stadtvermessung, dem magistratsinternen Geodatenprovider, geleitet. In einem offenen Verfahren wurde die Firma iNovitas AG (www.inovitas.ch) im Sommer 2017 als Auftragnehmerin für die Mobile Mapping Kampagne bestimmt. Die Firma setzte für die Erfassung des Stadtgebietes zwei idente Vermessungsfahrzeuge ein, die mit Positionierungssensoren, wie einem Global Navigation Satellite System (GNSS) und einer Inertial Measurement Unit (IMU), ausgestattet waren. Das Kamerasetup ist, wie in Abbildung 1 zu sehen, so gewählt, dass für eine weiterführende Auswertung möglichst alle Richtungen abgedeckt sind.



Abb. 1: Links: Wiener Straßennetz mit rund 3000 km Achsen, Rechts: Unterschiedliche Kamerasysteme auf den Mobile Mapping Fahrzeugen (3 Stereokameras, 1 Panoramakamera, 1 Monokamera). Bildrechte: MA41 / Stadt Wien.

Die Kamerasysteme wurden in einer Höhe von rund zwei Metern auf einem stabilen Rahmen am Dach der Vermessungsfahrzeuge montiert und erstmalig vor der Kampagne kalibriert. Auch während der Fahrt überwachte ein Programm mit Computer Vision Ansatz permanent die Kalibrierung, um ein versehentliches Verstellen der Kameras detektieren zu können. Drei RGB Stere-

okameras (ausgerichtet nach vorne und schräg hinten) ermöglichen eine komfortable Messung im 3D Raum per Monoplotting (RGB Bild + Tiefenkarte). Eine RGB Panoramakamera mit einem Blickwinkel von 360° gibt einen guten Überblick über die Situation um das Fahrzeug. Für eine visuelle Abschätzung des Straßenzustandes ist eine hochauflösende RGB Monokamera rückwertig ausgerichtet. Die aufgenommenen Bild- und Navigationsdaten wurden im Fahrzeug auf gesicherten Systemen aufgezeichnet.

Der Fortschritt der Befahrung wurde in einem GIS System im Fahrzeug sichtbar gemacht. Die Fahrzeuge nahmen ihre Daten im normalen Straßenverkehr bei rund 50 km/h Fahrgeschwindigkeit auf, wodurch keine speziellen Absicherungsmaßnahmen nötig waren. Spezialbereiche wie zum Beispiel Fußgängerzonen wurden ebenfalls erfasst, wobei das Fahrzeug durch ein oranges Warnlicht gekennzeichnet war.

2.2 Befahrung 2017-2018

Die Erfassung der Mobile Mapping Daten des Wiener Stadtgebietes startete Ende September 2017 und wurde im Mai 2018 fertig gestellt. In den Wintermonaten wurde die Befahrung wegen des zu niedrigen Sonnenstandes und kurzer Tageslängen unterbrochen. Insgesamt konnte an 70 Befahrungstagen 4200 km öffentlicher Raum digitalisiert werden. Im Vergleich zu den rund 3000 km Straßennetz resultiert die größere Anzahl an Kilometern aus den teilweise mehrfach zu befahrenden Straßenachsen (mehrspurige Fahrbahnen, Gegenfahrbahnen, Abbiegerelationen). In Summe wurden rund 90 Terrabyte an Mobile Mapping Daten und rund 30 Millionen Bilder generiert.

Die Datenaufbereitung der Rohdaten bis zur fertigen Datenbasis erfolgte durch die Auftragnehmerin iNovitas. Parallel zur Befahrung wurde durch die Firma in einer Processing Cloud das Datenmaterial georeferenziert, anonymisiert und für den Bilddienst aufbereitet. Nach dem Datentransfer auf die Server der Stadt Wien führte die MA 41 die Abnahme inklusive einer Qualitätskontrolle durch. Nach erfolgter Datenabnahme wurden die Bilddaten magistratsintern in einer Web-Applikation verfügbar gemacht.

3 Digitaler Bilddienst "Kappazunder"

Um mit den Mobile Mapping Daten sinnvoll arbeiten zu können, wurde der Bilddienst "Kappazunder" in der lokalen IT der Stadt Wien aufgesetzt. Es handelt sich dabei um eine Realisierung der infra3D Technologie aus EUGSTER & BRANDSTÄTTER (2018). Der plattformunabhängige browserbasierte Viewer ermöglicht den Mitarbeiterinnen und Mitarbeitern des Magistrats einen sehr komfortablen Zugriff auf die zentral gelagerten Bilddaten. Die Wartung des Werkzeuges Kappazunder kann ebenfalls zentral erfolgen. Der Name Kappazunder ist ein Wortspiel des Wiener Begriffs "Kapazunder" (=Koryphäe) in Kombination mit dem photogrammetrischen Drehwinkel Kappa.

Die Benutzeroberfläche des Kappazunder ist im Browser bewusst übersichtlich gehalten (siehe Abbildung 2). In einem großen Bildfenster (8) wird das RGB Bildmaterial an der aktuellen Fahrzeugposition dargestellt. Für die Navigation zu einer Adresse steht eine 2D Karte (3) oder eine Adresssuche (2), mit dem behördlichen Adressregister der Stadt Wien, zur Verfügung. Da vor allem auf größeren Straßen wie der Wiener Ringstraße mehrere Fahrspuren erfasst wurden, können diese in einem eigenen Auswahlfenster (6) selektiert werden. In der Kameraauswahl (7) ist der Wechsel zwischen verschiedenen Sensoren möglich. Mit der Videoleiste (4) können die aufgenommenen Bilder wie in einem Video abgespielt werden. Je nach Abspielgeschwindigkeit kann man sich auf diese Art und Weise sehr gut einen virtuellen Überblick über den öffentlichen Raum verschaffen. Die blauen Richtungspfeile (5) im Bildfenster zeigen vorhandene Abbiegemöglichkeiten an. In der Funktionsleiste (1) stehen verschiedene Werkzeuge zur Verfügung. Es gibt zum Beispiel die Möglichkeit photogrammetrische Messungen durchzuführen.



Abb. 2: Aufbau der Kappazunder Oberfläche im Webbrowser. Bildrechte: MA41 / Stadt Wien.

Der Kappazunder ist mit der bestehenden Geodaten-Infrastruktur (GDI) der Stadt Wien verbunden. Somit können Featureklassen eingebunden und dessen Geometrien und Attribute bearbeitet werden. In Abbildung 3 sind als Beispiel die 3D Linien der Mehrzweckkarte (MZK) in grüner Farbe ersichtlich. Die MZK-Daten wurden mit einer klassischen terrestrischen Vermessungsmethode aufgenommen und sind in den Bilddienst per Augmented Reality eingespiegelt. Anhand der Lagegenauigkeit kann die hohe Genauigkeit der Georeferenzierung der Bilddaten im Vergleich zu den vermessenen Linien gut nachvollzogen werden.

Ebenso ist es möglich, neue Featureklassen zu erstellen und für die Erfassung von Geometrien und Attributen zu nutzen. Da die Featureklassen in einer zentralen Geodatenbank gelagert sind, ist der Datenstand für alle berechtigten Nutzerinnen und Nutzer sofort sichtbar. Die zentrale Datenhaltung erlaubt zusätzlich eine gewisse Integrität in Bezug auf Datenqualität und Datensicherheit. Die digitalisierten Geodaten können weiterführend für diverse Anwendungen und GIS Analysen verwendet werden. Vorrangige Ziele für die Nutzung des Kappazunder sind die Möglichkeit von digitalen Begehungen (Lokalaugenscheinen) eines Ortes und eine effiziente Auswertung von 3D Objekten (Punkte, Linien, Flächen) durch Interpretation von Merkmalen in den Bilddaten (STRONDL et al. 2018).



Abb. 3: Linien der Mehrzweckkarte der Stadt Wien per Augmented Reality in den Bilddienst in grüner Farbe eingespiegelt. Die gute Georeferenzierung der Bilder wird im Vergleich zu den vermessenen Linien sichtbar. Bildrechte: MA41 / Stadt Wien.

4 Auswertungsprojekte

Im Rahmen des Programmes Wien gibt Raum haben sich in einer frühen Programmphase zwei unterschiedliche Pilotprojekte angeboten, wobei in beiden Fällen eine Objekt-Inventarisierung auf Basis des Kappazunder durchgeführt wird. In Zusammenarbeit mit den Magistratsabteilungen MA19 (Architektur und Stadtgestaltung), MA28 (Straßenverwaltung und Straßenbau) und MA46 (Verkehrsorganisation und technische Verkehrsangelegenheiten) wurde evaluiert, ob eine Mobile Mapping basierte Inventarisierung der Objekte technisch, zeitlich und wirtschaftlich durchführbar ist. Basierend auf Stichproben konnte bestätigt werden, dass die Fragestellung der Pilotprojekte durch die vorhandenen großflächigen Mobile Mapping Daten effizient und qualitativ hochwertig gelöst werden können. Des Weiteren zeigt sich, dass durch diese Technologie eine wirtschaftlich sinnvolle Bearbeitung der Fragestellung erstmals möglich ist. Nachfolgend werden die beiden Pilotversuche näher beschrieben.

4.1 Aufgabenstellung

Obwohl die beiden Pilotprojekte sich in ihren Aufgabenstellungen unterscheiden, behandeln beide Aktivitäten die Lösung von zeitkritischen Fragestellungen im öffentlichen Raum. Ziel ist es, bestimmte Objekte im Bilddienst Kappazunder mit Koordinaten zu erfassen, mit Attributen zu erweitern und direkt in der GDI der Stadt Wien verfügbar zu machen. Die räumliche Ausdehnung bei beiden Projekte beinhaltet den kompletten Straßenbereich der Stadt Wien.

4.1.1 Pilotprojekt-A: Werbeträger

Aufgabenstellung ist eine komplette Inventarisierung von verschiedenen Werbeträgertypen im öffentlichen Raum unter knappem Zeitbudget. Die Daten (Geometrie und Semantik der Werbeträger) dienen als wichtige Grundlage für die Erstellung eines neuen Werbekonzeptes der Stadt Wien. Der Bedarf eines Konzeptes ergibt sich durch rechtliche Rahmenbedingungen, die eine Neuregelung der auf Werbung bezogenen Genehmigungen nötig machen. Das Konzept dient letztgültig auch als Entscheidungsgrundlage für politische Entscheidungsträger. In allen Punkten spielen die mittels Kappazunder erhobenen Geodaten eine wichtige Rolle. Herausfordernd sind bei dieser Aufgabenstellung die große Anzahl an verschiedenen Werbeträgertypen sowie der mögliche Zeitrahmen.

4.1.2 Pilotprojekt-B: Verkehrszeichen und Wegweiser

Im Unterschied zu Pilotprojekt-A gibt es bei diesem Projekt bereits bestehende Daten aus der Verwaltungshistorie. Dieser Datenstand weist jedoch eine ungenaue Verortungs-Qualität auf. Ziel in Pilotprojekt-B ist eine großflächige Inventarisierung der rund 100.000 Standorte von Verkehrszeichen und Wegweisern, wobei speziell die Verknüpfung mit den bestehenden Daten aus einem Verkehrs-Management-System (VMS) herausfordernd ist. Die Verknüpfung mit dem Altbestand ist wichtig, da rechtliche Aspekte wie zum Beispiel Bewilligungen und Bescheide an den Altbestand geknüpft sind. Aus diesem Grund liegt der Fokus auf eine möglichst vollständige Zuordnung der bestehenden Altdaten zu den in der Natur tatsächlich vorhandenen Positionen.

Da für die Verwaltung der Objekte im VMS neben Sachdaten und Geometrien auch Bilder der jeweiligen Standorte benötigt werden, sollen zusätzlich zur Erfassung der Position auch Bildausschnitte des jeweiligen Objektes angefertigt werden. Unter Zuhilfenahme des Bilddienstes sollen zum einen die Bildausschnitte erstellt und zum anderen eine Qualitätssteigerung der Verortungs-Qualität erzielt werden.

4.2 Objektkatalog

Voraussetzung für eine qualitativ hochwertige Auswertung der Objekte ist ein detaillierter Objektkatalog. Darin werden die zu erfassenden Objekte aufgelistet und detailliert beschrieben. Die Selektion der geforderten Objekte, die Generalisierung und Interpretation der Geometrien und der zu erfassenden Attribute muss klar dargelegt werden. Werden verschiedene Geometriearten verwendet, ist eine klar nachvollziehbare Layerstruktur zu definieren. Aus dem Objektkatalog sollte demnach auch ersichtlich sein, welche Objekte nicht Teil der Auswertung sind. Ziel ist es, dass während einer laufenden Auswertung bei der jeweiligen dateninterpretierenden Person keine Fragen über die Objektkategorie entstehen. Der jeweilige Objektkatalog stellt somit eine verbindliche Grundlage für die Digitalisierung dar.

Gemeinsam mit den zuständigen Dienststellen wurde in einem iterativen Prozess geklärt, welche Objekte für die Auswertung relevant sind und welche nicht. Im Fall der Werbeträger musste abgeklärt werden, bei welchen Objekten es sich um Werbung im öffentlichen Raum handelt oder wie diese abzugrenzen sind. Die Werbeträger wurden deshalb in 20 verschiedene Gruppen wie z.B. Litfaßsäulen, Plakatwände oder Rollingboards eingeteilt. Um die jeweilige Gruppe bestmöglich definieren zu können, helfen Beispielbilder wie sie in Abbildung 4 zu sehen sind.



Abb. 4: Ausschnitt aus dem Objektkatalog für Werbeträger. Definiert sind die Gruppenzugehörigkeit und die Geometrie. Bildrechte: MA41 / Stadt Wien.

Beim Objektkatalog für das Pilotprojekt-B lag der Fokus auf der Definition von offiziell im Verwaltungsbereich der Stadt Wien liegenden Verkehrszeichen und Wegweisern. Hier wird auch genau geregelt, wie die Verknüpfung mit den bestehenden Daten im VMS herzustellen ist und wie die Bildausschnitte der Objekte auszusehen haben.

4.3 Objekterfassung im Bilddienst

Aufgrund der Komplexität der Aufgaben und Objekte wurden semi-automatische Erfassungsprozesse gewählt. Mehrere Personen in der MA 41 führen im Bilddienst Kappazunder auf Grundlage des Objektkataloges die 3D-Objekterfassung und Verknüpfung mit dem VMS durch. Die Aufgabe besteht darin, im Bilddienst virtuell das Straßennetz abzufahren, und alle sichtbaren und laut Objektkatalog relevanten Objekte mit den Digitalisierungswerkzeugen zu erfassen. Die Gebietsaufteilung für die Auswerterinnen und Auswerter erfolgt dabei Bezirksweise. Die abgearbeiteten Straßenachsen werden in GIS Layern nachgeführt, um jederzeit den Fortschritt der Auswertung visuell überwachen und berechnen zu können. Dies wird speziell aus wirtschaftlicher Sicht und zu Zwecken der Projektkommunikation benötigt.

Das Ergebnis der Bearbeitung sind qualitativ hochwertige 3D Geodaten, die zentral in der GDI der Stadt Wien abgelegt sind.

4.3.1 Ablauf und Ergebnis der Werbeträger

Die Auswertung der Werbeträger in Pilotprojekt-A konnte zeitgerecht abgeschlossen werden. Weiterführend wurden basierend auf diesen Daten unter anderem politische Entscheidungen und Vergabeprozesse durchgeführt. Innerhalb von nur zwei Monaten konnten rund 32.000 Objekte im Kappazunder inventarisiert werden. In Abbildung 5 ist eine exemplarische Szene des Ergebnisses zu sehen. In der linken Bildhälfte sieht man die Szene im Kappazunder aus Fahrzeugsicht, und in der rechten Bildhälfte sieht man dieselbe Szene im Orthophoto. Die grünen Linien entsprechen der Oberkante der inventarisierten Plakatwände und werden dem zugehörigen Mobile

Mapping Bild und dem Orthophoto überlagert. Anhand der Lage der Linien zeigt sich eine sehr gute Verortung der Daten, wobei in der Regel Genauigkeiten besser als 10cm erzielt werden können. Für jedes digitalisierte Feature wurden in der Datenbank Metadaten abgelegt, die in einem Abfragefenster ausgegeben werden können (Mitte linke Bildhälfte).



Abb. 5: Ergebnis der Werbeträger Digitalisierung. Rechts: Grüne Linien der Plakatwände, Links: 2D Ansicht mit dem Orthophoto im Hintergrund. Jede Geometrie ist mit Metadaten hinterlegt und in der GDI gespeichert. Bildrechte: MA41 / Stadt Wien.

Auf Grund der erfolgreichen Auswertung kann die Stadt Wien ab sofort auf eine hervorragende Datengrundlage aller sichtbaren Werbeträger zurückgreifen. Die Verortung und Klassifizierung der erfassten Objekte in unterschiedliche Gruppen des Objektkataloges konnte anhand der Bilddaten durchgeführt werden. Als Qualitätskontrolle wurde eine Stichprobe von 10% der erfassten Objekte manuell durch mehrere Personen überprüft, wobei die geprüften Objekte räumlich zufällig im Stadtgebiet verteilt sind. In Summe zeigt sich, dass 95% der Daten korrekt verortet und klassifiziert wurden. Die Fehlerquote von 5% lässt sich auf eine Unschärfe des Objektkataloges und auf Auswertefehler (z.B. Objekt nicht gefunden oder falsch klassifiziert) zurückführen. Für Pilotprojekt-A ermöglicht die gewählte Methodik somit zufriedenstellende und weitgehend vollständige Ergebnisse.

Eine Aussage über die Eigentümer der Werbeträger (also ein rechtlicher Aspekt) lässt sich anhand der Bilddaten nur bedingt realisieren. Sind die Werbeträger mit dem Logo des Eigentümers beschriftet und lässt es die Auflösung der Bilddaten zu, kann der Eigentümer visuell festgestellt werden. Eine rechtsverbindliche Aussage über den Eigentümer ist jedoch nur unter Zuhilfenahme der zugehörigen Bescheide oder durch eine anlassbezogene Ermittlung möglich. Die komplexe Fragestellung, auf welchen Grundstücken sich die Werbeträger befinden, ist mit den Bilddaten alleine nicht lösbar. Die erfassten Geodaten werden in weiteren Bearbeitungsschritten analysiert, wobei speziell der Verschnitt mit Polygonen der Stadt Wien Grundstücke von Interesse

ist. An dieser Stelle wird auch klar, wie wichtig eine gute Georeferenzierung der Mobile Mapping Daten und der daraus resultierenden Geodaten ist.

Mit den Grundlagendaten der Werbeträger können verschiedene Analysen durchgeführt werden. In der Dichtekarte in Abbildung 6 sind zum Beispiel Bereiche mit hoher Werbungsdichte sehr gut zu erkennen. Vor allem bei stark frequentierten Straßen wie zum Beispiel der Ringstraße oder dem Gürtel bestätigt sich die visuelle Informationsüberflutung in der Naturszene. In der in Abbildung 6 dargestellten Karte werden nur Gruppen des Objektkataloges dargestellt. Werbeschilder an Fassaden, wie sie bei Geschäften oft vorkommen, waren nicht Teil der Auswertung. Aus diesem Grund sticht zum Beispiel die belebte Einkaufsstraße, die Mariahilfer Straße (grüner Bereich in der Abbildung), nicht heraus.



Abb. 6: Ausschnitt einer Dichtekarte der Werbeträger im Straßennetz Wien. Je heller der Bereich, desto höher ist die Dichte der Werbung. Die grüne Markierung kennzeichnet die Mariahilfer Straße. Bildrechte: MA41 / Stadt Wien.

4.3.2 Ablauf und Ergebnis der Verkehrszeichen und Wegweiser

Die Auswertung der Standorte von Verkehrszeichen stellt eine herausfordernde Aufgabe dar. Altdaten und bestehende Geometrien müssen mit den in der Natur tatsächlich vorhandenen Positionen zusammengeführt werden. Auf Grund historischer Gegebenheiten wurden die Standorte der Verkehrszeichen bis vor kurzem nur mit Adressbezug verortet. Da Adressen in vielen Fällen keinen eindeutigen Raumbezug darstellen, und vor allem in Kreuzungsbereichen oder bei langen Grundstücken keine eindeutige Positionsangabe von Verkehrszeichen-Standorten ermöglichen, haben die in den Altdaten vorhandenen Standorte oftmals dieselbe Adresse und somit Position. Unter Zuhilfenahme der Mobile Mapping Daten wird nun eine Verknüpfung der tatsächlich vorhandenen Verkehrszeichen-Standorte durchgeführt, wobei die Bilddaten des Kappazunder mit den Informationen des VMS kombiniert werden. Diese herausfordernde Aufgabe ist nur bedingt automatisierbar, wodurch aus Qualitäts- und Zeitgründen ein manueller Ansatz gewählt wurde. Die Anzahl der Schilder sowie der zugehörigen Zusatztexte auf einem Steher müssen zusammenpassen, um ein Objektmatching durchführen zu können. Ist dies der Fall, wird eine eindeutige Verkehrszeichen-Standortnummer aus dem VMS als Attribut der Punktgeometrie im Kappazunder hinzugefügt. Da die Altdaten initial stark lokal geclustert sind, stellt diese Aufgabe eine hohe Anforderung an die auswertenden Personen.

Über eine Schnittstelle werden die im Kappazunder erfassten und verknüpften Geometrien mit dem VMS System ausgetauscht, wobei die Position der jeweiligen Standortnummer richtig gestellt wird. Praktisch bedeutet dies, dass die oftmals stark geclusterten Kreuzungsbereiche Schritt für Schritt korrigiert und entwirrt werden. In Abbildung 7 ist auf der linken Abbildung die erfasste Position eines Stehers im Kappazunder mit den zugehörigen Attributen ersichtlich, wobei der grüne Punkt die generalisierte Geometrie des Standortes darstellt. Auf der rechten Seite der Abbildung ist im VMS der bestehende Datensatz samt Visualisierung bereits an die korrekte Position gewandert. Zuvor war die alte Position in der Mitte des Gebäudes, da dort der Zentroid des Adressbezuges liegt. Je nach Gebäudedimension können die Schiebevektoren zwischen Alt und Neu in einem Bereich von 10 Meter und mehr liegen.



Abb. 7: Ausschnitt Ergebnis der Verkehrszeichen und Wegweiser Digitalisierung. Rechts: die Position im Kappazunder in grün, Links: Korrigierte Positionen im VMS in 2D. Bildrechte: MA41 / Stadt Wien.

Ist keine Zuordnung eines Objektes zu einer bestehenden Standortnummer des VMS möglich, wird für das neu gefundene Objekt automatisiert eine neue Standortnummer vergeben und im VMS neu angelegt. In diesem Fall gibt es per Definition keine zugehörigen Altdaten, wodurch das neue Objekt im Laufe der Zeit im VMS System überarbeitet und mit einer entsprechenden Genehmigung versehen werden muss. Aus dem Monitoring der aktuell noch laufenden Datenbearbeitung geht hervor, dass die Quote der nicht zuordenbaren Standorte bei ca. 18 Prozent liegt. Auf Grund Komplexität der Aufgabenstellung sowie der hohen Fluktuation und der bestehenden

Historie bei Verkehrszeichen und Wegweisern ist dieser Wert nicht verwunderlich. Nach Rücksprache mit Experten wird dies als ein gutes Ergebnis interpretiert.

Mit der dargelegten Methodik konnten in einem Zeitraum von nur vier Monaten rund 30.000 Standorte bearbeitet und korrigiert werden. In Abbildung 8 ist ein Ausschnitt der Wiener Innenstadt zu sehen, wobei über die Hintergrundkarte die korrigierten Standorte als blaue Punkte überlagert sind. Beeindruckend ist, dass sich alleine im ersten Wiener Gemeindebezirk rund 3.700 Verkehrszeichen-Standorte auf einer Fläche von 2,8 km² befinden.



Abb. 8: Standort-Karte der Wiener Innenstadt: Über die Hintergrundkarte sind die korrigierten Standorte der Verkehrszeichen als blaue Punkte überlagert. Bildrechte: MA41 / Stadt Wien.

5 Fazit & Ausblick

Mit der Mobile Mapping Kampagne des Programmes Wien gibt Raum wurde ein neuer Geobasisdatensatz für die Stadt Wien geschaffen, der für verschiedenste Aufgabenstellungen herangezogen werden kann. Auf Grund der einfachen Ausprägung des Bilddienstes Kappazunder als Webviewer können die Daten an einen großen Personenkreis ausgerollt werden. Anbindungsmöglichkeiten an eine bestehende GDI erlauben die Bearbeitung von neuen und bestehenden Datensätzen. Das bisher enorme positive Echo der ersten rund 600 Nutzerinnen und Nutzern des Bilddienstes deutet auf eine zukünftige umfangreiche Nutzung innerhalb des Magistrats der Stadt Wien hin. Von einer einfachen Betrachtung der Bilder bis hin zu komplexeren Bearbeitungen
werden mehrere Möglichkeiten offeriert. Um die Aktualität der Mobile Mapping Daten zu gewährleisten, sind weitere Befahrungen in regelmäßigen Abständen geplant. Durch die Vermessungsfahrzeuge beschränkt sich der Einsatzbereich aktuell auf das Wiener Straßen- und Wegenetz. Da es aber bereits Wünsche für eine Ausdehnung auf weitere Bereiche wie Parks, Radwege oder die Donauinsel gibt, setzt sich die MA 41 auch mit alternativen Formen mobiler Datenerfassung auseinander. Das Potential bildgebender Verfahren wird aus Sicht der Geodäsie als sehr hoch gesehen.

Die Ergebnisse der beiden Pilotprojekte zeigen, dass die gewählte Mobile Mapping Methodik qualitativ zufriedenstellende und weitgehend vollständige Ergebnisse ermöglicht. Im Bereich Werbeträger konnte eine Qualität von 95% korrekt erfasster Objekte erreicht werden. Im Bereich Verkehrszeichen liegt die Zuordnungsrate bei guten 82%. Durch die komplexen Aufgabenstellungen sind, vor allem bei der Fusion mit Altdaten, mehrheitlich noch manuelle oder semiautomatische Arbeitsschritte notwendig. Für zukünftige Projekte ist eine Erweiterung der Merkmalsextraktion aus Bilddaten, aufbauend auf einem Deep-Learning Ansatz, angedacht. Studien im Bereich Mobile Laserscanning zeigen, dass speziell Geometrien gut aus Punktwolken ableitbar sind, wobei jedoch die Integration von Bildern als großer Mehrwert gesehen wird (JAAKKO-LA et al. 2008; ZHENWEI et al. 2018). Deep Learning Methoden und die Integration von Laserscanning können möglicherweise in Zukunft die Erfassung von einfachen Objekten wie zum Beispiel Kanaldeckel oder Masten unterstützen. Komplexe Objekte, dessen Erscheinungsbilder stark variabel sind, werden für die neuen Technologien jedoch eine Herausforderung sein. Tiefgehendes Verständnis über die Stärken neuer Ansätze wird über praktische Tests gefunden werden müssen. Jedenfalls ist das Potential von hochauflösenden Bilddaten mit Raumbezug enorm, und wird voraussichtlich so manchen Bereich einer modernen Stadtverwaltung verändern.

6 Literaturverzeichnis

- STRONDL, M., VLADAR, D., DÜRAUER, S., EYSN, L., FALKNER, J., JAKLI, T. & OBLIN, A., 2018: Wien gibt Raum: Vom Konzept einer innovativen Verwaltung des öffentlichen Raums zum Mobile Mapping einer Großstadt. VGI: Österreichische Zeitschrift für Vermessung & Geoinformation, 1/2018, Tagungsband 13. Geodätentag, 44-52.
- EUGSTER, H. & BRANDSTÄTTER, W., 2018: Bildbasierte 3D-Geo-Strassen- und Schienen-Webdienste als Basis für ein Infrastrukturmanagement 4.0 – Technologie und Anwendungsmöglichkeiten von infra3D. VGI: Österreichische Zeitschrift für Vermessung & Geoinformation, 1/2018, Tagungsband 13. Geodätentag, 53-57.
- JAAKKOLA, A., HYYPPÄ, J., HYYPPÄ, H., KUKKO, A., 2008: Retrieval Algorithms for Road Surface Modelling Using Laser-Based Mobile Mapping. Sensors, 8(9), 5238-5249, doi:10.3390/s8095238, http://www.mdpi.com/1424-8220/8/9/5238.
- ZHENWEI, S., ZHIHONG, K., YI, L., YU, L., WIE, C., 2018: Automatic Recognition of Pole-Like Objects from Mobile Laser Scanning Point Clouds. Remote Sensing, 10(12), http://www.mdpi.com/2072-4292/10/12/1891.

Kombination von terrestrischem Laserscanning, UAV- und Nahbereichsphotogrammetrie zur Erfassung komplexer Kirchenbauwerke in Georgien

THOMAS LUHMANN¹, MARIA CHIZHOVA², DENYS GORKOVCHUK³, ANNIKA JEPPING¹ & HEIDI HASTEDT¹

Zusammenfassung: Im Rahmen einer Messkampagne zur dreidimensionalen Erfassung mehrerer historischer Kirchen in Tbilisi (Georgien) wurden im September 2018 photogrammetrische Aufnahmen und terrestrische Laserscans durchgeführt. Ziel war die vollständige räumliche Rekonstruktion mit einer Ortsauflösung und Genauigkeit von ca. 1cm unter teilweise erschwerten äußeren Bedingungen, die den Einsatz unterschiedlicher Messtechniken erforderten.

Die örtlichen Messdaten wurden durch jeweils zwei Laserscanning-Kampagnen (Leica BLK360 und Faro Focus 3D X330), zwei UAV-Flüge sowie zwei terrestrischen Bildverbänden (Canon D6 und Canon 200D) erfasst. Die photogrammetrischen Punktwolken wurden mit den SfM-Programmen Agisoft PhotoScan und RealityCapture unter Berücksichtigung der Passpunkte aus dem Faro-Laserscan berechnet. Die mittleren Restfehler aus den Registrierungen bzw. photogrammetrischen Auswertungen liegen bei 4-12 mm, abhängig von der gewählten Softwarelösung.

Im Beitrag werden erste Ergebnisse von einem der Messobjekte hinsichtlich Registrierungsgenauigkeit bzw. Genauigkeit der Bündelausgleichung analysiert. Die mit den einzelnen Verfahren berechneten Punktwolken werden verglichen und daraufhin untersucht, inwieweit eine Fusion der Punktwolken zur Vervollständigung von Lücken und Abschattungen genutzt werden kann.

1 Einleitung

1.1 Motivation

Im Rahmen einer Messkampagne zur dreidimensionalen Erfassung mehrerer historischer Kirchen in Tbilisi (Georgien) wurden im September 2018 photogrammetrische Aufnahmen und terrestrische Laserscans durchgeführt. Dabei wurden vorab mehrere Projektziele definiert:

- 3D-Rekonstruktion der Außenhaut der Bauwerke
- 3D-Rekonstruktion eines exemplarischen Innenraums
- Einsatz und Untersuchung des neuen Leica BLK360 Scanners unter Praxisbedingungen
- Kombination von UAV- und terrestrischen Bildaufnahmen mit terrestrischen Laserscans
- Nutzung der Daten für zukünftige Untersuchungen zur semantischem Segmentierung
- Nutzung der Daten für Restaurierungsarbeiten und touristische Zwecke

¹ Jade Hochschule Oldenburg, Institut für Angewandte Photogrammetrie und Geoinformatik, Ofener Str. 16, D-26121 Oldenburg, E-Mail: [Thomas.Luhmann, Annika.Jepping, Heidi.Hastedt]@jade-hs.de

² Otto-Friedrich-Universität Bamberg, Lehrstuhl für Digitale Denkmaltechnologien, Am Zwinger 4, D-96047 Bamberg, E-Mail: maria.chizhova@uni-bamberg.de

³ Kiev National University for Construction and Architecture, E-Mail: denys.gorkovchuk@spm3d.com

Im Rahmen einer laufenden Forschungsarbeit zur modellbasierten Rekonstruktion von zerstörten Bauwerken aus Punktwolken russisch-orthodoxer Kirchen (CHIZHOVA et al. 2016) entstand eine Zusammenarbeit der Jade Hochschule in Oldenburg und der Universität Bamberg mit der Georgian Technical University in Tbilisi. In einer einwöchigen Messkampagne wurden dazu insgesamt drei Kirchen mit verschiedenen Messsystemen aufgenommen:

- Terrestrischer Laserscanner Faro Focus 3D X330
- Terrestrischer Laserscanner Leica BLK360
- UAV DJI Mavic Pro mit FC 220 f=4.7mm
- DSLR-Kamera Canon EOS 200D, f=10-20mm
- DSLR-Kamera Canon EOS D6 Mark II, f=24-50mm

Ziel war die vollständige räumliche Rekonstruktion mit einer Ortsauflösung und Genauigkeit von ca. 1 cm unter teilweise erschwerten äußeren Bedingungen, die die Kombination unterschiedlicher Messtechniken erforderten. Dazu werden zunächst einzelne Gesamtpunktwolken aus den Laserscans und den photogrammetrischen Aufnahmen erzeugt. Die photogrammetrischen Auswertungen werden zum einen mit Agisoft PhotoScan, zum anderen mit RealityCapture durchgeführt. In einem anschließenden Fusionsschritt werden die Punktwolken in ein Gesamtmodell zusammengeführt.

Aufgrund logistischer und organisatorischer Gründe konnte keine detaillierte Vorplanung durchgeführt werden. Die erstmalige Ortsbegehung fand am ersten Tag der Kampagne statt. Auf die Auswahl der durch die lokalen Partner bereitgestellten Messgeräte (Faro Scanner, DJI Drohne) bestand keine Einflussmöglichkeit. Messtechnische Hilfsmittel mussten zum Teil kurzfristig vor Ort organisiert werden, ebenso wie Zugangsberechtigungen für die religiösen Stätten.

Eine tachymetrische Passpunktmessung konnte aus technischen Gründen nicht durchgeführt werden. Es wurde daher entschieden, Passpunkte für die photogrammetrische Orientierung aus den registrierten Gesamtpunktwolken des Laserscannings zu entnehmen.

Im folgenden Beitrag werden die einzelnen Messobjekte kurz vorgestellt. Die Untersuchungen beschränken sich jedoch zunächst auf die Andreas-Kirche im Lurdji-Kloster. Die Auswertung der anderen Messobjekte erfolgt in naher Zukunft.

Untersuchungen ähnlicher Art liegen aus der Literatur vor, beziehen sich jedoch auf andere verfügbare hard- und softwaretechnische Lösungen. Beispiele geben KERSTEN et al. (2015), NAUMANN et al. (2015) oder REMONDINO (2011).

1.2 Messobjekte

Ursprünglich geplant waren die Aufnahmen der georgischen Kirchenanlagen Lurdji und Sioni in der georgischen Hauptstadt Tbilisi (Tiflis). Aufgrund der räumlichen Nachbarschaft zur Andreas-Kirche in Lurdji wurde kurzfristig entschieden, auch die benachbarte russisch-orthodoxe Johannes-Kirche zu erfassen. Beschreibungen zu den Objekten finden sich in ANCHABADZE & VOLKOVA (1990).

1.2.1 Kloster Lurdji

Das Lurdji-Kloster ist ein orthodoxer Architekturkomplex in Tbilisi (Abb. 1). Das Kloster wurde im 7. Jhd. begründet, die ältesten Teile lassen sich auf das 12. Jhd. datieren, was der Regierungs-

zeit von Königin Tamar entspricht. Im 16. Jhd. wurde das Kloster von Persern zerstört und erst ein Jahrhundert später restauriert. Aktuell besteht der Sakralkomplex aus zwei Kirchen und dem dazu gehörigen historische Park Vera. Nach der Zerstörung wurde die Andreas-Kirche im 17. Jhd. als Basilika mit dreiteiligem Kirchenschiff restauriert. Ende des 19. Jhd. wurde die Kirche nach einem Projekt von A. Chizhov rekonstruiert und hat eine neue kugelförmige Kuppel bekommen, was für die georgische Architektur nicht typisch war. Diese Kuppel wurde 1995 durch eine klassische georgische Kegelkuppel ersetzt. Das früher mit blauen Keramikdachziegeln bedeckte Dach besteht heute aus blaugrau lackierten Zinkblechen. Die Dachgestaltung hat den Namen des Klosters bestimmt: "lurdji" bedeutet blau. Aufgrund der verschiedenen Umbauten liegt die Kirche nicht im typischen Stil georgischer Kirchen mit kreuzförmigen Grundriss vor, sondern in einer rechteckigen Grundform mit Kreuzdach und zentrischem Turm, der einen zylindrischen Sockel und eine kegelförmige Spitze besitzt.

Die Kirche hat eine Größe von ca. 17 m x 12 m x 25 m (B x L x H). Sie befindet sich inmitten eines bebauten Wohngebiets und der am Rand liegenden Parkanlage, so dass teilweise eine sehr eingeschränkte Zugänglichkeit vom Boden aus gegeben war. Das Material der Außenmauern besteht aus Sandstein bzw. Ziegelsteinen und besitzt eine natürliche Rauigkeit sowie ausgebrochene Stellen, die durch Witterung und Abnutzung entstanden sind. Die Dachflächen sind weitgehend texturarm.

Für das Objekt wurde die Aufnahme der Außenhaut festgelegt. Dazu wurden zwei UAV-Flüge, zwei terrestrische DSLR-Bildverbände sowie zwei Laserscanmessungen durchgeführt.



Abb. 1: Andreas-Kirche im Kloster Lurdji

1.2.2 Weitere Kirchenbauwerke

Während der Messkampagne wurde zusätzlich die Johannes-Kirche, die sich ebenfalls im Lurdji-Komplex befindet, aufgenommen (Abb. 2 links). Sie wurde als klassische russische Kirche mit Schiff-Konstruktion und fünf Zwiebelkuppeln im Zeitraum 1898-1901 durch den russischen Statthalter G. Golitsyn erbaut. T. Luhmann, M. Chizhova, D. Gorkovchuk, A. Jepping & H. Hastedt



Abb. 2: Links: Russische Kirche im Kloster Lurdji; vorne: Leica BLK360; hinten auf Stativ: Faro Scanner Rechts: Sioni-Kathedrale

Aufgrund der erheblichen Einschränkungen in der Zugänglichkeit wurde von diesem Objekt lediglich ein Laserscan mit dem Leica BLK360 aufgenommen. Eine photogrammetrische Aufnahme konnte nicht realisiert werden.

Die Sioni-Kathedrale ist die historische Hauptkirche der Stadt Tbilisi und eines der wichtigsten Zentren der georgischen Orthodoxie (Abb. 2 rechts). Der Bau der ursprünglichen Kirche begann im 6. Jhd. durch den georgischen König Wachtang I. Gorgassali, ihr endgültiger Aufbau ist auf das 7. Jhd. datiert. Diese erste Kirche wurde nach der Einrichtung des Tiflis-Emirates vollständig zerstört. Die neue Sioni-Kathedrale wurde im 12. Jhd. aufgebaut und hat mehrmals Zerstörungen und Rekonstruktionen erlebt (17. und 18. Jhd.). Nach der Restaurierung 1980-1983 hat die Kirche ihre mittelalterliche Gestalt behalten und entspricht stilistisch der klassischen prämongolischen Sakralarchitektur Georgiens.

Das Objekt wurde sowohl von außen als auch von innen aufgenommen. Insgesamt wurden wiederum zwei UAV-Flüge sowie jeweils zwei terrestrische DSLR-Bildverbände und zwei Laserscanmessungen im Innen- und Außenraum durchgeführt.

2 Laserscanning

2.1 Signalisierung

Alle aufgenommenen Objekte wurden mit Zielmarken ausgestattet, die in ca. 10 m Abstand zueinander am Gebäude sowie in der rückwärtigen Umgebung angebracht wurden (Übersicht in Abb. 3). Wie zuvor erwähnt, konnten die Zielmarken nicht tachymetrisch eingemessen werden, sondern dienten für das Laserscanning ausschließlich als Verknüpfungspunkte, deren 3D-Koordinaten anschließend für die photogrammetrische Datumsfestlegung verwendet werden. Die Zielmarken bestanden aus den üblichen schachbrettartigen Signalen (A4-Größe). Zusätzlich konnten einige Profi-Targets (Tilt & Turn Targets) eingesetzt werden, die auf Stativen ohne Änderung des Markenzentrums um die eigene Achse gedreht werden, um sie von verschiedenen Standpunkten und aus der Luft optimal messen zu können.

2.2 Faro Focus 3D X330

2.2.1 Übersicht

Der Faro Focus 3D X330 (Abb. 2 im Hintergrund) ist ein geodätischer Laserscanner mit Phasenvergleichsverfahren und einer 3D-Punktgenauigkeit von ca. 2 mm in 10 m Entfernung. Die maximale Reichweite beträgt 330 m. Für das Laserscanning wurde ein Punktabstand von 6 mm in 10 m Entfernung gewählt. In diesem Modus benötigt ein vollständiger Scan ca. 25 Minuten, inklusive der zusätzlich aufgenommenen Bilddaten. Eine Übersicht der erfassten Daten kann Tab. 1 entnommen werden.

	Standpunkte	3D-Punkte	Scandauer	Mittl. Auflö-	Reg.präzision	Reg.präzision
				sung	Register360	Cyclone
Faro	11	350 Mio.	4.5 h	3 – 5 mm	6 mm	4 mm
BLK	20	784 Mio.	1.5 h	4 – 8 mm	9 mm	4 mm

Tab. 1: Laserscanning-Daten Lurdji Andreas

Neben 9 bodennahen Standpunkten wurden zwei Scans von erhöhten Standpunkten aus dem benachbarten Glockenturm sowie von einer Dachterrasse eines anderen Gebäudes aufgenommen (Abb. 2). Aufgrund der eingeschränkten Standortwahl konnten einige Teile des Daches sowie des Turmkegels nicht erfasst werden. Die Punktauflösung am Objekt schwankt bedingt durch sehr unterschiedliche Messdistanzen zwischen ca. 1 mm und 30 mm (Turm).



Abb. 3: Übersicht der registrierten Faro-Gesamtpunktwolke mit Verteilung der Standpunkte (rot) und Zielmarken (Papiermarken grün, Profitargets blau)

2.2.2 Registrierung

Vor Ort wurde mit dem Programm Faro Scene eine erste Registrierung aller einzelnen Punktwolken durchgeführt, um die Güte und Vollständigkeit der Daten zu überprüfen. Später wurden die Daten mit Register 360 registriert, das die Extraktion der 3D-Koordinaten der automatisch gemessenen Zielmarken erlaubt. Die Registrierung erfolgte über die abgebildeten Zielmarken mit anschließender Optimierung (ICP). Die mittlere Standardabweichung der Gesamtregistrierung beträgt dabei 6 mm. Bei einer Registrierung mit der Software Cyclone verbessert sich die Präzision der Registrierung auf 4 mm. Die mittlere Punktauflösung liegt zwischen 3 m in bodennahen Bereichen und 5 mm am Turm. Abb. 3 stellt eine Übersicht der registrierten Punktwolke mit den Standpunkten und den erfassten Zielmarken dar.

2.3 Leica BLK360

2.3.1 Übersicht

Der seit 2017 auf dem Markt erhältliche Laserscanner Leica BLK360 (Abb. 2 im Vordergrund) zeichnet sich vor allem durch seine Handlichkeit aus. Das Gerät ist vorwiegend für einfache Gebäudeaufnahmen mit mittleren Genauigkeitsansprüchen bei einer Reichweite bis 60 m konzipiert. Erste Untersuchungen zur Leistungsfähigkeit zeigen BLASKOW et al. (2018), die die vom Hersteller angegebene 3D-Punktgenauigkeit von ca. 6 mm in 10 m Entfernung bestätigen. Die Distanzmessung erfolgt nach dem WFD-Prinzip (wave form digitizer). Das Laserscanning mit dem BLK wurde im Betriebsmodus High Density vorgenommen, der einem Punktabstand von 6 mm in 10 m Entfernung entspricht. In diesem Modus benötigt ein vollständiger Scan inklusive der Aufnahme von Panoramabildern ca. 5 Minuten. Eine Übersicht der erfassten Daten kann Tab. 1 entnommen werden.

Beim BLK werden die erfassten Daten zunächst im Scanner gespeichert und anschließend per WLAN auf einen Tablet-Computer übertragen.



Abb. 4: Übersicht der registrierten BLK-Gesamtpunktwolke mit Verteilung der Standpunkte (rot) und Zielmarken (Papiermarken grün, Profitargets blau)

2.3.2 Registrierung

Auch für die BLK-Scans wurde noch vor Ort eine erste Registrierung mit dem Programm ReCap durchgeführt. Die spätere Gesamtregistrierung mit Register 360 ergab eine mittlere Standardabweichung von 9 mm. Die Registrierung erfolgte über die abgebildeten Zielmarken mit anschließender Optimierung (ICP). Die mittlere Punktauflösung liegt zwischen 4 mm in bodennahen Bereichen und 8 mm am Turm. Bei einer Registrierung mit der Software Cyclone verbessert sich auch hier die Präzision der Registrierung auf 4 mm. Abb. 4 stellt eine Übersicht der registrierten Punktwolke mit den Standpunkten und den erfassten Zielmarken dar.

2.4 Vergleichende Analyse

2.4.1 Registrierung

Beide untersuchten Scanner konnten problemlos in einem praktischen Projekt eingesetzt werden. Auch wenn der BLK360 eher als Scanner für Innenräume oder einfache BIM-Anwendungen mit geringeren Genauigkeitsanforderungen eingestuft werden muss, hat er sich in diesem Projekt auch als geeigneter Scanner für eine vollständige Bauwerkserfassung gezeigt. Besonders vorteilhaft im Vergleich zum Faro ist dabei die schnelle und einfache Datenerfassung. Nachteilig ist die sehr begrenzte Batteriekapazität sowie die fehlende USB-Schnittstelle bzw. SD-Speicherkarte zur Übertragung der Daten ohne Netzwerkverbindung sowie eine geringere Robustheit gegenüber Sonneneinstrahlung und Hitze. Aufgrund der technischen Spezifikationen und den Registrierungsergebnissen kann davon ausgegangen werden, dass die Faro-Punktwolke qualitativ besser ist als die des BLK.

Die aus beiden Punktwolken extrahierten Zielmarkenkoordinaten konnten über eine 3D-Ähnlichkeitstransformation mit 7 Parametern bei einer mittleren Standardabweichung (RMS der transformierten Objektkoordinaten) von 3.7 mm aufeinander transformiert werden. Die berechneten Transformationsparameter

$X_0 = 75164.16740 \text{ mm}$	$Y_0 = -669.69493 \text{ mm}$	$Z_0 = 37.12007 \text{ mm}$
$\omega = -0.036144$ °	$\phi = 0.004090$ °	$\kappa = 174.755005$ °
m = 0.999821		

zeigen beim mitberechneten Maßstabsfaktor *m* und bei der Abweichung von der Horizontalen (ω, φ) zunächst auffällige Abweichungen zu den erwarteten Werten 1 bzw. 0, die bei Anwendung auf eine Distanz von 20 m jedoch im Bereich von <1 cm liegen. Eine 3D-Transformation ohne mitgeführten Maßstab verschlechtert das Ergebnis jedoch.

Noch nicht endgültig geklärt sind softwareabhängige Abweichungen in den Registrierungen, die nicht auf die Datenqualität, sondern auf Probleme bei Datenexport und Systemeinstellungen zurückzuführen sind.

2.4.2 Qualität der Punktwolken

Die Qualität der Punktwolken wird in diesem Beitrag zunächst nur exemplarisch analysiert. Dabei werden die Abweichungen des BLK-Scans gegenüber der Faro-Punktwolke als Referenz untersucht. Ein erster Vergleich der beiden registrierten und ungefilterten Punktwolken an einer Außenwand durch Cloud-to-Cloud-Vergleich (Funktion Cloud/Cloud-Distance in CloudCompare) zeigt in Abb. 5 links, dass beide Punktwolken im Bereich von 5-10 mm zueinander passen, mit Ausnahme der Bereiche der Objektkanten (Gebäudeecken, Dachfirst). Auffällig sind geradenförmige Sprünge in den Abweichungen, die vermutlich auf systematische Restabweichungen der Registrierungen oder die Einbeziehung der Scans von benachbarten Gebäuden zurückzuführen sind. Weiterhin ist eine mit zunehmender Gebäudehöhe wachsende Diskrepanz deutlich erkennbar. Ihre Ursache liegt vermutlich in der ausschließlich horizontalen Verteilung der Zielmarken sowie ggf. einer mit zunehmendem Vertikalwinkel schlechteren Scangenauigkeit.



Abb. 5: Vergleich Faro-BLK: links Cloud-to-Cloud; rechts Cloud-to-Cloud nach ICP (in mm)

Abb. 5 rechts zeigt das Ergebnis nach einer zusätzlichen Anpassung der Scans über ICP. Die höhenabhängigen Systematiken werden dadurch weitgehend eliminiert, was erneut auf ein Datumsproblem bei der Registrierung hindeutet. Das Histogramm der Abweichungen zeigt ein Maximum bei 4 mm, was eine gesamthafte systematische Verschiebung der Punktwolken zu einander indiziert.



Abb. 6: Vergleich Faro-BLK: links Cloud-to-Cloud nach der Registrierung mit Register (in mm); rechts Cloud-to-Mesh für Registrierung mit Cyclone (max. 17 mm)

Abb. 6 zeigt die Restabweichungen am Kirchturm, die dort bis zu 30 mm für die Registrierungen aus Register 360 betragen. Bei einem Cloud-to-Mesh Vergleich mit der in Cyclone registrierten Punktwolke sinken dort die Abweichungen auf maximal 17 mm. Auch hier ist es offensichtlich, dass mit zunehmender Gebäudehöhe die Diskrepanz zwischen beiden Scans zunimmt.

3 Photogrammetrie

3.1 UAV-Flüge

Durch die örtlichen Partner konnte eine Drohne vom Typ DJI Mavic Pro mit FC 220 Kamera (12 Mpixel) bereitgestellt werden (Brennweite f=4.7 mm, Bildformat 6.4 mm \times 4.8 mm, Pixelgröße 1.6 µm). Die maximale Flugdauer beträgt ca. 22 Minuten. Die Drohne wurde von einem erfahrenen Operateur manuell gesteuert. Es wurden ein mäanderförmig geflogener Bildverband mit Senkrechtbildern sowie zwei kreisförmig angelegte Bildverbände mit Schrägluftbildern aus unterschiedlichen Entfernungen aufgenommen. Tab. 2 stellt die wichtigsten Flugdaten zusammen. Abb. 7 zeigt Beispielbilder der drei verschiedenen Flüge.

	Bilder	Flughöhe ü.G.	Distanz	Bildmaßstäbe	GSD
Nadir	173	30 – 35 m	20 – 30 m	4600 - 6500	7 – 10 mm
Kreis 1	88	15 – 20 m	14 – 25 m	3600 – 5200	5 – 9 mm
Kreis 2	107	45 – 55 m	25 – 32 m	5500 – 6700	9 – 11 mm
Gesamt	368	15 – 55 m	14 – 32 m	3600 - 6700	5 – 11 mm

Tab. 2: UAV-Flugdaten Lurdji Andreas-Kirche



Abb. 7: Beispielbilder der UAV-Flüge

3.2 Terrestrische Bildaufnahme

Vom Boden aus wurden jeweils zwei handgehaltene Bildverbände mit digitalen Spiegelreflexkameras (DSLR) aufgenommen, die jeweils Zoom-Objekte besaßen. Fix-Fokus-Objektive standen nicht zur Verfügung. Die Objektive wurden jeweils in der kürzesten Brennweite mit Klebeband fixiert, Autofokus und Sensorreinigung wurden deaktiviert. Die Bilder wurden im JPEG-Format bei niedriger Kompression gespeichert.

Die Kamera Canon EOS D6 Mark II mit Objektiv f=24-50mm besitzt einen 26.2 Mpixel Vollformat-CMOS-Sensor (35.9 mm x 24.0 mm, Pixelgröße 6.5 μ m) und einen integrierten GPS-Empfänger. Die Canon EOS 200D mit Objektiv f=10-20mm besitzt einen 24.2 Mpixel CMOS Sensor im DX-Format (22.3 mm x 14.9 mm, Pixelgröße 3.7 μ m). Die Öffnungswinkel bezogen auf die Bilddiagonale betragen 84° (Canon D6) bzw. 106° (Canon 200D).

	Kamera	Bilder	Entfernung	Bildmaßstäbe	GSD
1	Canon 200D	206	1 – 20 m	120 – 2000	2 – 15 mm
2	Canon D6	465	1 – 20 m	120 – 2000	2 – 15 mm

Tab. 3: Terrestrisch aufgenommene Bilddaten

Die Bildaufnahme wurde so angelegt, dass folgende Kriterien weitgehend erfüllt wurden:

- Hohe gegenseitige Überlappung für sichere Messung und Zuordnung von Verknüpfungspunkten (feature points);
- Kombination von Senkrecht- und Schrägaufnahmen für optimale geometrische Schnittbedingungen;
- Kombination von Nahaufnahmen (Abstand 1–2m) und Fernaufnahmen (5–20m) zur Erzielung einer hohen Objektauflösung bei gleichzeitig großer Objektabdeckung durch Übersichtsbilder;
- Zahlreiche gekantete (um jeweils 90° gedrehte) Aufnahmen zur sicheren Kamerakalibrierung;
- Kleine Blendenöffnung zur Erreichung einer hinreichend großen Schärfentiefe.

Die Dauer der Bildaufnahme betrug pro Bildverband ca. 1,5 Stunden.

3.3 Auswertung

3.3.1 Agisoft PhotoScan

Das Softwareprogramm PhotoScan (neu: MetaShape) der russischen Firma Agisoft ist ein weit verbreitetes Structure-from-Motion (SfM) Programm, das in vielen verschiedenen Anwendungsfeldern eingesetzt wird. Es liefert neben der Orientierung (Alignment) und den ausgeglichenen Objektkoordinaten (dünne Punktwolke) eine über semi-globales Matching erzeugte dichte Punktwolke, Oberflächenvermaschungen und True-Orthophotos. Die Software bietet gut dokumentierte Kalibriermodelle für die innere Orientierung und zahlreiche Import- und Exportformate sämtlicher Orientierungsparameter.

Die automatische Prozessierung aller Bilddaten aus den verschiedenen Datensätzen, die sehr stark unterschiedlichen Aufnahmebedingungen und Qualitäten für die Auswertung in PhotoScan unterliegen, erwies sich als schwierig. Die automatisierte Zuordnung aller Bilddaten verwies auf einen deutlichen Orientierungsfehler, der sich durch die Ausbildung zweier zueinander rotierter Rekonstruktionen der Kirche zeigte (Abb. 8). Dargestellt sind Ansichten einer reduzierten dichten Punktwolke nach der automatischen Orientierung aller Bilder. Erkennbar ist hier, neben der doppelten Rekonstruktion des gleichen Gebäudes, dass diese stark zueinander rotiert sind (Abb. 8 links) und auch ein Maßstabsunterschied (Abb. 8 rechts) vorliegt. Es konnte identifiziert werden, dass die Bilder der DJI sowie der Canon EOS 200D eine gleichwertige und gleichgelagerte Rekonstruktion ergeben, während die Bilder der Canon EOS 6D zu einer deutlichen Fehlorientierung führen. Aus diesem Grund wurde in PhotoScan zunächst nur ein reduzierter Bilddatensatz für die Auswertung hinzugezogen.

Die Auswertung der Bilddaten des UAV-Fluges sowie der Canon EOS 200D erfolgte in getrennten Blöcken (chunks). Jeder Bilddatensatz einer Kamera wurde in der Auflösungsstufe "medium" (reduziert um Faktor 4 zur Originalauflösung) automatisch orientiert und geprüft. Es wurden lediglich 2 Bilder im automatischen Orientierungsprozess eliminiert. Abb. 9 rechts zeigt alle Bildpositionen nach der Zusammenführung der getrennten Blöcke. Die Messung der Passpunkte erfolgte vorab getrennt in den Blöcken. Insgesamt wurden in der Erstprozessierung 20 Passpunkte zur Georeferenzierung genutzt (vgl. Abb. 9 links). Aufgrund des Bewuchses um die Kirche wurden auf der nördlichen Seite der Kirche zunächst keine Passpunkte identifiziert, da sie lediglich im terrestrischen Bildmaterial abgebildet sind. Der Großteil der nördlichen Fassade unter dem Blattwerk des angrenzenden Bewuchses ist für Luftaufnahmen verdeckt.



Abb. 8: Ansichten einer reduzierten dichten Punktwolke nach fehlerhafter automatischer Orientierung in PhotoScan 1.4.1; links: fehlerhafte Orientierung mit doppelter Rekonstruktion und Rotation; rechts: Frontansicht der Kirche mit doppelter Rekonstruktion des Eingangsportals



Abb. 9: links: Passpunktübersicht der Auswertung in PhotoScan; rechts: Übersicht der Bilddaten aus Bildflug und terrestrischen Bilder nach der Bildorientierung in PhotoScan

Die zusammengefassten Blöcke wurden abschließend gemeinsam ausgeglichen unter Berücksichtigung der Simultankalibrierung der beteiligten Kameras (f, cx, cy, k1-k3, p1-p2, b1-b2). Der mittlere Rückprojektionsfehler (RMS_{rück}) beträgt 5.2 Pixel für alle Verknüpfungspunkte und 0.5 Pixel für die Passpunkte. Der RMS_{rück} erscheint damit relativ hoch, er konnte jedoch zu diesem Zeitpunkt noch nicht abschließend geprüft werden. Zu identifizieren sind vereinzelte Bilder, die sehr hohe Rückprojektionsfehler aufweisen, so dass davon auszugehen ist, dass noch fehlerhafte Zuordnungen vorliegen, die in einer weiteren Analyse näher zu analysieren sind. Die mittlere 3D-Abweichung an den Passpunkten beträgt 3.1 cm bei einer a priori Genauigkeit von 7 mm (siehe Tab. 4). Die erzielten Abweichungen erscheinen in Bezug auf die Passpunkte realistisch, da zum einen insbesondere die Abbildungsqualität der A4-Marken bei den Drohnenbildern sehr stark eingeschränkt ist und zum anderen wenige Passpunkte mit guten Schnittbedingungen erfassbar waren. Die Betrachtung der Passpunktabweichungen allein über die Bilder der Canon EOS 200D ergibt eine mittlere Abweichung von 6 mm, was ihrer Messqualität entspricht und bei diesem Nahbereichsbildverband zu erwarten war. Die Abweichungen im cm-Bereich resultieren damit vornehmlich aus dem UAV-Flug.

Das texturierte und vermaschte 3D-Modell aus Abb. 10 basiert auf der Generierung einer dichten Punktwolke in der Qualitätsstufe "high" (reduziert um Faktor 2 zur Originalauflösung, 42 Mio. Punkte). Die anschließende 3D-Vermaschung erfolgte mit höchster Vermaschungsdichte (resultierend in etwa 8.4 Mio. Flächen). Zur Texturierung wurde die "adaptive orthophoto" Methode gewählt, die eine Bildauswahl auf Grundlage der Flächenausrichtung wählt und gleichzeitig eine Mosaikierung erlaubt.



Abb. 10: Texturiertes und vermaschtes 3D-Modell aus PhotoScan (Alignment "medium", Dense cloud "high", Mesh "arbitrary, high", Texture "adaptive orthophoto, mosaic")

Aufgrund der komplexen Bildanordnung und der Objektgeometrie entstehen hohe Rechenzeiten im Bereich von ca. 1 Stunde für das Alignment und ca. 5 Tagen für die Berechnung der dichten Punkwolke (Qualitätsstufe "high", Prozessor i7-4770 mit 3.4 GHz CPU, GeForce GTC 645, 16 MB RAM). Das 3D-Modell verweist außerdem auf Unsicherheiten in der Rekonstruktion der nördlichen Seite der Kirche (Verdeckungen durch Bewuchs, schwierige Bildverknüpfungen) sowie an der westlichen Seite (kurze Seite in Abb. 10 links). Das westliche Kirchenportal kann in der dichten Punktwolke noch mit ausreichender Qualität rekonstruiert werden, jedoch erfolgt zum einen durch die Vermaschung eine Fehlkonstruktion des halbrunden Ausbaus und zum anderen sind die Modelltexturen sehr verschwommen und blass. Diese Abweichungen sind näher

zu untersuchen und ggf. alternative Modellierungswerkzeuge einzusetzen. Die Generierung einer dichten Punktwolke mit höherer Auflösung könnte ergänzend zu einem qualitativ höheren 3D-Modell führen. Diese sowie weitere statistische Untersuchungen zur Simultankalibrierung und Ausgleichung der Bildverbände folgen in weiteren Analyseschritten.

3.3.2 RealityCapture

Die Software RealityCapture der slowakischen Firma Capturing Reality ist ein weiteres auf Structure-from-Motion basierendes Photogrammetrie-Programm, das sich durch seine schnelle hochparallele, GPU-basierte Datenverarbeitung und die Möglichkeit zur direkten Integration von Laserscanning-Punktwolken auszeichnet. Dabei wird implizit davon ausgegangen, dass die Laserscanning-Daten von höherer Genauigkeit als die photogrammetrische Auswertung sind, d.h. die Orientierung der Bilder und die damit berechnete Punktwolke wird der Laserscanreferenz angepasst. Nachteilig am Programm sind die begrenzte Einflussmöglichkeit der Parametersteuerung und die sehr sparsame Angabe von statistischen Kenngrößen der Ergebnisse. Je nach gewähltem Modell zur Kameramodellierung werden die Parameter der inneren Orientierung ggf. für jedes Bild getrennt ermittelt (bildvariante innere Orientierung). Photogrammetrische Orientierungs- und Kalibrierungsdaten können nicht in andere Formate exportiert werden. RealityCapture ermöglicht nicht den Export von Punktwolken, sondern ausschließlich die Ausgabe fertig vermaschter Oberflächenmodelle.



Abb. 11: Vermaschtes Gesamtmodell mit RealityCapture 1.0.3

Im Lurdji-Projekt wurden sämtliche vorhandenen Bilder aller Kameras gemeinsam mit der Faro-Punktwolke unter Einbeziehung von 22 Passpunkten ausgewertet. Dabei wurden 1100 Bilder von insgesamt 1570 Bildern vollautomatisch orientiert. Lediglich in einigen Bereichen des Daches und der Turmspitze versagte die automatische Bildverknüpfung, da entweder überwiegend homogene (texturlose) Objektoberflächen oder starke Variationen zwischen Vordergrund (Kreuz) und Hintergrund (Bäume, Park) auftraten. Zur Kamerakalibrierung wurde ein Modell mit drei radial-symmetrischen und zwei tangentialen Verzeichnungsparametern gewählt. Der mittlere Rückprojektionsfehler (entspricht Sigma 0) beträgt 0.6 Pixel.

Die in RealityCapture triangulierte Oberfläche besteht aus ca. 140 Mio. Dreiecken. Die Rechenzeit für das Alignment auf einem Rechner mit i7 6700k Prozessor, 64GB RAM und NVIDIA GTX980TI Grafikkarte betrug ca. 7 Minuten, die Berechnung der dichten Punktwolke und Vermaschung (Qualitätsstufe Medium) benötigte ca. 4,5 Stunden.

Abb. 11 zeigt das berechnete 3D-Modell als Vermaschung (ca. 5 mm Punktabstand). Die hohe Qualität wird zum einen durch die Vollständigkeit des Gesamtmodells belegt, zum anderen zeigen aber auch Detailansichten, wie gut das Objekt rekonstruiert wurde, besonders auch an den schwer zugänglichen Seiten (Abb. 12). Feinste Details wie das Baugerüst am Eingang oder filigrane gusseiserne Zäune im rückwärtigen Bereich werden korrekt modelliert. Lediglich das Kreuz auf der Turmspitze zeigt noch "Schleier", die durch die o.g. Hintergrundproblematik sowie einige fehlende Bilder verursacht sein dürften.



Abb. 12: Detailansichten aus der mit RealityCapture gerechneten Punktwolke

3.4 Ergebnisanalyse

Die aufgenommenen Bildverbände besitzen unterschiedlich gute Konfigurationen für SfM. Während die UAV-Bildflüge durch sehr hohe Überlappungen und gleichmäßige Orientierungsdaten problemlos orientiert werden können, gelingt dies für die einzelnen terrestrischen Bildverbände nur mit manuellen Eingriffen. Insbesondere der D6-Datensatz zeigt einige Problemstellen mit zu geringen Überlappungen oder aus der Reihe fallenden Bildern, die in der Sequenz das Objekt nicht kontinuierlich abdecken. Problematisch sind auch um 90° gekantete Aufnahmen, welche die Orientierung (Alignment) offenbar stören können. Werden jedoch alle vorhandenen Bilder gemeinsam ausgewertet, gelingt die Orientierung weitgehend ohne Probleme bis auf wenige, nicht orientierbare Bilder.

Software	Orientierte	Rückprojektions-	RMS Objekt	Rechenzeit	Rechenzeit
	Bilder	fehler	-	Alignment	Dichte Punktwolke
Agisoft	572	5.2 px	31 mm	1 h	6 T
RealityCapture	1100	0.6 px	5 mm	7 min	4.5 h

Tab. 4: Photogrammetrische Auswertungen

Zum jetzigen Zeitpunkt der Auswertungen ist das Ergebnis von RealityCapture deutlich besser als von PhotoScan (Tab. 4). Zudem sind die Rechenzeiten erheblich kürzer. Aufgrund der beobachteten systematischen Restfehler bei der Kamerakalibrierung in PhotoScan sowie einer noch nicht endgültigen Analyse der Programmeinstellungen und der Passpunktkonfiguration können diese Aussagen jedoch noch nicht als endgültig aufgefasst werden. Insgesamt zeigt das Ergebnis der photogrammetrischen Auswertung, dass eine mittlere Standardabweichung der 3D-Punkte von ca. 5 mm im Objektraum erreichbar ist. Dabei ist zu beachten, dass die Passpunkte aus dem Faro-Scan entnommen und nicht tachymetrisch (mit möglicher höherer Genauigkeit) eingemessen worden sind.

Die dargestellten Punktewolken sind Rohdaten ohne jede Filterung oder manuelle Nachbearbeitung. Für nachfolgende Anwendungen oder Visualisierungen müssen sie jedoch noch aufbereitet und gesäubert werden. Der exemplarische Vergleich beider Punktwolken zeigt im Detail einige gravierende Unterschiede, die jedoch ebenfalls noch nicht endgültig bewertet werden können. Da in RealityCapture die Faro-Punktwolke in den SfM-Prozess integriert ist, entsteht im Anschluss auch eine Punktwolke, die sehr nahe am Laserscanergebnis liegt. Die Detailanalysen zeigen aber auch, dass z.B. filigrane Objektdetails, die nicht vom Laserscanning aufgelöst werden, durch die Bildauswertung sehr gut rekonstruiert werden.

4 Weitere Versuche

Abb. 13 zeigt einen ersten Vergleich (Cloud-to-Mesh) der mit RealityCapture gerechneten Punktwolke mit allen Bilddaten und der Punktwolke des BLK360 gegenüber der Referenz aus dem Faro-Scan. Abb. 13a,b stellen die Abweichungen über die gesamte Objektfassade dar. Die mittlere Abweichung beträgt lediglich 2 mm, d.h. die gemeinsame Auswertung der TLS- und der Bilddaten führt zu einem hochwertigen 3D-Modell der Oberfläche. Zu sehen sind hier flächenhafte Bereiche ähnlicher Abweichungen, die zumindest zum Teil auch durch die sehr unterschiedliche Oberflächentextur des Mauerwerks erklärt werden können (vgl. Abb. 1). Abb. 13c und d zeigen die entsprechenden Histogramme der Abweichungen im unteren und im oberen Bereich der Fassade. Hier wird sichtbar, dass im oberen Bereich schlechtere Daten vorliegen, analog zum Ergebnis aus Abs. 2.4.2.



Abb. 13: Vergleich der mit RealityCapture gerechneten Punktwolke (BLK360 und Bilder) gegenüber Faro

5 Zusammenfassung und Ausblick

Dieser Bericht stellt erste Ergebnisse einer umfangreichen Messkampagne dar, die im September 2018 in Georgien durchgeführt wurde. Ziel war zunächst, den Einsatz unterschiedlicher Messsysteme und Auswerteverfahren für die 3D-Rekonstruktion historischer Kirchen unter realen Bedingungen zu evaluieren, um daraus Erfahrungen für kommende Projekte zu sammeln. Grundsätzlich liefern terrestrisches Laserscanning und auf Structure-from-Motion basierende photogrammetrische Verfahren ähnliche Ergebnisse.

Der Vorteil des Laserscannings besteht zunächst aus der zuverlässigen Erfassung von Punktwolken, ohne dass besonderes ingenieurtechnisches Know-how erforderlich ist. Voraussetzung sind stabile Standpunkte sowie hinreichende gegenseitige Überlappungen bzw. Passpunktkonfigurationen. Der Zeitaufwand für die Aufnahme im Feld betrug mit den eingesetzten Geräten das 2-5fache einer photogrammetrischen Bildaufnahme. Der Laserscanner Leica BLK360 hat sich als taugliches Gerät erwiesen, das im Rahmen seiner Spezifikationen verlässliche Ergebnisse liefert. Die UAV-gestützte photogrammetrische Aufnahme erlaubt die Vermessung der Dach- und Turmflächen, die durch TLS nicht zu erkennen sind. Der Einsatz selbst einer einfachen, preiswerten Drohne hat sich hier als praxistauglich gezeigt. Die terrestrisch aufgenommenen Bilder sind dann problemlos auswertbar, wenn eine große gegenseitige Überlappung (ca. 80-90%) sichergestellt ist und zwischendurch auf gekantete oder aus der Reihe fallende Aufnahmen verzichtet wird. Die Auswertung einzelner terrestrischer Bildverbände ist nicht ohne weiteres möglich gewesen, in der Kombination mit den UAV-Bildern konnten aber alle Aufnahmen orientiert werden. Die Gesamtgenauigkeit liegt mit ca. 5 mm im Bereich des Laserscannings. Je nach Hard- und Softwareausstattung sind die erforderlichen Rechenzeiten bei der hohen Anzahl von Bildern zum Teil nicht praktikabel. Die hochparallelisierte Lösung von RealityCapture zeigt jedoch, das optimierte Implementierungen mit gleichzeitiger TLS-Datenfusion möglich ist und zu hochqualitativen Ergebnissen führt.

Ausblickend sind weitere Untersuchungen geplant, in denen intensivere Analysen der Kamerakalibrierung und der dichten Punktwolken vorgenommen werden. Es ist weiterhin vorgesehen, weitere SfM-Programme mit diesen Daten zu testen.

Schließlich werden in naher Zukunft auch die anderen aufgenommenen Kirchen ausgewertet.

6 Literaturverzeichnis

- ANCHABADZE, Y.D. & VOLKOVA, N.G., 1990: Altes Tbilisi. Miklucho-Maklai (Hrsg.), Russische Akademie der Wissenschaften, Institut für Ethnografie.
- BLASKOW, R., LINDSTAEDT, M., SCHNEIDER, D. & KERSTEN, T., 2018: Untersuchungen zum Genauigkeitspotential des terrestrischen Laserscanners Leica BLK360. Photogrammetrie, Laserscanning, Optische 3D-Messtechnik - Beiträge der Oldenburger 3D-Tage 2018, Luhmann, T. & Schumacher, C. (Hrsg.), VDE Verlag GmbH, Berlin/Offenbach, 284-296.
- CHIZHOVA, M., KOROVIN, D., BRODOWSKII, M., BRUNN, A., STILLA, U. & LUHMANN, T., 2017: Probabilistic reconstruction of orthodox churches from precision point clouds using cellular automata. 3D Arch 2017, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(2/W3), 187-194, doi: 10.5194/isprs-archives-XLII-2-W3-187-2017.
- KERSTEN, T., LINDSTAEDT, M., MAZIULL, L., SCHREYER, K., TSCHIRSCHWITZ, F. & HOLM, K., 2015: 3D-Erfassung, Modellierung und Visualisierung der Festungsanlage Kristiansten in Trondheim durch Photogrammetrie und terrestrisches Laserscanning im Rahmen von ERASMUS-Programmen. 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, 78-89.
- NAUMANN, M., GRENZDÖRFFER, G., NIEMEYER, F. & FRANK, A., 2015: Symbiose von UAS-Photogrammetrie und TLS zur Vermessung und 3D-Modellierung von Kirchenbauwerken am Beispiel des Greifswalder Domes. Photogrammetrie, Laserscanning, Optische 3D-Messtechnik - Beiträge der Oldenburger 3D-Tage 2015.
- REMONDINO, F., 2011: Heritage Recording and 3D Modeling with Photogrammetry and 3D Scanning. Remote Sensing, **3**(6), 1104-1138.

Evaluation of the Microsoft HoloLens for the Mapping of Indoor Building Environments

PATRICK HÜBNER¹, STEVEN LANDGRAF¹, MARTIN WEINMANN¹ & SVEN WURSTHORN¹

Abstract: Mobile augmented reality devices like the Microsoft HoloLens are capable of simultaneously tracking the device location and mapping its environment in real-time. Thus, they offer potential for acquiring at least coarse point clouds and meshes of single rooms or even complete building structures that can be used in the context of building information modelling (BIM) in the future instead of manually modelling existing buildings based on 2D floor plans or manual measurements with laser scanners or computationally expensive image-based 3D reconstruction techniques. For this reason, we provide an extensive quantitative evaluation of the mapping results of the Microsoft HoloLens against terrestrial laser scanning (TLS) ground truth. We show that while the geometry of single rooms can be mapped quite accurately with the HoloLens, deviations may occur in the spatial arrangement of multiple rooms relative to each other.

1 Introduction

Mobile augmented reality (AR) devices like the Microsoft HoloLens (MICROSOFT 2018) that are capable of accurate real-time inside-out tracking offer potential for the in situ visualization of building information modelling (BIM) data, e.g. in the domains of facility management (GHEISARI & IRIZARRY 2016), cultural heritage (BARAZZETTI & BANFI 2017) or education (ARASHPOUR & ARANDA-MENA 2017). This of course implies the availability of building model data for the building environments to be augmented.

While, in recent years, building construction projects are increasingly conducted with the aid of BIM techniques (GHAFFARIANHOSEINI et al. 2017) which results in building models arising together with their corresponding physical buildings, there are many existing buildings for which building model data does not exist. As manually modelling existing buildings based on two-dimensional floor plans or manual in situ measurements is a laborious and costly endeavor, the automatic or semi-automatic acquisition of three-dimensional building model geometry is currently an active field of research (LU & LEE 2017; MA & LIU 2018). In this context, the acquisition of indoor building geometry is mostly done by specialized highly accurate active sensors like laser scanners or computationally expensive image-based 3D reconstruction techniques (DAI et al. 2013).

On the other hand, the HoloLens is not only capable of augmenting indoor environments with corresponding building model data (HÜBNER et al. 2018), but it is also equipped with four cameras and a depth sensor for simultaneously mapping the environment and tracking the movements of the device. The geometric information captured by moving around within an indoor environment are accessible to the user as sparse triangle meshes. While the geometric accuracy and resolution

¹ Karlsruhe Institute of Technology (KIT), Institute of Photogrammetry and Remote Sensing, Englerstraße 7, D-76131 Karlsruhe, E-Mail: [patrick.huebner, martin.weinmann, sven.wursthorn]@kit.edu, steven.landgraf@student.kit.edu



Fig. 1: Overlaying the 2D floor plan (right) with the corresponding mesh (left) captured with the HoloLens shows that the overall geometry of the hallway can be captured without noticeable distortions (LANDGRAF 2018)

of these meshes can certainly not compete with the accuracy of laser scanners, derived building models should principally be sufficient for the task of visualizing building-related information with mobile AR devices like the HoloLens itself.

Compared to laser scanners, the HoloLens as a mapping device enables comfortable and timeefficient indoor mapping. A mesh of a hallway of approximately 30 meters as depicted in Figure 1(left) for example can be captured in a matter of few minutes just by casually walking through it with the device and looking around. Overlaying a 2D floor plan with the derived mesh as depicted in Figure 1(right) shows that its overall geometry is quite accurate. Furthermore, the HoloLens can be considered as a comparatively low-cost mapping device in relation to laser scanners.

In this paper, we evaluate the indoor mapping capacity of the Microsoft HoloLens. While evaluation in terms of the accuracy of various distances measured within indoor data captured with the HoloLens has been done e.g. by LIU et al. (2018) or HUANG et al. (2018), we provide an extensive, quantitative evaluation of the mapping accuracy of complete room-scale meshes against a ground truth provided by a terrestrial laser scanner (TLS). As test environment, an empty apartment consisting of five rooms of different size and a hallway was selected.

The focus of this paper lies on the evaluation of the indoor mapping capacity of the HoloLens with respect to the mapping of coarse indoor building geometries. Our use-case is the mapping of the building geometry itself - i.e. the dimensions of rooms and walls including windows and door openings - and not the mapping of fine geometric details. For this reason, an empty apartment without furniture was chosen as test environment.

Anyhow even finer geometric details can be captured up to a certain degree by the HoloLens device. Figure 2 for example shows various views on a mesh captured with the HoloLens for a stairwell that has interior window casement frames with a width of about 10 cm and handrail rods with a diameter of about 2 cm. The views on the mesh in the Figure 2 show, that enough of these fine geometric details could be captured to at least allow an easy visual interpretation of the scene represented by the mesh. However, especially the detail view in Figure 2(right) demonstrates that the window frames and handrail bars are not captured geometrically complete or accurate.



Fig. 2: Different views on the mesh of a stairwell captured with the HoloLens (LANDGRAF 2018)

The key contributions of this paper are:

- We provide an extensive quantitative evaluation of the mapping results of the Microsoft HoloLens against TLS ground truth.
- We show that while the geometry of single rooms can be mapped quite accurately with the HoloLens, deviations may occur in the spatial arrangement of multiple rooms relative to each other.
- We show that drift effects can occur when mapping large indoor spaces.
- We show that the mapping results of the Microsoft HoloLens are affected by a constant scale factor.

2 Evaluation Method

For the purpose of evaluating the capacity of the Microsoft HoloLens as a device for the rapid and easy-to-use mapping of indoor building geometry, an empty apartment consisting of five rooms of different size and one central hallway was used as test environment. An overview of this apartment is provided in Figure 3.

For acquiring ground truth data to evaluate the HoloLens meshes against, a terrestrial laser scanner (Leica HDS 6000) was used. The obtained point clouds from the different positions of the laser scanner visible as circles in the plan view in Figure 3(left) were registered in a common coordinate system by means of artificial planar and spherical markers placed in the apartment. The complete point cloud of the whole apartment was consecutively cleaned, thinned to an average point distance of 1 cm and meshed with the Poisson surface reconstruction algorithm (KAZHDAN et al. 2006) implemented in the software MeshLab (CIGNONI et al. 2008). The resulting triangle mesh is depicted in Figure 3.

This complete apartment was mapped five times with the HoloLens for obtaining evaluation data. Between each consecutive mapping, all environment data on the device was deleted to ensure five independent measurements. For this mapping process, the commercially available HoloLens App SpaceCatcher (SPACECATCHER 2018) was used for recording the triangle mesh data of the apartment.



Fig. 3: Ground truth data acquired with a terrestrial laser scanner for the test environment

Although the HoloLens device always measures the geometry of the environment via its depth sensor to assist its tracking and the resulting triangle meshes can be obtained via the web interface of the device, it is advantageous to use an app for the cause of recording triangle meshes, because this allows for configuring a desired spatial resolution of the obtained triangle meshes. Furthermore, with suchlike apps, triangle meshes are directly visible for the operator while they get recorded as depicted in Figure 4. In the standard using mode of the HoloLens device, this is not possible in such a comfortable way. Furthermore, the spatial resolution cannot be set by the user outside custom apps that make use of the respective HoloLens SDK functionality. Therefore, the triangle meshes obtainable with apps like the SpaceCatcher as depicted in Figure 4(c) are much smoother and more complete in comparison to meshes recorded without the use of an app depicted in Figure 4(b).

For comparison of the resulting HoloLens meshes against a respective reference mesh, each HoloLens mesh was registered on the reference mesh via the software CloudCompare (CLOUDCOM-PARE 2018) by means of manually selected pass points and subsequent fine registration via the Iterative Closest Point (ICP) algorithm (BESL & MCKAY 1992; ZHANG 1994). The registered



Fig. 4: View of a corridor (a) overlaid with the triangle mesh captured by the HoloLens app SpaceCapture (b) or without scanning app (c)

meshes were then compared via the Hausdorff distance (CIGNONI et al. 1998) in MeshLab (CIGNONI et al. 2008).

Firstly, the results of these five mappings of the apartment with the HoloLens Device were compared to each other in the above described manner to get an impression of the variability of indoor mapping results acquired with the HoloLens. Furthermore, the HoloLens meshes were also evaluated against the ground truth data acquired by terrestrial laser scanner.

3 Results and Discussion

An averaged mesh color-coded with the averaged Hausdorff distances across all 10 possible combinations for comparing the five meshes recorded with the HoloLens device with each other is depicted in Figure 5.

It is clearly visible, that the deviations between the compared HoloLens meshes are in the range of few centimeters for most parts of the apartment. They only reach higher values near the ceiling, where some of the meshes have holes (the ceiling itself was not scanned in the course of all experiments in this work). It can thus be concluded that the HoloLens device performs spatial mapping of indoor environments with a low level of variance between independent measurements even for indoor environments consisting of multiple rooms.

Furthermore, while registering the HoloLens meshes on each other, the scale was kept fixed. So, larger non-constant scale errors can be excluded for the HoloLens. While registering the HoloLens



Fig. 5: Mean Hausdorff distances of all five HoloLens meshes compared with each other



Fig. 6: Mean Hausdorff distances of all five HoloLens meshes compared with the TLS ground truth



Fig. 7: Schematic overview (left) of devations between the rooms of the test environment that are hidden in Figure 6 because of the situation depicted on the right

meshes on the TLS ground truth mesh however, there proved to exist a significant scale factor. In our experiments, this scale factor was estimated to 1.0120 ± 0.002 across 18 tested meshes in total. The averaged point clouds color-coded with averaged Hausdorff distances across all five HoloLens meshes compared to the TLS ground truth mesh are depicted in Figure 6.

Here, the deviations are again comparatively low on the floor and on the outer walls of the mapped apartment. Especially on the inner walls, however, larger deviations between the HoloLens meshes and the ground truth data are visible. In this case, the situation is in fact even worse than indicated by the color-coded Hausdorff distance values. As exemplified on the right-hand side in Figure 7, it can happen that the deviated mesh of the wall of one room is compared to the ground truth mesh of the wall of a bordering room instead of comparing it to the ground truth mesh of the same respective wall if the deviation is large enough in relation to the space between both wall surfaces. So for this use-case of evaluating meshes from indoor building geometry where the spatial deviations can lie in the same range as the distance between wall surfaces of neighboring rooms, the Hausdorff distance as presented by CIGNONI et al. (1998) is not directly applicable. Here, an evaluation procedure is needed that takes into account the topology of indoor building structures and ensures that wall meshes only get compared to meshes of corresponding walls even if there are walls of neighboring rooms with a lower distance.

The schematic overview on the left-hand side of Figure 7 depicts the actual deviations of the HoloLens meshes in comparison to the TLS ground truth data. Here, walls where the HoloLens meshes fit well to the ground truth data are indicated in green, regions of medium differences are indicated in yellow and walls with strong deviations from the ground truth meshes are indicated in red. The arrows indicate the direction in which the respective wall mesh from the HoloLens is shifted in relation to the wall in the ground truth data. The length of the arrow is only chosen for reasons of visibility and does not indicate the amount of deviation.

P. Hübner, S. Landgraf, M. Weinmann & S. Wursthorn



Fig. 8: Registration of every single room independently registered against ground truth with same scale factor

Figure 7 shows that deviations from the ground truth data mainly occur on inner walls parallel to the doors connecting the rooms of the apartment with each other. In fact, the transitions from one room to another through doors are weak spots in the tracking of the HoloLens device and thus the spatial correctness of its mapping data. So it can be assumed that, while the geometry of the respective rooms by themselves is captured well, the overall correctness of the whole measured apartment suffers from weak accuracy in the spatial connection between the respective rooms.

To demonstrate that this is in fact the case, the single rooms from the averaged HoloLens mesh of the apartment have been extracted. Those single-room meshes were then registered separately on the respective ground truth room from the TLS data while keeping the scale fixed to the scale factor that was estimated when registering the HoloLens mesh of the whole apartment on the TLS ground truth. As shown in Figure 8, the resulting single-room meshes fit the ground truth room meshes well.

So it can be concluded, that the high deviations in the walls of the apartment interior in Figure 7 do not result from the scale factor of the HoloLens not being constant but merely from the fact that spatial connections between the rooms that are all mapped quite correctly by themselves are captured with large errors.

Anyhow there definitely are situations where drift in the spatial mapping with the HoloLens is occurring. The left part of Figure 9 for example shows the mesh of a large loop of a hallway with a total length of over 200 m. On the right, the same mesh is mapped on a floor plan. In this case, large deviations contrasting the good results from Figure 1 are clearly visible, and loop closure is not appropriately achieved.

4 Conclusion and Outlook

In this paper, we have presented a first look on the quantitative evaluation of the Microsoft HoloLens in the context of the mapping of indoor building geometry. It shows that the HoloLens allows the scanning of the basic geometry of single rooms consisting of walls with windows and door openings with an accuracy of few centimeters. In the case of indoor environments consisting



Fig. 9: Overlaying the 2D floor plan (right) with the mesh on the left shows that there exist drift effects in the mapping of large-scale spaces with the HoloLens (LANDGRAF 2018)

of multiple rooms connected by narrow passages like doors, there occur larger deviations of the resulting meshes relative to ground truth building geometry. However we have shown, that the respective rooms constituting the indoor environment by themselves still are captured with sufficient accuracy. Nonetheless we also have shown, that in very large, continuous indoor spaces there can also occur drift effects.

In conclusion, the HoloLens reveals high potential for rapid, easy-to-use mapping of basic indoor building geometry. This is of interest, especially in the field of the automated creation of as-built BIM models of existing buildings, which in turn can provide a valuable source for spatially located information to be visualized in augmented reality applications on devices like the HoloLens itself. However, getting from the triangle meshes recorded by the HoloLens device to semantically enriched, topologically correct building models is still challenging and holds much potential for further research.

Also the specific topic of evaluating the indoor mapping capacity of mobile self-tracking mapping devices like the HoloLens still holds open research questions to be addressed in future work. The spatial mapping process of the HoloLens can still be considered as a black-box system, whose exact inner workings are not publicly known and well understood. Furthermore, there is a need for a suited evaluation metric for evaluation scenarios like the one discussed in this work, where there is the need to ensure, that meshes representing specific walls are only compared to ground truth meshes representing the same wall and not other walls nearby, even if those are situated in closer proximity because of the inaccuracy of the meshes to be evaluated. This again presupposes a high level of semantic and topological information to be extracted from the basic triangle meshes delivered by a device like the HoloLens.

5 Literature

ARASHPOUR, M. & ARANDA-MENA, G., 2017: Curriculum Renewal in Architecture, Engineering, and Construction Education: Visualizing Building Information Modeling via Augmented Reality. In: Proceedings of the 9th International Structural Engineering and Construction Conference (ISEC), 1-6.

- BARAZZETTI, L. & BANFI, F., 2017: Historic BIM for Mobile VR/AR Applications. In: Ioannides, M., Magnenat-Thalmann, N. & Papagiannakis, G. (Eds.), Mixed Reality and Gamification for Cultural Heritage. Springer, 271-290.
- BESL, P. & MCKAY, N. D., 1992: A Method for Registration of 3-D Shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(2), 239-256.
- CIGNONI, P., CALLIERI, M., CORSINI, M., DELLEPIANE, M., GANOVELLI, F. & RANZUGLIA, G., 2008: MeshLab: An Open-Source Mesh Processing Tool. In: Scarano, V., Chiara, R. D. & Err, U. (Eds), Sixth Eurographics Italian Chapter Conference, 129-136.
- CIGNONI, P., ROCCHINI, C. & SCOPIGNO, R., 1998: Metro: Measuring Error on simplified Surfaces. IComputer Graphics Forum, 17(2), 167-174.
- CLOUDCOMPARE, 2018: CloudCompare 2.10-alpha, https://www.danielgm.net/cc/. last accessed: 12/2018.
- DAI, F., RASHIDI, A., BRILAKIS, I. & VELA, P., 2013: Comparison of Image-Based and Time-of-Flight-Based Technologies for Three-Dimensional Reconstruction of Infrastructure. Journal of Construction Engineering and Management, 139(1), 929-939.
- GHAFFARIANHOSEINI, A., TOOKEY, J., GHAFFARIANHOSEINI, A., NAISMITH, N., AZHARD, S., EFIMOVA, O. & RAAHEMIFARB, K., 2017: Building Information Modelling (BIM) Uptake: Clear Benefits, Understanding its Implementation, Risks and Challenges. Renewable and Sustainable Energy Reviews, 75, 1046-1053.
- GHEISARI, M. & IRIZARRY, J., 2016: Investigating Human and Technological Requirements for Successful Implementation of a BIM-based Mobile Augmented Reality Environment in Facility Management Practices. Facilities, 34(1/2), 69-84.
- HUANG, J., YANG, B. & CHEN, J., 2018: A Non-Contact Measurement Method based on HoloLens. International Journal of Performability Engineering, **14**(1), 144-150.
- HÜBNER, P., WEINMANN, M. & WURSTHORN, S., 2018: Marker-Based Localization of the Microsoft HoloLens in Building Models. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(1), 195-202.
- KAZHDAN, M., BOLITHO, M. & HOPPE, H., 2006: Poisson Surface Reconstruction. In: Proceedings of the Symposium on Geometry Processing, 1-10.
- LANDGRAF, S., 2018: Evaluierung der Qualität von HoloLens-Punktwolken. Bachelor thesis, Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology (KIT).
- LIU, Y., DONG, H., ZHANG, L. & SADDIK, A. E., 2018: Technical Evaluation of HoloLens for Multimedia: A First Look. IEEE Multimedia, 25(3), 1-7.
- LU, Q. & LEE, S., 2017: Image-Based Technologies for Constructing As-Is Building Information Models for Existing Buildings. Journal of Computing in Civil Engineering, 31(4), 04017005/1–14.
- MA, Z. & LIU, S., 2018: A Review of 3D Reconstruction Techniques in Civil Engineering and their Applications. Advanced Engineering Informatics, **37**, 163-174.
- MICROSOFT, 2018: Microsoft HoloLens. https://www.microsoft.com/de-de/hololens. last accessed: 10/2018.
- SPACECATCHER, 2018: SpaceCatcher HoloLens App. http://spacecatcher.madeinholo.com/. last accessed: 12/2018.

ZHANG, Z., 1994: Iterative Point Matching for Registration of Free-Form Curves and Surfaces. International Journal of Computer Vision, **13**(2), 119-152.

Filter-based Pose Estimation for Electric Vehicles Relative to a Ground-based Charging Platform Using On-board Camera Images

ALEXANDER HANEL¹, PRAJWALA SUDI², STEPHAN PFENNINGER², ECKEHARD STEINBACH³ & UWE STILLA¹

Abstract: Efficient inductive charging of electric vehicles requires the accurate alignment of the vehicle above a ground-based inductive charging platform. This can be achieved using onboard cameras observing the charging platform during the approach manoeuvre. In this contribution, a method for vehicle pose estimation relative to the charging platform is proposed. Vehicle dynamics data obtained from independent pose estimates relying on either markers or markerless features are fused by Kalman filtering to mutual vehicle pose estimates, wherefore the method can be used even when the vehicle is close to the charging platform and the markers are no longer visible in the on-board camera images. The method has been tested with simulated data along trajectories of typical approach manoeuvers to parking lots with a charging platform at the end of the trajectories. Independent from the vehicle motion model and the tested trajectory, deviations of estimated positions from the ground truth of around 0.5 m at the charging platform have been achieved when using dynamics data from both independent pose estimates with an extended Kalman filter.

1 Inductive charging of electric vehicles at charging platforms

Development of electro mobility leads to new requirements for vehicles and external infrastructure. Among these requirements, supplying power to the vehicles plays an important role. Nowadays, plug-in electric vehicles carry a battery providing energy to one or more electrical engines. Typically, batteries of such vehicles are charged using a wired connection between the vehicle and a power socket at a charging station. Alternatives to wired connections are exchanging the battery at battery swapping stations (either stationary or mobile) (SHAO et al. 2017) or doing inductive charging (HWANG et al. 2016). With inductive charging, townscapes don't get changed by many charging stations along the streets and charging becomes more comfortable for the drivers of electric vehicles, to mention just two advantages. Power can be supplied via inductive charging both if the vehicles are being driven or being parked. In both cases, vehicles and roads or parking lots have to be equipped with electromagnetic coils for the power transfer to the vehicle. According to PANCHAL et al. (2018), two aspects have strong negative influence on the efficiency of the inductive charging process of electric vehicles: A large air-gap (ground clearance of vehicles) and misalignment between the coils at the vehicle and in the roads or parking lots. Methods for reducing

¹ Technical University of Munich, Photogrammetry & Remote Sensing, Arcisstraße 21, D-80333 Munich, E-Mail: [alexander.hanel, stilla]@tum.de

² ServiceXpert Gesellschaft für Service-Informationssysteme mbH, Wandsbeker Allee 77, D-22041 Hamburg, E-Mail: [prajwala.sudi, stephan.pfenninger]@servicexpert.de

³ Technical University of Munich, Chair of Media Technology, Arcisstraße 21, D-80333 Munich, E-Mail: eckehard.steinbach@tum.de

the air-gap to increase the efficiency are often limited, for example to vehicles with adaptive air suspension. In contrast, the degree of misalignment can be reduced by placing the vehicle accurately above the coils. Accurate placement can be realized especially for inductive charging platforms installed in parking lots, where specific approach manoeuvres can be driven. When approaching the platform with the vehicle, the induced voltage can be used up to a maximal misalignment of around 60 cm (induced voltage has almost reached zero) as a measure for fine alignment (HWANG et al. 2016). Coarse alignment enabling a smooth approach manoeuvre requires therefore a different method. One possibility would be to use the vehicle position obtained by GPS to determine the relative pose to the charging platform and to guide the vehicle closer to the charging platform. This approach is not always practical as for example in dense urban areas with limited space vehicles are often parked in parking garages, where typically no GPS signal is available. For such situations, a possibility would be to detect the charging platform in images acquired by an environment-observing optical on-board camera in the vehicle during the approach manoeuvre and to derive the relative pose between vehicle and charging platform. Compared to other typical environment-observing sensors in vehicles like LiDAR or radar, optical cameras are relatively cheap (e.g. BMW parts catalogue, e.g. available at BMWFANS.INFO 2017) and mounted at several places of a car providing a good overview over the environment around the car on different sides (e.g. ZIEBINSKI et al. 2016). Among all cameras, front-looking ones are considered as most important for this task, as they observe the upcoming driveway.

For both autonomously-driven cars relying on cameras for environment perception and human drivers, modern charging platforms installed completely below the road surface (e.g. Bombardier primove, BOMBARDIER TRANSPORTATION 2013), could be made visible for example by special markers attached to the ground. Common markers for photogrammetric applications, e.g. for vehicle geometry inspection (AICON 3D SYSTEMS GMBH 2018), are circular markers (e.g. described by NAIMARK & FOXLIN 2002) providing one reference point per marker. Other markers, like the squared ArUco markers (ROMERO RAMIREZ et al. 2018), provide four reference points per marker, one per corner. Reference points with known 3D object coordinates and corresponding 2D image coordinates allow estimating the camera pose relative to the markers by solving the perspective-n-point problem, for example. In the case markers are representing a charging platform, the relative pose between the camera-carrying vehicle and the platform can be obtained.

Regardless of the camera position at the vehicle, markers attached to the ground will become invisible in images when the vehicle is close to the platform. In such situations, marker-based pose estimation becomes impossible. Nevertheless, algorithms like structure-from-motion (e.g. FITZ-GIBBON & ZISSERMAN 1998), visual SLAM (e.g. LEMAIRE et al. 2007) or visual odometry (e.g. NISTER et al. 2004) allow for vehicle pose estimation also in such situations (HANEL & STILLA 2017). Therefore, typically markerless image features like SIFT (LOWE 1999), SURF (BAY et al. 2008) or ORB (RUBLEE et al. 2011) are used to obtain matches between images and to create a map of the environment around the vehicle. Consequently, pose estimation becomes possible for images taken in mapped parts of the environment. However, estimated poses are not directly related to the charging platform.

In this paper, a method for iteratively estimating the vehicle pose during an approach manoeuvre to a charging platform using images from a front-looking on-board camera is proposed, which can

be used for both cases that markers representing the charging platform can be detected in the acquired images or not. The key contribution of this paper is the evaluation of the performance of the proposed method for different vehicle motion models, for the cases that special markers have been detected or not and for different filtering methods for data fusion between poses obtained using markers and poses obtained without seeing markers on simulated data of trajectories of typical parking manoeuvers.

2 Method for vehicle pose estimation during approach manoeuvres to a charging platform

Objective of the method proposed in this paper is to estimate the vehicle pose relative to a charging platform iteratively during an approach manoeuvre to the platform using images taken with a front-looking on-board camera. It is assumed that the charging platform is located in a parking lot of a parking garage or another place, where the beginning of the approach manoeuvre is known, for example by the known entry point to the parking garage. It is further assumed that the charging platform is represented by special markers attached to the ground, which can be detected in images and used for estimating the vehicle pose. By fusing vehicle dynamics data (e.g. vehicle position or vehicle velocity, depending on the dynamics model) obtained independently by pose estimation using special markers with vehicle dynamics data obtained by pose estimation using image features, the proposed method allows to estimate the vehicle pose even when the vehicle is close to the charging platform and the markers are not visible in the camera images.

The algorithm for pose estimation consists of the following two steps, which are supposed to be performed iteratively for every image *i* acquired with the on-board camera during an approach manoeuvre. In the first step, depending on whether a sufficient number of markers has been detected, either stage 1 or stage 2 is performed for obtaining vehicle dynamics data: Stage 1 (Fig. 1, sufficient number of markers detected): Vehicle position estimates x_i and y_i are obtained as vehicle dynamics data from pose estimation using markers on the charging platform ("MB method"). Independently, vehicle velocity estimates v_{x_i} and v_{y_i} are obtained as dynamics data from pose estimation using a markerless method like SLAM or structure-from-motion ("ML method") relying on image features. Stage 2 (Fig. 2, not sufficient number of markers detected): No position estimates are obtained by the MB method as no markers are visible, vehicle velocity estimates are



Fig. 1: Stage one workflow for vehicle pose estimation with markers and markerless image features assuming that the charging platform is visible in the acquired on-board camera images.

obtained by the ML method as in stage 1. In the second step, mutual vehicle pose estimates according to a vehicle motion model describing vehicle dynamics are obtained by fusing available position and velocity estimates using a Kalman filter.



Fig. 2: Stage two workflow for vehicle pose estimation with markerless image features only assuming that the charging platform is not visible in the acquired on-board camera images.

2.1 Marker-based vehicle pose estimation ("MB method")

Marker-based estimation of vehicle poses relies on reference points provided by special markers representing the charging platform. The vehicle dynamics data x_i and y_i describing the planar vehicle position are obtained from pose estimation and are fed to the Kalman filter as measurement at time point t_i .

Many common pose estimation methods rely on known 3D object coordinates and 2D image coordinates of reference points. Using reference points provided by markers has multiple advantages for the proposed application of approach manoeuvres: Pose estimates are linked to the charging platform, allowing to determine the relative pose between vehicle and platform. The 3D object coordinates can be obtained, for example by tacheometry, in an Euclidean parking garage coordinate system with metric scale, which could be used as common coordinate system for data fusion with data obtained by other methods (see Section 2.2). Using ArUco markers providing four reference points per marker can increase the robustness in case of partial visibility or bad viewing conditions, as only a single marker has to be detected to make pose estimation possible.

While 3D object coordinates need to be obtained only once, 2D image coordinates of reference points have to be obtained for every acquired image. Consequently, pose estimation using the MB method can be done for every image, in which a sufficient number of markers can be detected. For detecting markers in images and extracting image coordinates of the reference points, there are many methods available in the literature (e.g. GARRIDO-JURADO et al. (2014) for ArUco markers). For estimating vehicle poses, methods solving the perspective-n-point problem (e.g. WU & HU 2006; LEPETIT et al. 2008) can be used, for example.

2.2 Markerless vehicle pose estimation ("ML method")

With monocular visual SLAM or structure-from-motion algorithms applied to images taken with the front-looking on-board camera, vehicle poses can be estimated using a sufficient number of image features extracted from images taken during the approach manoeuvre, independent from how many special markers can be detected in these images. As in contrast to the MB method no direct link to the charging platform is given, the vehicle dynamics data v_{x_i} and v_{y_i} describing the

vehicle velocity are derived from two consecutive poses and are fed to the Kalman filter as measurement at time point t_i .

E.g. using SLAM or structure-from-motion, a map of 3D points representing the environment in front of the vehicle is created from matches of image features between multiple images. If using a front-looking on-board camera, the created map will cover the environment in the upcoming driveway of the vehicle. Therefore it is likely that later acquired images show parts of the already mapped environment and matches can be obtained for image features between these images and previous ones. By these matches, correspondences between the 3D map points and the later acquired images are established, allowing to use the 3D coordinates of the map points and the 2D image coordinates of the matched features for pose estimation.

With no further information available, the origin, orientation and scale of the map coordinate system are determined by the SLAM or structure-from-motion algorithm and do not correspond with the parking garage coordinate system, wherefore the direct link to the charging platform is not given. For fusing data from the ML method and the MB method, both data should be given in the metric parking garage coordinate system. The metric scale can be obtained for the map coordinate system from a vehicle odometer or being extracted from objects with known size in the environment, like traffic signs (HANEL et al. 2018). The origin and orientation of the map coordinate system can be aligned to the parking garage coordinate system assuming that the vehicle position and driving direction are known at the entry point into the parking garage. As the position of the entry point is prone to large uncertainties (e.g. driveway has multiple lanes), the vehicle velocities v_{x_i} and v_{y_i} and not the vehicle position coordinates x_i and y_i are used as measurement for Kalman filtering.

2.3 Mutual vehicle pose estimates using Kalman filtering with vehicle dynamics data obtained from the MB and ML methods

Kalman filtering is applied to obtain mutual vehicle pose estimates in two ways: 1) Pose prediction for time points without any measurements given (time update only), especially if the markers at the charging platform are not visible in the images and 2) pose estimation by fusing vehicle dynamics data provided as measurements by the MB and the ML method (time update and measurement update).

The Kalman state vector contains parameters describing the vehicle dynamics, like planar position or velocity, using special motion models for vehicles, as described by SCHUBERT et al. (2008). Measurement update is done consecutively for different types of measurements (MB, ML) and not in a common update step to be able to perform measurement updates even if not all measurements are available at a time point. A Kalman filter method for non-linear motion models and measurement models is used.

As vehicles typically follow trajectories with multiple straight and curved parts, different motion models might be the most appropriate ones for describing the vehicle dynamics at different time points. Therefore, the motion model used for Kalman filtering should be adaptive: At every time step, all motion models are updated (time update, measurement update) and the motion model providing the best state accuracy is selected, using the state accuracy as the measure for determining the most appropriate model. The steps for model adaption at time point t_i are:

1) State estimates $X_{i,j}$ for all motion models j are calculated from the state estimate in the current motion model using model transfer equations derived from physical relations between the different motion models (e.g. calculating velocity v_x in x direction from total velocity v and vehicle heading angle θ), 2) state covariance matrices $P_{i,j}$ are calculated using error propagation based on the transfer equations and the state covariance matrix of the current motion model, 3) updated state estimates $\hat{X}_{i,j}$ and covariance matrices $\hat{P}_{i,j}$ are obtained either by time update or measurement update for all motion models separately, 4) updated state variances are extracted from $\hat{P}_{i,j}$ and normalized using the state values $\hat{X}_{i,j}$ assuming that larger state values correlate with larger state variances, 5) state accuracy is calculated for each model by averaging the normalized state variances and 6) the motion model for the next Kalman filtering step is adapted to the model with the lowest state accuracy value, the final state \hat{X}_i and covariance matrix \hat{P}_i are given by the state and covariance matrix of this model.

3 Simulation of approach manoeuvres to a charging platform

Purpose of the simulation is to compare different motion models, measurements and Kalman filter methods with regard to estimated vehicle poses and their deviation from the ground truth for typical approach manoeuvres to a charging platform.

3.1 Simulation settings

In the following, the cases and settings used for simulating approach manoeuvres are described.

Simulation cases

The simulation covers the following cases, which are considered as most relevant for the problem addressed in this paper. Additional cases with variations of parameters like the length of the trajectory, the vehicle velocity or the measurement frequency are considered as less relevant.

- 1) Motion models with the following state parameters describing vehicle dynamics (SCHUBERT et al. 2008): CV ("constant velocity"; position x_i, y_i , velocity v_{x_i}, v_{y_i}), CTRV ("constant turn rate and velocity"; position x_i, y_i , heading angle θ_i , velocity v_i , yaw angle ω_i), CTRA ("constant turn rate and acceleration"; position x_i, y_i , heading angle θ_i , velocity v_i , acceleration a_i , yaw angle ω_i), CCA ("constant curvature and acceleration"; position x_i, y_i , heading angle θ_i , velocity v_i , heading angle θ_i , velocity v_i , acceleration a_i , curvature c_i) and ADAP (the best motion model selected among CV, CTRV, CTRA and CCA for each time point using the approach described in Section 2.3).
- Measurements: MB (vehicle positions x_i and y_i from marker-based pose estimation), ML (vehicle velocities v_{xi} and v_{yi} from markerless pose estimation), CAR (velocity v_i and yaw rate ω_i from standard vehicle sensors for comparison purpose), MB + ML (combination of x_i, y_i and v_{xi}, v_{yi}) and MB + ML + CAR (combination of x_i, y_i and v_{xi}, v_{yi} and v_i and ω_i).
- 3) Kalman filter methods: EKF (extended Kalman filter) and UKF (unscented Kalman filter).
- 4) Approaching manoeuvre trajectories representing typical parking situations: STRAIGHT (driving straight from the driveway into a parking lot at its end) and CURVED (turning right from the driveway into a parking lot on the right side, "transverse parking").

Trajectories

The ground truth trajectories (Fig. 3) start at the entry point into the virtual parking garage and end in a parking lot equipped with a charging platform. The charging platform is placed few meters before the end of the trajectory to simulate that markers can't be seen in some images. Both trajectories have the same length of around 30 m.



Fig. 3: Ground floor view of a virtual parking garage with the virtual straight trajectory (left) and curved trajectory (right). The trajectories (blue crossed lines) start at the entry point to the parking garage. The charging platform installed in a parking lot is visualized in red.

The virtual vehicle is moving along the trajectory with a constant velocity of 1 m/s. It is assumed that the dynamics of the virtual vehicle are equal to the virtual front-looking on-board camera.

Virtual camera

For specifying the virtual front-looking on-board camera (specifications see Tab. 1), the properties of the front-looking camera used to record the Cityscapes dataset of road scene images (CORDTS et al., 2016) are used.

Tab. 1: Specifications of the virtual cam-	-
era used for simulation.	

Geometric resolution	1024 x 2048 px	
Principal point	513 px, 1097 px	
Focal length of the lens	5 mm	

Tab. 2: Specifications of the virtual markers repre-
senting the charging platform.

Geometry of markers	7 x 5 grid
Distance between mark- ers	20 cm
Marker edge size	10 cm
Reference point distance on a marker	5 cm
Min. marker size in im- ages	10 px

Measurements for Kalman filtering

Virtual MB measurements x and y are obtained by performing pose estimation (solving the PnP problem) using virtual ArUco markers (specifications see Tab. 2) representing the charging platform. For reference points on the ArUco markers, 3D object coordinates and corresponding 2D image coordinates are simulated. White zero-mean Gaussian noise is added for the sake of realism. Points being far away from the camera tend to lead to large pose errors (PENTENRIEDER et al. 2006), whereas markers which would be imaged with a pixel size below a minimal marker size (threshold see Tab. 2) are not considered for MB measurements. Points which would not be visible in images (e.g. lying behind the camera) are not considered, as well. Virtual ML measurements v_x and v_y are obtained from position differences between consecutive points along the ground truth trajectory with noise added. ML measurements are not given during ML initialization in the first seven seconds to account for the time needed for correctly initializing scale and orientation (see Section 2.2). Virtual CAR measurements v and ω are obtained from the ground truth trajectory with noise added. CAR measurements are used for comparison between images as data from additional vehicle sensors and velocity and yaw rate as data from standard vehicle sensors. All virtual measurements are given with a frequency of 2 Hz.

Kalman filter methods

An extended Kalman filter and an unscented Kalman filter are used as filtering methods. Both methods allow for non-linear motion and measurement models. The extended Kalman filter requires a smaller number of filter parameters to be determined, which might make it more robust to use. In contrast, the unscented Kalman filter does not require Jacobian matrices for time and measurement updates. Therefore, this filter method is not restricted to first-order approximations of the non-linear motion and measurement models, which could lead to sub-optimal performance or sometimes divergence of the extended Kalman filter (WAN & VAN DER MERWE 2000).

Kalman filter initialization and noise characteristics

For initializing the Kalman filter, the first trajectory position at the entry point into the parking garage is used as initial vehicle position x_1 , y_1 in the parking garage coordinate system. The heading angle θ_1 is set to the trajectory direction at the entry point. The initial velocity v_1 is taken from the first CAR measurement, v_{x_1} and v_{y_1} are calculated thereof using θ_1 . The initial yaw angle ω_1 is taken from the first CAR measurement, as well, assuming that CAR measurements are available always when the vehicle is in use. As accelerations are not observed by any sensor, a_1 is initialized to zero. White zero-mean Gaussian noise is added to all initial values.

The state covariance matrix P_1 is initialized with a pessimistic approximation of variance values reported in previous work (e.g. SCHUBERT et al. 2008). A small value of 0.01 is added as covariance between different state parameters taken from the same measurement source (MB, ML, CAR). No covariance is assumed between state parameters from different measurement sources.

Setting process noise covariance matrix Q and measurement noise covariance matrix R for Kalman filtering can be a challenging task (ABBEEL et al. 2005), as often no information about the specific noise characteristics is available. Different strategies for setting those matrices have been proposed (e.g. ODELSON et al. 2006; ÅKESSON et al. 2007). To avoid additional complexity in the simulation, which might correlate with the obtained results, values for Q and R are selected based on assumptions and not on one of the aforementioned strategies. For state parameter variances in Q, the same values as for P_1 are used. In contrast to P_1 , the covariance values on the minor diagonals are assumed to be random values obtained from a zero-mean Gaussian distribution with a small sigma of 0.01. Measurement variance values for R are selected based on persimistic approximation of values reported by PENTENRIEDER et al. (2006), the covariance values assumed to be random values obtained form a zero-mean Gaussian distribution with a small sigma of use reported by PENTENRIEDER et al. (2006), the covariance values assumed to be random values obtained form a zero-mean Gaussian distribution with sigma equal to 0.01. When obtaining the random values for Q and R, it is ensured that the symmetry property for covariance matrices is not violated.
3.2 Deviation of estimated vehicle positions from the ground truth

The different simulation cases are compared based on the deviation of the estimated mutual vehicle positions from the ground truth positions. The estimates are obtained from 100 executions of each case to avoid biases which might result from specific random values or noise, which are used for the measurements or for initializing the Kalman filter, for example, as described above.

MB measurements

For the cases with MB measurements only (Fig. 4), deviations are increasing in the centimetre range as long as measurements are available. When measurements become unavailable (markers not detected anymore \rightarrow predicted poses, $> \sim 27$ s), the deviations are increasing in the meter range in only a few seconds. This strong increase can be observed especially for the curved trajectory. A reason for this might be that the prediction begins shortly after the curve in the trajectory, whereas the Kalman filter might not have settled with regard to the new movement direction. Another observation is a stronger increase of the deviation for predicted poses for the UKF than for the EKF.





ML measurements

For cases with ML measurements only (Fig. 5), a stronger increase of the deviation can be observed for predicted poses (< -7 s; time needed for ML initialization) compared to the following time points with measurements available. For the curved trajectory, the deviation increases in a step-like shape at the beginning of the curve (-20 s), which was not observed for the case with MB measurements only. Again, estimates obtained by UKF show larger deviations compared to estimates obtained by EKF.



Fig. 5: Position deviation for mutual pose estimates obtained with ML measurements. The vertical dashed line indicates the beginning of the curve in the trajectory. For the first time points (white background color), mutual poses have been predicted only due to the time needed for initializing the ML method. For later time points (red background), measurements have been available.

CAR measurements

For cases with CAR measurements only (Fig. 6), the motion models CV and ADAP show almost constant strong increase of the deviation of more than one meter per second. A disadvantageous property of the CV model for the given measurements might be that neither the heading angle nor yaw rate are modelled as state parameters. CTRV, CTRA and CCA models having heading angle and yaw rate as state parameters show a remarkably lower increase of the deviation.



Fig. 6: Position deviation for mutual pose estimates obtained with CAR measurements. The vertical dashed line indicates the beginning of the curve in the trajectory. CAR measurements have been available along the complete trajectory (red background color).

MB and **ML** measurements

For cases using MB and ML measurements together (Fig. 7), lower deviation values can be observed at the end of the trajectory (where the charging platform is located) in comparison to cases using MB or ML measurements only. Furthermore, the deviations at the end of the trajectory are similar (max. ~1 m difference) for all of the four cases in Fig. 7 and for all motion models in these cases. The most remarkable increase in the deviation can be observed for the last time points (> \sim 27 s), for which no MB measurements are available anymore (cf. the red background color in Fig. 4). In contrast to using MB and ML together, for other measurements some of the four cases show remarkably larger deviations than others: Cases with UKF if using MB or ML measurements only, motion models CV and ADAP if using CAR measurements only, for example.



Fig. 7: Position deviation for mutual pose estimates obtained with MB and ML measurements. The vertical dashed line indicates the beginning of the curve in the trajectory. Measurements have been available along the complete trajectory (red background color).

MB, ML and CAR measurements

For cases using MB, ML and CAR measurements together (Fig. 8), the deviations are larger than for using MB and ML measurements together. In contrast to previous cases for which UKF has shown worse performance, using the EKF shows larger deviations at the end of the trajectory using MB, Ml and CAR measurements. A possible explanation might be that when fusing multiple different types of measurements, the first-order approximation of non-linear motion and measurement models by the EKF might not be sufficient.



Fig. 8: Position deviation for mutual pose estimates obtained with MB, ML and CAR measurements. The vertical blue dashed line indicates the beginning of the curve in the trajectory. Measurements have been available along the complete trajectory (red background colour).

General discussion

For none of the cases it has been analysed whether the deviations are deviations in longitudinal or lateral direction (wrt. the vehicle movement direction). For all cases, the motion model CCA did diverge when using the EKF (deviation values in the range of 10⁶ m and above). As mentioned already in the paragraph above, the first-order approximation of the EKF filter might not be sufficient for this motion model. The adaptive motion model has shown varying performance: For some cases, deviations are lower than for the fixed motion models (CV, CTRV, CTRA, CCA), while for others they are larger. A correlation to simulation settings (like the Kalman method or the trajectory) can't be observed visually. Coming back to the task of aligning an electric vehicle at a charging platform, the results shown above lead to the recommendation to use MB and ML measurements together with EKF. If using MB and ML with EKF, none of the motion models has led to larger deviations than approximately 0.5 m at the end of the trajectory. That avoids the question which motion model to select, which is considered as important for robust application of the proposed method in daily use of vehicles, even though in certain other cases lower deviations (min. ~ 0.2 m) have been achieved with other measurements. Considering the induced voltage for fine alignment of the vehicle above the charging platform within a maximal misalignment of 0.6 m, a position deviation of approximately 0.5 m can be seen as sufficient for coarse alignment.

4 Conclusion

In this paper, a method for vehicle pose estimation relative to a ground-based inductive charging platform represented by special markers has been proposed. In a first step of the method to be performed iteratively during an approach manoeuvre, vehicle dynamics data are independently obtained from camera pose estimation using the markers detected in images acquired by a front-looking on-board camera and from camera pose estimation based on a markerless algorithm like visual SLAM. In a second step, vehicle dynamics data are fused to achieve mutual vehicle pose estimates. By fusing the data from different sources, the proposed method is capable to estimate vehicle poses even if no markers are available, for example because the vehicle is close to the charging platform at the end of the approach manoeuvre. Experiments with simulated camera pose estimates and vehicle dynamics data have shown for typical approach manoeuvers to parking lots that deviations of estimated vehicle positions from the ground truth of around 0.5 m can be achieved using both independent estimations. Obviously, goal of future work could be to test the proposed method in real-world experiments. A further goal could be to test and, if necessary, improve the robustness of the method for bad conditions, which can occur on roads, like rain.

5 References

ABBEEL, P., COATES, A., MONTEMERLO, M., NG, A. Y. & THRUN, S., 2005: Discriminative Training of Kalman Filters. Proceedings of Robotics: Science and Systems, **2**, 1.

AICON 3D SYSTEMS GMBH, 2018: ProCam | AICON 3D Systems. Website. https://www.aicon3d.com/en-GB/products/vehicle-testing/procam, last access 2018-10-30.

- ÅKESSON, B. M., JØRGENSEN, J. B., POULSEN, N. K. & JØRGENSEN, S. B., 2007: A Tool for Kalman Filter Tuning. Computer Aided Chemical Engineering, **24**, 859-864.
- BAY, H., ESS, A., TUYTELAARS, T. & VAN GOOL, L., 2008: Speeded-Up Robust Features (SURF). Computer Vision and Image Understanding, **110**(3), 346-359.
- BMWFANS.INFO, 2017: BMW Parts Catalog. Distance Systems, Cruise Control, BMW 750iX G11. http://bmwfans.info/parts-catalog/G11/Europe/750iX-N63R/browse/distance_systems_cruise_control, last access 2017-10-29.
- BOMBARDIER TRANSPORTATION, 2013: primove true e-mobility. Brochure. Available online: http://primove.bombardier.com/fileadmin/primove/content/MEDIA/Publications/BT_Brochure PRIMOVE 210x280 2013 final upd 110dpi SP.pdf, last access 2018-12-19.
- CORDTS, M., OMRAN, M., RAMOS, S., REHFELD, T., ENZWEILER, M., BENENSON, R., FRANKE, U., ROTH, S. & SCHIELE, B., 2016: The Cityscapes Dataset for Semantic Urban Scene Understanding. IEEE Conference on Computer Vision and Pattern Recognition, 3213-3223.
- FITZGIBBON, A. W. & ZISSERMAN, A., 1998: Automatic camera recovery for closed or open image sequences. European Conference on Computer Vision, Springer, 311-326.
- GARRIDO-JURADO, S., MUÑOZ-SALINAS, R., MADRID-CUEVAS, F. J. & MARÍN-JIMÉNEZ, M. J., 2014: Automatic generation and detection of highly reliable fiducial markers under occlusion. Pattern Recognition, 47(6), 2280-2292.
- HANEL, A. & STILLA, U., 2017: Calibration of a vehicle camera system with divergent fields-ofview in an urban environment. 37. Wissenschaftlich-Technische Jahrestagung der DGPF, 26, 160-169.
- HANEL A., MITSCHKE, A., BOERNER, R., VAN OPDENBOSCH, D., HOEGNER, L., BRODIE, D. & STILLA, U., 2018: Metric scale calculation for visual mapping algorithms. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, XLII-2, 433-440.
- HWANG, K., CHO, J., PARK, J., KIM, D., KWON, J.H., KWAK, S.I., PARK, H.H. & AHN, S., 2016: An Autonomous Steering System for Wireless Charging Electric Vehicles to Minimize Power Degradation. 2016 3rd Asia-Pacific World Congress on Computer Science and Engineeringha, 60-65.
- LEMAIRE, T., BERGER, C., JUNG, I.-K. & LACROIX, S., 2007: Vision-Based SLAM: Stereo and Monocular Approaches. International Journal of Computer Vision, 74(3), 343-364.
- LEPETIT, V., MORENO-NOGUER, F. & FUA, P., 2008: EPnP: An Accurate O(n) Solution to the PnP Problem. International Journal of Computer Vision, **81**, 155-166.
- LOWE, D. G., 1999: Object recognition from local scale-invariant features. Proceedings of the Seventh IEEE International Conference on Computer Vision, **2**, 1150-1157.
- NAIMARK, L. & FOXLIN, E., 2002: Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker. Proceedings of the International Symposium on Mixed and Augmented Reality, 27-36.
- NISTER, D., NARODITSKY, O. & BERGEN, J., 2004: Visual odometry. Proceedings of the IEEE 2004 Computer Society Conference on Computer Vision and Pattern Recognition, 1, I-652-I-659.
- ODELSON, B. J., RAJAMANI, M. R. & RAWLINGS, J. B., 2006: A new autocovariance least-squares method for estimating noise covariances. Automatica, **42**(2), 303-308.

- PANCHAL, C., STEGEN, S. & LU, J., 2018: Review of static and dynamic wireless electric vehicle charging system. Engineering Science and Technology, **21**(5), 922-937.
- PENTENRIEDER, K., MEIER, P. & KLINKER, G., 2006: Analysis of Tracking Accuracy for Single-Camera Square-Marker-Based Tracking. Proceedings of Dritter Workshop "Virtuelle und Erweiterte Realität der GI-Fachgruppe VR/AR".
- ROMERO RAMIREZ, F. J., MUÑOZ-SALINAS, R. & MEDINA-CARNICER, R., 2018: Speeded Up Detection of Squared Fiducial Markers. Image and Vision Computing, **76**, 38-47.
- RUBLEE, E., RABAUD, V., KONOLIGE, K. & BRADSKI, G., 2011: ORB: An Efficient Alternative to SIFT or SURF. Proceedings of the 2011 International Conference on Computer Vision, 2564-2571.
- SCHUBERT, R., RICHTER, E. & WANIELIK, G., 2008: Comparison and Evaluation of Advanced Motion Models for Vehicle Tracking. 11th International Conference on Information Fusion, 730-735.
- SHAO, S., GUO, S. & QIU, X., 2017: A Mobile Battery Swapping Service for Electric Vehicles Based on a Battery Swapping Van. Energies 2017, **10**, article number 1667.
- WAN, E. A. & VAN DER MERWE, R., 2000: The Unscented Kalman Filter for Nonlinear Estimation. Proceedins of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium, 153-158.
- WU, Y. & HU, Z., 2006: PnP Problem Revisited. Journal of Mathematical Imaging and Vision, 24, 131-141.
- ZIEBINSKI, A., CUPEK, R., ERDOGAN, H. & WAECHTER, S., 2016: A Survey of ADAS Technologies for the Future Perspective of Sensor Fusion. Computational Collective Intelligence, 135-146.

Automatic Generation of Large Point Cloud Training Datasets Using Label Transfer

TORBEN PETERS¹ & CLAUS BRENNER¹

Abstract: In this work, we describe a framework for the automatic annotation of large scale point clouds. Our input data is generated using a mobile mapping system, which features LiDAR as well as camera image acquisition. In a first step, we labelled images of the measurement campaigns using pre-trained CNN (convolutional neural network) models for semantic segmentation. To that end, we ran PSPNet, which was trained on the cityscapes dataset. Then, since all camera poses are known from the mobile mapping GNSS/ IMU system, the image labels were transferred to the 3D points measured by the LiDAR scanners. Using this approach, we are able to automatically generate very large amounts of labelled point clouds, which can be used as training data. However, the dataset contains label noise, mainly because of calibration- and classification-errors, label policy, and occlusions which occur due to platform or object movements. We investigate different types of label noise and show how to recover the erroneous labelled 3d points. We do this by learning to identify wrong patterns in the label- aggregation and by calculating temporal features for every 3d-point by aligning large scale point clouds from different measurement campaigns into a voxel grid.

1 Introduction

Many state-of-the-art solutions to recognition, interpretation or planning problems are based on deep learning techniques. For example, in the domain of autonomous driving, deep learning is used for object segmentation and classification, motion planning, and even end-to-end learning. In classical supervised learning, networks are trained with data of a specific domain for the given task. However, if one domain intersects with another domain, the knowledge can be transferred between tasks. This procedure is called transfer learning. Autonomous cars often use different sensors in order to solve related tasks, for example cameras and LiDAR sensors, which makes the transfer of knowledge a viable approach. In this paper, we are presenting a framework for label transfer between images and 3d points. We are showing to what extend the quality of the labels is degrading by investigating different types of label noise. We do this by comparing the transferred labels with a manually annotated reference dataset. Furthermore, we are improving the label transfer by implementing a full ray tracing. Additionally, by storing aggregated labels in histograms, we are able to show that some types of noise are identifiable in the labels itself. Other types of noise are separable from the true classes by analyzing the temporal recurrence of a 3dpoint. We do this by aligning point clouds from different measurement campaigns in a voxel grid. By determining the occupancy count of the voxels, we can calculate a feature that helps us to identify wrongly labeled points. In the end, we are investigating different approaches for label noise reduction and show that we are able to recover the label noise by manually labeling only a small amount of data.

¹ Leibniz University Hannover, Institute of Cartography and Geoinformatics

2 Related work

Semantic segmentation attempts to segment and classify parts of a scene by doing pixel- or pointwise classification. Since the rise of deep learning, popular approaches are using fully convolutional neural networks for the semantic segmentation of images (LONG et al. 2015). More advanced approaches are still relying on neural networks like PSPNet (ZHAO et al. 2017) or Deeplab (CHEN et al. 2018) but improved the performance by modifying their architecture and components. In order to compare different approaches many benchmarks are available. The so called cityscapes benchmark for semantic segmentation was released by CORDTS et al. (2016). They include annotated data from 50 different german cities and annotated 25000 images with up to 30 classes. In our work we are trying to use this information in order to enrich our own dataset. With Transferring labels we refer to the general procedure of mapping label information from one domain (e.g. 2d) into another (e.g. 3d). The transfer between 3d- and 2d-space has been done before, XIE et al. (2016) transferred human annotated point clouds into images in order to create arbitrary amounts of training images and corresponding labels. Transferring labels from 2d images to 3d point clouds is also not completely new. Notable contributions are made by MCCORMAC et al. (2016) and HERMANS et al. (2014). Both are considering static scenes and transfer labels by labeling RGB-D images wich are then mapped into 3d-scenes. Another way of labeling 3d point clouds from images was proposed by BOULCH et al. (2017). They are projecting the point cloud into images which are then classified by convolutional neural networks. In order to classify the point cloud the classified image pixel are remapped from 2d into 3d. Recently in 2018, ZHANG et al. used semantic segmented images in order to project them into point clouds. However, in comparison to our work they used stationary terrestrial laser scanners, which do not exhibit the same amount of label noise as is the case with a mobile mapping system or laser scanners of autonomous cars. The handling of dynamic objects in stereo images was investigated by KOCHANOV et al. (2016). They are using scene flow in order to propagate the semantics of dynamic objects into 3d maps.

Label noise describes the presence of wrong labels in a data set. There are many different types of noise described in literature (NETTLETON et al 2010; DRORY et al. 2018). One type of noise are randomly flipped labels. Some studies are showing that deep neural networks are robust to this type of noise as long as the noise level does not exceeds a certain level (DRORY et al. 2018; ROLNICK et al. 2018). Another type is the flip label-noise. In this case a each label is confused to a certain degree with a different class.

Label noise cleaning deals in general with correcting the label noise in a data set. Today many state of the art procedures are based on deep neural networks (XIAO et al. 2015; LEE et al. 2017; PATRINI et al. 2016; HENDRYCKS et al. 2018). While Lee et al. tried to correct the label noise by demoting or removing wrong instances XIAO et al. (2015) corrected the labels directly. Both rely on a small "clean" dataset that helps them to estimate the noise in order to clean the corrupted labels. Patrini et al. were able to correct the label noise in an unsupervised way by estimating a transition matrix which describes the probability that a class is flipped to another. Hendrycks et al. improved this approach by incorporating supervised information. However the topic of label noise handling has been extensively studied as shown by the survey of FRENAY et al. (2014). Additionally we would like to note that it is not uncommon to deal with label noise cleaning in the

scenario that is presented in this work. Some of the previous mentioned authors used conditional random fields (CRF) in order to improve the quality of the mapped labels in 2d and 3d (XIE et al 2016; MCCORMAC et al 2016; HERMANS et al 2014).

3 Dataset

Our dataset was created within fourteen different measurement campaigns in the time between March and October of 2017. We used a Riegl-VMX 250 mobile mapping system mounted on a van and equipped with two full-circle laserscanners, two cameras and an inertial measurement/global navigation satellite system (IMU/GNSS) unit. Each measurement campaign covered the same route and was done at least once a month in an urban environment, which resulted in ~1 billion points and ~10.000 images per campaign.

We first automatically annotated the (RGB camera) images taken during our measurement campaigns using pre-trained convolutional neural network (CNN) models for semantic segmentation. To that end, we ran PSPNet (Zhao et al.) which was trained on the *cityscapes* dataset (CORDTS et al. 2016)). Then, using the known pose of all cameras, we transferred all labels to the 3d points. The dataset constructed in this way contains about 15 billion labelled points.

3.1 Methodology

Figure 1 shows an overview of our methodology. As mentioned before, our first step is to use PSPnet by Zhao et al. to classify each pixel in every image taken by our mobile mapping van. According to the results reported in Zhao et al., PSPNet reached a mean intersection over union (IoU) of 81.2 % for the cityscapes dataset. For each image we calculated the pose of the camera in UTM coordinates and the view direction vector *d*. Furthermore, the intrinsic parameters of each camera were available. By projecting the 3d points into all semantically segmented images, their labels can be picked and we are able to map each label to a 3d-point.



Fig. 1: Overview methodology (f.l.t.r): camera images, semantic segmented images, point cloud containing transferred labels colored by majority label.

Because a 3d-point may be projected into multiple images, it is also possible that different labels are associated with this point. In order to accumulate these labels, we define one histogram h_i per 3d-point p_i , with $i \in \{1, ..., n\}$ where n is the number of measured points. Each histogram has 20 bins, and if a class $c = j \in \{1, ..., 20\}$ is observed, we are incrementing the bin j by one. The

resulting histogram can contain contradictory information, for example if a car drives through the scene while the images are taken, we are accumulating car and street labels. The following picture shows an example of the transferred labels. We adhere to the color conventions used in the cityscapes project in order to make our results easier comparable.



Fig. 2: (a) Point cloud colored by RGB images,

(b) Point cloud colored by transferred majority label

3.2 Label noise

By *label noise*, we term the effect that a certain amount of points will have non-correct labels. Although we do not know the exact process that leads to these label flips, we can identify several reasons. Examples are calibration errors or poor prediction quality of PSPNet due to different sensor types or different settings from the original data. Besides those errors we identified several systematic causes for label noise in our dataset. This is noise due to

- the label policy of cityscapes,
- occlusion in 3d-point clouds, and
- the difference in capture time between the laserscanner and camera sensors.

In the following sections we are describing each case in detail.

3.2.1 Noise due to label policy

As shown in figure 3(a) the image label policy of cityscapes requires that surfaces which are visible behind tree crowns, such as building facades, have to be assigned the tree label. However, many of the laser rays will go through the tree crown and thus the points on the facade will be labelled as tree. An example for this kind of label noise is shown in figure 3(b). In the depicted scene, the tree labels are erroneously projected onto the building behind.



Fig. 3: (a) Semantic segmented image

(b) Point cloud colored by transferred majority label

One can see this effect also in figure 2 (b). By storing the aggregated labels in a histogram we are also accumulating labels from the object in the front.

This information can be used in order to segregate the noisy histograms and thus counteract the noise induced by the label policy. We can verify this by clustering the histograms using k-means clustering with k = 32, shown in figure 4. In this figure, each histogram is assigned to the nearest centroid. It is colorized according to the cluster number, therefore histograms with the same color belong to the same cluster.



Fig. 4: Histograms clustered using k-means clustering. Classes are randomly colored. Red class on the building visualizes histograms which contains tree and building labels.

By looking at the bright red clusters on the facades it becomes apparent that this kind of noise is separable from pure tree and background classes. To that end, we can directly use the histograms h_i as features in order to learn if the histogram is affected by the noise and therefore the class derived from the peak of the histogram is probably wrong.

3.2.2 Noise due to occlusion

In order to transfer labels from the images to the scan strips, the strips are processed one after the other, using all images that were taken along this strip. Since all single laser points and all captured images are time stamped, the complete geometry can be recovered using the IMU/GNSS data, so that every laser point can be projected to every image, using exterior and interior orientation, as well as lens distortion terms.

However, since the laser and image rays do not coincide, occlusions are quite frequent, which results in 3d points being assigned the label of an occluding object rather than the correct object, which will introduce label noise. This effect is mitigated to a certain extend by the fact that we are accumulating all transferred labels in a histogram, as described above. Therefore, if the point is unoccluded for most of the time, the histogram will still peak at the correct label.

In order to improve this further, we have implemented a full ray tracing for the label transfer. All points of a scan strip are sorted into a voxel grid. Then, when determining the label of a 3d-point, the ray to each camera center is traced in this grid and the point is considered to be occluded if an occupied cell is found along the ray. We used 10*cm* grid cells for this operation. In addition, a voxel pyramid was computed to speed up the computation. Although this process is only a crude approximation of the complex interactions between laser and image rays, it generally improves the label transfer.



Fig. 5: (a) point cloud colored without raytracing (b) point cloud colored with raytracing

An example is shown in figure 5. Here, the colorization with ray tracing prevents the building in the background from being colored with tree (stem) pixels. The red car on the street is also less scattered to the ground if ray tracing is used. The colorization shown here has direct impact on the label transfer because we are aggregating class labels by mapping classified image pixels to 3d points using the very same mechanism.

3.2.3 Noise due difference in capture time

Further label transfer errors occur if the recorded object is not static. This ensues from the fact that our cameras are facing backwards while the laserscanners are vertically inclined. We tried to create a feature that gives us a measure of how dynamic a captured 3d-point is. We aligned all points between different measurement campaigns using the strip adjustment approach described by Brenner (2016). By creating a voxelgrid using an edge length of 5 cm, we grouped different points

and histograms in 3d-space. We counted the number $v_i \in \{1, ..., 14\}$ of how often a voxel was measured in distinct measurement campaigns. This value is assigned to each point that is present in the corresponding voxel. After normalization, a score of $v_i = 0$ means that p_i is highly dynamic and $v_i = 1$ means that p_i is highly static.



Fig. 6: Point cloud colored by (a) intensity, (b) class label, and (c) voxel count v_i black color means measured once or never, bright color means measured 14 times.

Figure 6 shows a scan strip with a cyclist on the street and two parking cars. Because of the difference in capture time the cyclist is labelled as street in figure 6(b). By counting the number of occurrences v_i , we can create a feature that can be used to identify dynamic objects. This feature is visualized in figure 6(c), where we can see, that the cyclist who introduced label noise can easily be removed. Furthermore this feature preserves the parking cars in the background because they were measured multiple times.

4 Evaluation of the Quality of the Dataset

In order to measure the quality of our dataset, we labeled a subset of 75283713 3d points manually. However, due to unbalanced occurrence and the different physical size of the classes the ground truth is highly unbalanced.



Fig. 7: (a) distribution of classes in the dataset

(b) distribution of classes in the ground truth

Figure 7(a) shows the relative frequency of each class in the dataset. It covers all classes that are included in the evaluation of cityscapes. In the ground truth (figure 7(b)) the classes sky, motorcycle, truck, train and bus are missing. It is obvious that we cannot label any points as sky, because this class is not measurable with a laserscanner, therefore it must be label noise and can be excluded from the dataset. As can be seen in figure 7(a), the other four classes are extremely rare, so we didn't find enough (or any) examples and had to exclude these classes from the evaluation.

We measured the quality of the dataset by evaluating the IoU which is calculated from the wellknown true/ false positives/ negatives TP, FP, FN as follows:

(1)
$$IoU = \frac{TP}{TP+FP+FN}$$

In order to extract a class c from a histogram h_i we used the majority class $c = argmax(h_i)$. The IoU is first calculated per category, which gives us an idea of how well each category intersects with the ground truth. We estimated the overall label noise by calculating the mean IoU. The results are shown in the following figures. In order to compare them to the performance of the pretrained network, we included the IoU results of PSPnet from cityscapes.



Fig. 8: (a) IoU with majority label

(b) IoU if at least one point in the histogram intersects with the ground truth (best)

Figure 8(a) shows the estimated IoU per category for our dataset compared to the IoU of PSPNet. The mean IoU in our dataset is \sim 35% and the mean IoU of PSPnet is 81.2%. In figure 8(b) we estimated the "best" case in which at least one value intersects with the ground truth. In contrast to the majority label, a true positive is counted as valid if the histogram contains a value greater zero wherever it intersects with the ground truth class. The mean IoU rises in the "best" case to about \sim 57%. We would like to note that in this case the IoU for terrain rises above the one of PSPNet. It could be possible that the aggregation through different pictures leads to a denoising effect. In the next step we calculated a confusion matrix for the majority labels.



Fig. 9: (a) Confusion matrix normalized per column, (b) Confusion matrix normalized per row

The table in Figure 9(a) is normalized per column therefore the diagonal axis contains the precision of the dataset. Figure 9(b) is normalized per row, thus the diagonal axis contains the recall of the dataset per class. As can be seen in Figure 9(a), moving objects are often confused with their surroundings. We believe that this type of noise is due to the difference in capture time. For example the prediction of the class car is to 44% true but in 43% of the cases it is road. Also the classes person and bicycle are often confused with sidewalk. Except fence, most static object are predicted with a high precision. As shown in Figure 9(b), the recall differs in some cases. In many cases bigger objects are having a high recall for example road, buildings, vegetation, terrain and cars. Smaller objects like poles, traffic lights, persons and bicycles are lower in recall, which makes sense because it is harder to hit the right point in 3d if the object is small or moving. Additionally, classes with a low IoU like wall, fence and rider are generally having a bad recall.

5 Label noise cleaning

In order to rise the IoU of the dataset we are presenting different approaches. One way of dealing with label noise is to remove possibly wrong labels. One potential drawback is that this approach induces sparsity into the dataset which could lead to worse outcomes in a supervised learning task because spatial information are getting lost.

5.1 Filtering based on entropy

A possible measure for the uncertainty of labels is the entropy of the label histogram. A high entropy means that the histogram contains strong votes for different labels which could be an indicator for the presence of label noise. Histograms are removed if the entropy is above a certain threshold. The best IoU was calculated for zero entropy which means that there is only one peak

in the histogram. Therefore we are removing histograms if the entropy is not equal to zero. As shown in the following figure it worked well for some classes.



Fig. 10: Comparison of the IoU after the labels were filtered by an entropy threshold

The orange bars are showing that the IoU raised in some cases. The most significant difference was measured for the class cars. This class went from 37.9% to 61.5% IoU. However, many classes that were problematic in the first case dropped in IoU after the dataset was filtered. Furthermore by removing these histograms 27.5 % of the data in the test set was removed. The mean IoU raised from 34.9% to 35.2%. By using this filter only on classes that did raise in the IoU we are able to get an estimated mean IoU of 38.98%.

5.2 Filtering labels by learning the noise

Another way of estimating the noise is by learning it in a supervised manner. In order to separate the noisy labels from the good ones we introduced two classes $d = \{0,1\}$

(2)
$$d = \begin{cases} 1, \text{ if } l_i = argmax(h_i) \\ 0, \text{ if } l_i \neq argmax(h_i) \end{cases}$$

A histogram is labelled as 1 if the ground truth l_i is equal to the majority label in the histogram. Otherwise it is marked as zero and thus removed from the dataset. In order to learn the noise in the dataset we used a gradient boosting decision tree (GBDT) algorithm, LightGBM by Guolin, et al.. We separated 15% of ground truth as test set by picking the samples randomly. The hyperparameter tuning was done on the training set by using 3-fold cross validation. We optimized the parameter with random search by maximizing the area under the receiver operating characteristic curve (ROC AUC). The ROC AUC also known as c-statistic is a measure of goodness of fit for binary outcomes which is not sensitive to unbalanced classes. As shown in the following figure, we were able to improve the IoU for nearly every class.



Fig. 11: Comparison of the IoU after the labels were filtered by GBDT

We would like to note that the outcome for the training and test set were nearly identical - the mean IoU differs by only 0.2 %. By using the classifier we can raise the IoU to 48.77 %. In order to increase the IoU even further, more information is needed. We chose to add the voxel count v_i in order to distinguish between dynamic and static objects. We believe that the information of dynamics paired with a typical histogram can help to remove noisy labels. The training procedure is done in same manner as before. The following images show the difference between both classifiers.



Fig. 12: Comparison of different label noise removal techniques. (a) Shows the depicted scene from Figure 3 with no changes, we marked the bicycle driver as red. (b) Shows the noise removal using only the histogram as feature. In (c) we additionally used the voxel count v_i as feature

We used the scene from figure 6 to show the effect of using the voxel count v_i . A perfect classifier would remove the red marked cyclist completely. As shown in figure 12 (b), the classifier is able to remove some noisy labels but it fails to delete the cyclist when only the histogram information is available. As can be seen in figure 12 (c), we were able to remove almost all of the red points by incorporating the voxel count v_i . Still the classifier is able to preserve the correct classified (blue) cars. Furthermore, we were able to raise the mean IoU to about 51.22%. The following

figure shows how the IoU improved per class by using GBDT with voxel count against the filter based on the entropy.



Fig. 13: IoU per category after removing the faulty labels. We compared the cleaning by using the entropy (blue) vs. by learning the noise (orange).

5.3 Correction labels by learning the noise

Lastly we tried to "flip" the labels to the correct class instead of removing faulty histograms. We did this in the same manner as before by using GBDT. The hyperparameter tuning was done by maximizing the weighted F1-Score. Additionally to the histogram we included the voxel count as feature. By flipping the labels we were able to increase the mean IoU to about 63.79% which is only 18% worse than PSPNet performed on the testset in cityscapes. Interestingly, it is also better than the "best case" we presented in figure 8(b), which means that we can recover labels that are not even present in the histogram. In the following picture we are showing the qualitative improvement over the original label transfer.



Fig. 14: (b) shows that we were able to recover information that would otherwise be lost if we used the majority label (Fig. 14 (a)). As can be seen, the cyclist and some other points were flipped to the correct class.

5.4 Results

Lastly we are showing a table that compares all results we presented in this work.

Tab 1: Comparison of different noise reductions to PSPNet. The column "Original" contains the unchanged 3d data set. The next column "Entropy" shows the IoU if histograms without zero entropy are removed. "Removed" and "Flipped" are the learned approaches with GBDT including the voxel count.

IoU	Original	Entropy	Removed	Flipped	PSPNet
road	84.3	93.41	94.01	91.96	98.68
sidewalk	24.71	11.64	42.25	55.35	86.92
building	85.59	93.45	93.08	91.85	93.46
wall	16.15	4.89	31.13	52.5	98.68
fence	4.09	2.3	1.32	23.4	63.67
pole	20.6	14.65	31.62	45.15	67.67
traffic light	17.37	6.22	32.0	55.82	76.12
traffic sign	12.8	14.1	22.25	50.53	80.47
vegetation	78.23	87.85	87.4	86.51	93.63
terrain	56.99	57.73	78.22	75.34	72.2
person	12.51	10.74	31.47	59.56	69.32
rider	7.55	0.0	34.18	75.41	69.32
car	37.88	61.45	85.25	76.16	90.27
bicycle	29.91	34.79	52.91	53.48	63.5
Avg. IoU	34.91	35.23	51.22	63.79	81.19

We have shown that we were able to greatly improve the average IoU by learning the noise in the data set. The class "building" even reached nearly the original IoU. Some classes are better when the histograms are removed instead of being flipped. We could therefore imagine that a classifier that decides whether to remove the label or flip it could even improve the IoU further by taking the best of both worlds.

6 Conclusion

We have presented a framework for label-transfer from the image domain to point clouds. After identifying origins of label noise, we created new features in order to improve the IoU of our dataset. Our Framework has the potential to create very large amounts of labelled point cloud data. This data can be used to train deep neural networks in order to learn semantic segmentation in 3d. For future work, we would like to improve the IoU by using more advanced techniques for noise detection. We can imagine that by using a spatial neighborhood, the label cleaning can be improved. Additionally it could be helpful to adapt the pre trained neural network to our cameras by using adversarial discriminative domain adaptation (adda) (Tzeng et al. 2017). The improved classification error would minimize the label noise in an early stage of our framework.

7 Acknowledgements

This work was funded by the German Research Foundation (DFG) as a part of the Research Training Group GRK2159, 'Integrity and collaboration in dynamic sensor networks' (i.c.sens).

8 References

- BOULCH, A., GUERRY, J., LE SAUX, B. & AUDEBERT, N., 2018: SnapNet: 3D point cloud semantic labeling with 2D deep segmentation networks. Computers & Graphics, 71, 189-198.
- BRENNER, C., 2016: Scalable estimation of precision maps in a MapReduce framework. In International Conference on Advances in Geographic Information Systems, 27.
- CHEN, L.C., PAPANDREOU, G., KOKKINOS, I., MURPHY, K. & YUILLE, A.L., 2018: Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Transactions on Pattern Analysis and Machine Intelligence, **40**(4), 834-848.
- CORDTS, M., OMRAN, M., RAMOS, S., REHFELD, T., ENZWEILER, M., BENENSON, R., FRANKE, U., ROTH, S. & SCHIELE, B., 2016: The cityscapes dataset for semantic urban scene understanding. IEEE Conference on Computer Vision and Pattern Recognition, 3213-3223.
- DRORY, A., AVIDAN, S. & GIRYES, R., 2018: On the Resistance of Neural Nets to Label Noise. arXiv preprint arXiv:1803.11410.
- FRÉNAY, B. & VERLEYSEN, M., 2014: Classification in the presence of label noise: a survey. IEEE Transactions on Neural Networks and Learning Systems, 25(5), 845-869.
- HENDRYCKS, D., MAZEIKA, M., WILSON, D. & GIMPEL, K., 2018: Using Trusted Data to Train Deep Networks on Labels Corrupted by Severe Noise. arXiv preprint arXiv:1802.05300.
- HERMANS, A., FLOROS, G. & LEIBE, B., 2014: Dense 3d semantic mapping of indoor scenes from rgb-d images. IEEE International Conference on Robotics and Automation, 2631-2638.
- KE, G., MENG, Q., FINLEY, T., WANG, T., CHEN, W., MA, W., QIWEI YE, Q. & LIU, T. Y., 2017: Lightgbm: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems, 3146-3154.
- KOCHANOV, D., OŠEP, A., STÜCKLER, J. & LEIBE, B., 2016: Scene flow propagation for semantic mapping and object discovery in dynamic street scenes. IEEE/RSJ International Conference on Intelligent Robots and Systems, 1785-1792.
- LEE, K.H., HE, X., ZHANG, L. & YANG, L., 2017: CleanNet: Transfer Learning for Scalable Image Classifier Training with Label Noise. arXiv preprint arXiv:1711.07131.
- LONG, J., SHELHAMER, E. & DARRELL, T., 2015: Fully convolutional networks for semantic segmentation. IEEE Conference on Computer Vision and Pattern Recognition, 3431-3440.
- MCCORMAC, J., HANDA, A., DAVISON, A. & LEUTENEGGER, S., 2017: Semanticfusion: Dense 3d semantic mapping with convolutional neural networks. IEEE International Conference on Robotics and Automation, 4628-4635.
- NETTLETON, D.F., ORRIOLS-PUIG, A. & FORNELLS, A., 2010: A study of the effect of different types of noise on the precision of supervised learning techniques. Artificial intelligence review, 33(4), 275-306.

- PATRINI, G., ROZZA, A., MENON, A.K., NOCK, R. & QU, L., 2017: Making deep neural networks robust to label noise: A loss correction approach. IEEE Conference on Computer Vision and Pattern Recognition, 2233-2241.
- ROLNICK, D., VEIT, A., BELONGIE, S. & SHAVIT, N., 2017: Deep learning is robust to massive label noise. arXiv preprint arXiv:1705.10694.
- TZENG, E., HOFFMAN, J., SAENKO, K. & DARRELL, T., 2017: Adversarial discriminative domain adaptation. Computer Vision and Pattern Recognition, 1(2), 4.
- XIAO, T., XIA, T., YANG, Y., HUANG, C. & WANG, X., 2015: Learning from massive noisy labeled data for image classification. IEEE Conference on Computer Vision and Pattern Recognition, 2691-2699.
- XIE, J., KIEFEL, M., SUN, M.T. & GEIGER, A., 2016: Semantic instance annotation of street scenes by 3d to 2d label transfer. IEEE Conference on Computer Vision and Pattern Recognition, 3688-3697.
- ZHANG, R., LI, G., LI, M. & WANG, L., 2018. Fusion of images and point clouds for the semantic segmentation of large-scale 3D scenes based on deep learning. ISPRS Journal of Photogrammetry and Remote Sensing, 143, 85-96.
- ZHAO, H., SHI, J., QI, X., WANG, X. & JIA, J., 2017: Pyramid scene parsing network. IEEE Conference on Computer Vision and Pattern Recognition, 2881-2890.

3D-Punktwolken in PosgreSQL/PostGIS für Anwendungen in Geoinformationssystmen

THERESA MEYER¹ & ANSGAR BRUNN¹

Zusammenfassung: Neben dem dateibasierten Management von 3D-Punktwolken besteht seit einigen Jahren auch die Möglichkeit, diese gemeinsam mit anderen Geodaten in räumlichen Datenbanksystemen vorzuhalten und dabei von den Vorteilen der kombinierten Analyse von georeferenzierten Punktwolken-, Raster- und Vektordaten zu profitieren. In diesem Beitrag wird vorgestellt, wie die Punktwolkenerweiterung "Pointcloud" für die PosgreSQL/PostGIS Datenbank mit ihren speziellen Methoden und der Integration von PostGIS-Funktionen genutzt werden kann, um auch räumliche Abfragen in 3D zu realisieren. Im Besonderen werden erweiterte Möglichkeiten zur Punktwolkenkachelung durch Kombinationen von Funktionen der Point Data Abstraction Library (PDAL) aufgezeigt und dahingehend untersucht, in wie weit sich unterschiedliche Kachelmethoden und -konfigurationen auf Speicherplatzbedarf, Abfrageperformance und Interaktionen zwischen Punktwolken und anderen 2D und 3D Geoobjekten in der Datenbank auswirken. Die Ergebnisse zeigen, dass eine individuell gewählte Speicherstruktur für eine Punktwolke unter Berücksichtigung ihrer dreidimensionalen räumlichen Ausdehnung und Punktdichte ausschlaggebend für einen geringen Speicherverbrauch und performante Abfragen in PostGIS-Anwendungen ist.

1 Einleitung

Als Quelle für räumliche Informationen gewinnen 3D-Punktwolken in Anwendungsfeldern der Geodäsie, Geoinformatik, Architektur und Bauarchäologie zunehmend Bedeutung. Obwohl moderne Messinstrumente wie Laserscanner, Unmanned-Aerial-Systems (UAS) und Mobile Mapping Systeme die Erzeugung von Punktwolken immer mehr vereinfachen und eine große Auswahl von anwendungsspezifischer Auswertesoftware zur Verfügung steht, stellt die effiziente Verwaltung von charakteristisch großen Punktwolkendaten im Alltagsbetrieb von Vermessungsund Ingenieurbüros eine Herausforderung dar. Punktwolken dienen als Grundlage zur Lösung verschiedenster ingenieurtechnischer Aufgaben und haben schon lange Einzug in die Projekte auch von kleinen und mittleren Unternehmen gehalten. Sie werden z. B. zu Zwecken der BIM-konformen As-Built-Dokumentation von Gebäuden und Anlagen erzeugt (TANG et al. 2010), für Veränderungsdetektionen und Deformationsanalysen von Ingenieurbauwerken (MUKUPA et al. 2016) oder sie dienen zur hochauflösenden, dreidimensionalen Dokumentation von Urgelände (BURGER et al. 2015), weshalb sie mittlerweile auch zu einer unverzichtbaren Planungsgrundlage für Bauprojekte aller Art geworden sind.

Für die Verarbeitung, die Speicherung und den Austausch zwischen Anwendungsprogrammen werden für Punktwolken üblicherweise eigens entwickelte Dateiformate wie LAS (ASPRS 2013) PCD (RUSU & COUSINS 2011) oder E57 (HUBER 2011) genutzt, die es ermöglichen neben X-, Y- und Z-Koordinaten auch zusätzliche Punktattribute wie *Intensity*, *Return Number*, *Red*, *Green*,

¹ Hochschule für angewandte Wissenschaften Würzburg - Schweinfurt, Fakultät Kunststofftechnik und Vermessung, Röntgenring 8, D-97070 Würzburg, E-Mail: [Theresa.Meyer, Ansgar.Brunn]@fhws.de

Blue, NIR, Classification, Point Source ID verlustfrei komprimiert zu transportieren. Alternativ zur dateibasierten Verwaltung gibt es seit einigen Jahren jedoch auch die Möglichkeit 3D-Punktwolken gemeinsam mit anderen Geodaten in räumlichen Datenbanksystemen ("Geodatenbanken") vorzuhalten und dabei von der kombinierten Analyse von georeferenzierten Punktwolken-, Raster- und (2D- und 3D-) Vektordaten zu profitieren. Vorteil einer (Geo-) Datenbank-Lösung ist neben der zentralen Datenhaltung die Tatsache, dass räumliche und nicht räumliche Daten aller Art, wie Punktwolken, Rasterdaten, Vektordaten (z. B. Katasterdaten, Bestands- und Planungsdaten in Form von CAD- und BIM-Modellen), Fachdaten und administrative Daten jederzeit miteinander in Beziehung gesetzt werden können.

Dieser Beitrag handelt von 3D-Punktwolken in ihrer Rolle als Objekte in räumlichen Datenbanksystemen und speziell der Punktwolken-Erweiterung "Pointcloud", die seit dem Jahr 2013 für das freie und quelloffene Geodatenbanksystem PostgreSQL/PostGIS entwickelt wird. Pointcloud erweitert PostGIS um neue Datentypen und eine Reihe eigener Methoden, die es ermöglichen 3D-Punktwolken in Datenbanktabellen zu speichern und mittels GIS-Tools zu analysieren. Im Besonderen geht es um die Möglichkeiten der zwei- und dreidimensionalen Kachelung von 3D-Punktwolken mittels Funktionen der Point Data Abstraction Library (PDAL). Anhand eines beispielhaften PostGIS/Poincloud-Workflows für Punktwolken-Abfragen im dreidimensionalen Raum, werden die Ergebnisse von Untersuchungen der Auswirkungen unterschiedlicher Kachel-Methoden und -Konfigurationen auf Speicherplatzbedarf, Abfrageperformance und Interaktionen zwischen Punktwolken und Vektordaten präsentiert. Es wird außerdem eine Auswahl von open-source Tools und Programmcode-Bibliotheken für die Bearbeitung und Analyse von 3D-Punktwolken vorgestellt und skizziert wie sich diese in datenbankbasierte Arbeitsabläufe integrieren lassen. Der Artikel schließt mit einer Einschätzung, in wie weit sich das Geoobjekt "3D-Punktwolke" in die Strukturen der PostGIS-Datenbank integrieren lässt und einem Blick auf alternative Lösungen auch im Bereich kommerzieller Software.

2 3D-Punktwolken in GIS

Punktwolken werden z. B. für die Generierung von 3D-Gebäudemodellen zum Aufbau von 3D-Stadtmodellen (HAALA & BRENNER 1997), zur Vegetationsdetektion in urbanen Gebieten (HöFLE & HOLLAUS 2010), als Grundlage für Gelände- und Objekt-Modellierungen bei Hochwassersimulationen (MERWADE et al. 2008) oder für die Extraktion von Dach- und Fassadenflächen für Solarpotentialanalysen (VöGTLE et al. 2005; JOCHEM et al. 2011) benötigt. Im Umgang mit Punktwolken sind zwei grundsätzliche Fälle zu unterscheiden. Lassen sich die Punkthöhen (Z-Koordinaten) als Funktion der Lage modellieren besitzt, eine Punktwolke die charakteristische Eigenschaft von 2,5D Objekten, bei denen für jede Kombination von X und Y nur ein einziger Z-Wert existiert. Solche Punktwolken resultieren i. d. R. aus luftgestützten Erfassungsmethoden und können auch im Rasterformat gespeichert sowie mit herkömmlichen Raster-Tools in einem GIS verarbeitet werden. Handelt es sich dagegen um unstrukturierte Punktwolken mit ausgeprägtem 3D-Charakter, wie sie z. B. aus terrestrischem oder mobilem Laserscanning hervorgehen, entsteht der Bedarf nach "richtigen" 3D-Speicherstrukturen und –

Operationen, da eine Rasterkonvertierung einen enormen und irreversiblen Informationsverlust zur Folge hätte.

Allgemeiner Standard für GIS-Anwendungen in 3D ist das *Feature-Geometry-Modell*, das dreidimensionale Koordinaten erlaubt und eine Reihe von Klassen für 3D-Oberflächen, Dreiecksvermaschungen (TINs), 3D-Körper und 3D-Topologie bereitstellt. Auch das erweiterte Datenbankschema *SQL/MM Spatial* unterstützt seit der dritten Version dreidimensionale Koordinaten und umfasst, im Gegensatz zum *Simple-Feature-Modell*, entsprechende 3D-Methoden. Während es bewährte und standardisierte Modelle und Strukturen für Raster- und vor allem Vektordaten in Geodatenbanksystemen gibt, ist die Art und Weise wie die Integration von originalen 3D-Punktwolken zu handhaben ist jedoch noch nicht spezifiziert oder standardisiert. Derzeit widmet sich die OGC-Arbeitsgruppe "Point Cloud" der Definition von Anforderungen an Interoperabilität und der Implementierung von Standards im Anwendungsfeld von 3D-Punktwolken (OGC 2018).

3 Pointcloud-Erweiterung für PostgreSQL/PostGIS

Obwohl PostGIS mit POINT Z einen Geometrietyp für Punkte mit dreidimensionalen Koordinaten bereitstellt, ist dieser Datentyp für die Speicherung von 3D-Punktwolken ungeeignet. Grund hierfür ist schlicht die Tatsache, dass die Anzahl von Punkten in einer Punkwolke aus z. B. terrestrischem Laserscanning mit bis zu mehreren Milliarden Punkten viel zu groß ist, als dass jeder Punkt in einer einzelnen Tabellenzeile gespeichert werden könnte. Für die Speicherung von 3D-Punktwolken bei gleichzeitig effizientem Zugriff auf alle Punktattribute in einer PostgreSQL/-PostGIS Datenbank wurde deshalb die Erweiterung Pointcloud entwickelt. Sie erweitert PostGIS um die zwei neuen Objekttypen PcPoint und PcPatch. Räumlich benachbarte Punkte werden zu mehreren hunderten bis tausenden als sog. Patches (Kacheln/ Blöcke) in Form von Byte-Arrays organisiert. Jedes Punktwolken-Patch entspricht einer Instanz der Klasse PcPatch und wird in einer eigenen Tabellenzeile gespeichert. Die inhaltliche Struktur von Patches wird durch ein XML-Schemadokument beschrieben. Es enthält genaue Informationen darüber, welche Punkt-Attribute an welcher Stelle und mit welchem Datentyp vorliegen. Beispielsweise könnte ein Intensitätswert je nach Auflösung als 4-Byte Integer oder in einem einzelnen Byte gespeichert und X-, Y- und Z-Koordinaten durch Double-Werte oder als 4-Byte Integer mit Skalierungsfaktor repräsentiert werden.

3.1 Erweiterte Punktwolkenkachelung mit PDAL-Funktionen

Punktwolken werden beim Schreiben in die Postgres-Datenbank praktischerweise in Patches gekachelt, damit sie in Tabellen komprimiert gespeichert und gleichzeitig effizient abgefragt werden können. Die Punktwolkenkachelung wird standardmäßig während des Imports im Zuge der Ausführung von Kommandozeilen-Anweisungen und Funktionen der *Point Data Abstraction Library* (PDAL) realisiert. PDAL stellt zwei grundsätzliche Funktionen zur Punktwolken-kachelung bereit:

Ungleichmäßige Kachelung mit *filters.chipper()*:
Diese Funktion bündelt benachbarte Punkte in einzelne nicht überlappende Patches. Die

maximale Anzahl von Punkten in einem Patch wird durch einen Vorgabewert für den Funktions-Parameter *capacity* festgelegt. Als Standardwert sind 5000 Punkte pro Kachel vorgesehen (PDAL, 2018). Die Punkte werden so sortiert, dass eine lineare Ordnung im zweidimensionalen Raum entsteht. Die Z-Koordinaten bleiben gänzlich unberücksichtigt. Die sortierte Punktliste wird dann in einzelne Patch-Arrays gemäß der vorgegebenen Kapazität unterteilt. Der Algorithmus prüft für jedes Array dessen Ausdehnung in X- und Y-Richtung und unterteilt neu, falls diese zu unausgewogen ist.

- Gleichmäßige Gitter-Kachelung mit *filters.splitter()*:
 - Bei Anwendung dieser Funktion werden Punktwolkenpunkte in quadratische Patches, mit einer vom Anwender vorgegebenen Seitenlänge (in Meter) gebündelt. Der Kachelursprung wird entweder zufällig vom Algorithmus gewählt oder kann in Form einer X- und Y-Koordinate festgelegt werden. Auf diese Weise entstehen Patches mit derselben Ausdehnung in der X/Y-Ebene, die sich jedoch in der Anzahl ihrer Punkte unterscheiden.

Unabhängig von der gewählten Kachel-Methode und Patch-Größe muss sich der Anwender der Tatsache bewusst sein, dass die Punktwolken-Kachelung mit *filters.chipper()* und *filters.splitter()* in beiden Fällen keine 3D- sondern vielmehr eine 2D-Kachelung ist. Die individuelle Ausdehnung einer Punktwolke entlang der Z-Achse wird nicht berücksichtigt. Das hat zur Folge, dass auch Punkte, die aufgrund unterschiedlicher Z-Werte im 3D-Raum unter Umständen sehr weiter auseinanderliegen dennoch demselben Patch-Array zugeordnet werden. Besonders für Punktwolken mit ausgeprägtem 3D-Charakter können auf diese Weise sehr unregelmäßige und unausgewogene Patches entstehen (Abb. 1). Um Punktwolkenpatches auch in ihrer Ausdehnung entlang der Z-Achse zu limitieren, bietet sich ein neuartiger Kachel-Prozess unter Hinzunahme der PDAL-Funktion *filters.range()* an. Die Funktion ist ein Bereichsfilter, der in Anwendung für die Dimension ,Z' eine Punktwolke in beliebig viele Höhenschichten (z. B. 1 m oder 5 m) unterteilt. In Kombination mit *filters.chipper()* und *filters.splitter()* kann *filters.range()* dazu verwendet werden eine um ,Z' erweiterte dreidimensionale Punktwolkenkachelung zu erzeugen (Abb. 2).



Abb. 1: Kachelung einer Punktwolke mit 3 Millionen Punkten mit filters.chipper() und Kapazität 100



Abb. 2: Kachelung einer Punktwolke mit 3 Millionen Punkten mit filters.chipper(), Kapazität 100 und Beschränkung auf 1m entlang der Z-Achse

3.2 Pointcloud-Funktionen und Integration von PostGIS

Sämtliche Funktionen der Pointcloud-Erweiterung beziehen sich auf die beiden zentralen Pointcloud-Objekte PcPoint und PcPatch. PcPoints spielen insgesamt eine eher untergeordnete Rolle, denn es ist zwar möglich PcPoint-Tabellen nach gültigem Schema anzulegen und mit Punkten zu füllen, allerdings werden diese Einzelpunkte in erster Linie nur für spezielle Abfragen, z. B. als Zwischenschritt zur Konvertierung einer kleinen Ergebnismenge von Punktwolken-Punkten in PostGIS-Punkte vom Typ POINT Z benötigt. Ganze Punktwolken werden in PcPatch-Tabellen verwaltet und zu diesem Zweck stellt Pointcloud eine Reihe spezieller Funktionen bereit. Dazu gehören u. a. solche zum Erzeugen und Auflösen, Umsortieren und Komprimieren/

Dekomprimieren von Patches (z. B. *PC_MakePatch()*, *PC_Union()*, *PC_Sort()* und *PC_Compress()*) oder Funktionen zum Abruf von genauen Informationen über deren Inhalt (z. B. *PC_Summary()*, *PC_NumPoints()* oder *PC_PatchMax()* und *PC_PatchMin()*). Darüber hinaus gibt es die Pointcloud-Funktionen *PC_FilterGreaterThan()*, *PC_FilterLessThan()*, *PC_FilterBetween()* und *PC_FilterEquals()*, zum Filtern von Punkten auf Basis von Attributwerten wie Intensität, Klassifizierungscode, Z-Wert oder Zeitstempel.

Erst mit der zusätzlichen Erweiterung *Pointcloud_PostGIS* werden Interaktionen zwischen Pointcloud- und PostGIS-Objekten möglich. Unter Verwendung der Funktion *PC_Intersects()* können Patches dahingehend überprüft werden, ob sie die Geometrie eines PostGIS-Objekts überschneiden. Mit *PC_Intersection()* werden neue Patches erzeugt, die die Ergebnismenge einer Überschneidung zwischen einem PcPatch und z. B. einem Polygon darstellen. Sowohl *PC_Intersects()* als auch *PC_Intersection()* akzeptieren ausschließlich 2D-Objekte und ermöglichen deshalb keine räumlichen Abfragen in 3D. Die Funktion *PC_EnvelopeGeometry()* gibt die zweidimensionale Begrenzung eines Patches als PostGIS-Polygon zurück und *Pc_BoundingDiagonalGeomery()* ermöglicht durch die Rückgabe der 3D-Boundingbox-Diagonalen eines Patches als LINESTRING Z den Aufbau eines 3D-Indexes.



Abb. 3: Anwendung der Pointcloud_PostGIS-Funktion PC_Intersection()

3.3 Exemplarischer Workflow einer mehrstufigen Abfrage in 3D

Für Intersection-Abfragen mit PcPatch-Objekten bietet Pointcloud in Kombination mit der Pointcloud_Postgis-Erweiterung nur die Methoden *PC_Intersects()* und *PC_Intersection()*. Beide Funktionen akzeptieren aber nur PcPatch- und zweidimensionale PostGIS-Objekte und keine 3D-Körper vom Typ POLYHEDRALSURFACE Z. Somit können Pointcloud-Punkte zunächst nur projiziert auf die Grundrissebene abgefragt werden. Für die Selektion einer Punktwolken-Teilmenge, die sich innerhalb eines 2,5D-Körpers im 3D-Raum befindet, kann die Beschränkung auf den zweidimensionalen Raum durch eine mehrstufige Abfrageprozedur umgangen werden:

- 1. Mit der Pointcloud-Funktion *PC_BoundingDiagonalGeometry()* werden die Diagonalen der 3D-Boundingboxen aller Patches als PostGIS 3D-Linien vom Typ LINE Z erzeugt.
- 2. Um performante Abfragen in 3D zu ermöglichen wird für die Diagonalen bzw. deren 3D-Boundingboxen ein 3D-Index aufgebaut.

- 3. Durch Anwendung des Index-Operators &&& werden die 3D-Boundingboxen aller Patch-Diagonalen selektiert, welche die 3D-Boundingbox eines beliebigen 3D-Objekts vom Typ POLYHEDRALSURFACE Z überschneiden.
- 4. Die 3D-Boundingboxen der Patch-Diagonalen werden mit der PostGIS-Funktion *Box3D()* in Volumenkörper konvertiert, sodass der exakte Test auf Überschneidung unter Verwendung der *Operation ST_3DIntersects()* ausgeführt werden kann (nur mit SFCGAL Backend).



Abb. 4: Ergebnis von ST_3DIntersects()

- 5. Mit einem Inner Join über ein mitgeführtes Schlüsselattribut (ID) werden die selektierten Volumenkörper auf die ursprünglichen PcPatch-Objekte zurückgeführt.
- 6. Die vorselektierten Punktwolken-Patches werden erst jetzt mit Hilfe der speziellen Pointcloud-Funktion *PC_Intersection()* auf die Begrenzungen des Abfrage-Objekts in X und Y "zugeschnitten".



Abb. 5: Ergebnis von PC_Intersection()

7. Je nach Kachel-Methode verbleiben in den Patches immer noch 3D-Punkte, die ober- bzw. unterhalb des Abfrage-Objekts liegen. Für Abfragen mit 2,5D-Objekten bietet sich deshalb die Pointcloud-Funktion *PC_FilterBetween()* an. Mit ihr lässt sich die minimale und maximale Z-Ausdehnung der finalen Ergebnis-Patches exakt beschränken.



Abb. 6: Ergebnis von PC_FilterBetween()

Als Ergebnis dieser Abfragen liegt eine Auswahl von Punktwolken-Patches vor, welche ausschließlich 3D-Punkte beinhalten, die im Inneren oder auf der Außenseite eines Abfragekörpers

im 3D-Raum liegen. Es könnte sich dabei z. B. um ein Szenario handeln, bei dem eine Punktwolke und ein 3D-Gebäudemodell interagieren. Das Abfrageobjekt könnte einen Wandkörper im Inneren eines Gebäudes darstellen und es sollen alle 3D-Punkte aus einem terrestrischen Laserscan selektiert werden, die sich auf diesem Wandkörper befinden, damit sie getrennt vom Rest der Punktwolke weiterverarbeitet und analysiert werden (z. B. um Ausrichtung und Neigung as-built zu ermitteln und den Planungsdaten gegenüberzustellen). Auswahlpunkte können mit PDAL aus der Datenbank in ein Punktwolkenformat exportiert, mit externen Funktionen und direkter Datenbankanbindung weiterverarbeitet oder als PostGIS-Punkte vom Typ POINT Z konvertiert werden. Letzteres ermöglicht auch punktgenaue Abfragen mit beliebigen 3D-Objekten in PostGIS, funktioniert praktisch aber nur für unverhältnismäßig kleine Selektionsmengen von 3D-Punkten, die kaum noch dem Wesen einer 3D-Punktwolke entsprechen.

3.4 Vor- und Nachteile unterschiedlicher Kachel-Methoden

Die Entscheidung für eine Bestimmte Punktwolken-Patch-Größe und Kachel-Methode wird dem Nutzer überlassen. In der PDAL-Dokumentation gibt es nur die Empfehlung, dass der Speicherplatzbedarf eines einzelnen Patches die standardmäßige Seitengröße von 8 KB einer PostgreSQL-Datenbank nicht überschreiten sollte und dass dies für die meisten LiDAR-Punktwolken mit Patchgrößen von 400 bis 600 Punkten noch erreicht wird (PDAL 2018). Der Standardwert bei Anwendung von *filters.splitter()* entspricht jedoch mit 5000 Punkten etwa dem 10-fachen dieser Empfehlung.

Für den Anwender stellt sich unweigerlich die Frage welche Kachel-Methode (regelmäßig oder unregelmäßig? 2D oder 3D?) am geeignetsten für welche Art von Punktwolke, Projekt-/ Untersuchungsgebiet und die konkrete Aufgabenstellung ist. Untersuchungen des Speicherplatzbedarfs bei Anwendung von filters.chipper() und filters.splitter() in verschiedenen Konfigurationen, haben gezeigt, dass die Punkt-Kapazität bei Anwendung von filters.chipper(), bzw. die Gitterweite bei filters.splitter() möglichst groß sein sollte, wenn es unabhängig von weiteren Interaktionen und Operationen in der Geodatenbank nur um die verlustfreie Minimierung von Punktwolken-Daten geht. Je mehr 3D-Punkte in einem Patch gebündelt werden, desto weniger Patches entstehen insgesamt und desto effektiver wirkt die standardmäßige Kompressionsmethode "dimensional" innerhalb jedes einzelnen Patches. Bei dieser Kompressionsmethode werden die Punkt-Arrays nicht punkt- sondern dimensionsweise sortiert, sodass für jede Punktdimension/ Punktattribut die geeignetste Kompressionstechnik (z. B. Lauflängencodierung) angewandt werden kann (BOUNDLESS 2015).

Auch eine 3D-Kachelung wirkt sich deutlich positiv auf die Größe der Punktwolken-Tabellen in der Datenbank aus. Es ist anzunehmen, dass dies ebenfalls in der Kompression der Punkte innerhalb der Patches begründet liegt. Je ähnlicher die Ausprägungen von Punktmerkmalen wie Koordinaten, RGB- oder Intensitätswerte usw. sind, desto stärker lassen sich diese komprimieren. Räumlich eng benachbarte Punkte haben ähnliche Eigenschaften und Punkte, die innerhalb von "echten" 3D-Patches gebündelt werden liegen näher beieinander als 3D-Punkte, die gemeinsam in 2D-Patches organisiert werden. Vergleiche von 2D- und 3D-Kachelungen, mit und ohne Kompression haben das bestätigt.



Abb. 7: Auswirkung unterschiedlicher Kachelmethoden in 2D und 3D auf den Speicherplatzbedarf von Pointcloud-Tabellen für eine Beispielpunktwolke (ca. 33 Millionen 3D-Punkte mit RGB-Werten).

Werden Abfrageprozeduren für dieselbe Punktwolke in verschiedenen Kachelungen durchlaufen lassen sich auch Unterschiede hinsichtlich Ausführungsdauer und Ergebnismenge erkennen. Dabei ist zwischen Abfragen auf Patch-Level und solchen auf Einzelpunkt-Level zu unterscheiden. Vorselektionen von ganzen Patches auf Basis von Index-Operatoren und approximierten Geometrien sind als erste Näherungen zu betrachten (Stufe 1). In einer nächsten Verfeinerungsstufe wird die Kandidatenmenge von Patches dann mit der originalen Geometrie eines Abfrageobjekts betrachtet (Stufe 2) und erst im letzten Schritt mit Pointcloud-Funktionen auf Punkt-Level-Genauigkeit ausgewählt (Stufe 3) (vgl. Workflow aus Kap. 4.3). Bei jeder Kachelmethode ist die Auswahlmenge nach Stufe 1 am größten und wird dann in Stufe 2 bis zur finalen Ergebnismenge nach dem Durchlauf von Stufe 3 weiter reduziert. Die gleiche Abfrageprozedur liefert für jede Kachelmethode letztendlich dieselbe Punktmenge, dennoch sind deutliche Unterschiede in den Zwischenergebnissen nach einzelnen Filterstufen erkennbar, die sich auf die Performance von Anwendungen in der Praxis auswirken können. Je kleiner die Anzahl von Punkten in einem Patch, desto feiner die Kachelung und desto präziser sind die Ergebnismengen aber desto größer ist die Gesamtanzahl von Patches und der Aufwand für einzelne Operationen in der Datenbank.

Die Größe von Unterschieden beim Vergleich von verschiedenen 2D- und 3D-Kachelmethoden ist abhängig von der individuellen Punktwolken-Geometrie (vor allem vom Verhältnis von Ausdehnung in der X/Y-Ebene und zur Ausdehnung entlang der Z-Achse) und von der, des Abfrage-Objekts. Die Diagramme in Abbildung 8 zeigen die deutlichen Unterschiede in den Ergebnismengen der Filterstufen 1 und 2 für eine Auswahl von 3D-Punkten innerhalb eines Quaders im 3D-Raum aus einer Punktwolke mit ca. 3 Millionen Punkten. Diese Beispiel-Punktwolke erstreckt sich über mehrere Stockwerke im Inneren eines Gebäudes (siehe auch Abb. 1 u. 2) und hat daher einen sehr ausgeprägten 3D-Charakter. Die Ergebnisse von 2D- und 3D-Kachelmethoden, jeweils mit unterschiedlich großen Patches zeigen, dass sich eine 3D-Kachelung für möglichst präzise Ergebnismengen von Anfragen auf Patch-Level grundsätzlich besser geeignet ist als die standardmäßige 2D-Variante.



Abb. 8: Auswirkung unterschiedlicher Kachelmethoden in 2D und 3D auf die Ergebnismengen von Punktanfragen der Stufen 1 und 2 für eine Beispielpunktwolke mit ca. 3 Millionen Punkten.

4 Alternative Lösungen zur Integration von 3D-Punktwolken in GIS

Die Aufbereitung und Weiterverarbeitung von originalen 3D-Punkwolken ist keine klassische GIS-Aufgabe, sondern wird zu großen Teilen in spezieller Software für Punktwolken, Laserscanning, digitaler Photogrammetrie und/ oder 3D-Modellierung umgesetzt. *CloudCompare* (CLOUDCOMPARE, 2018) ist beispielsweise eine sehr vielseitige open-source Punktwolken-Software, die jedoch keine Programmier- oder Datenbankschnittstelle bereitstellt. Für eine hersteller- und programmunabhängige Aufbereitung, Verarbeitung und Analyse von Punktwolken existieren eine Reihe von freien Bibliotheken, die u. a. dafür genutzt werden können, den Funktionsumfang von offenen datenbankbasierten Punktwolken-Managementsystemen zu erweitern.

Im Zusammenhang mit PostGIS/Pointclound ist an vorderster Stelle die PDAL-Bibliothek zu nennen. Sie realisiert den praktischen Datenbankimport von Punktwolken durch Verarbeitungsanweisungen (sog. *PDAL-Pipelines*), die in JSON-Syntax geschrieben und über die Kommandozeile ausgeführt werden. PDAL ermöglicht es im Zuge einer Pipeline-Ausführung zusätzliche sog. *Filter*-Funktionen auf die Daten anzuwenden. Dazu gehören z. B. Berechnungen aus Punktnachbarschaften wie Normalenvektoren, Eigenwerte oder Voxel-Schwerpunkte. Andere Funktionen dienen dazu Punktwolken mit georeferenzierten Bildern zu colorieren, Dreiecksnetze

zu erstellen oder Punktwolken in andere Bezugssysteme zu transformieren. Es existiert auch eine Verbindung zwischen PDAL und der *Point Cloud Library* (PCL) (PCL, 2018), deren Fokus noch mehr auf der Entwicklung von neuen Algorithmen zur Verarbeitung von Punktwolken liegt. PDAL unterstützt außerdem die Python-Programmiersprache, denn eigene Python-Funktionen sind auf einfache Weise in PDAL-Pipelines einzubinden und mit der *PDAL*-Erweiterung für Python ist es wiederum auch möglich PDAL-Funktionen in Python-Anwendungen auszuführen. Eine weitere Programmbibliothek speziell für 3D-Daten und Punktwolken ist *Open3D* (ZHOU et al. 2018), die neben grundlegenden Verarbeitungsalgorithmen auch Möglichkeiten zur Visualisierung von Punktwolken inklusive Animationen umfasst. Open3D-Funktionen sind ebenfalls über eine Python-Erweiterung nutzbar und damit auch für die Verarbeitung von Punktwolken in der PostGIS-Datenbank. Die *LAStools* (RAPIDLASSO 2018) (nur teilweise frei) und *LibLAS*-Bibliothek (LIBLAS 2018) sind speziell für den Umgang mit LiDAR-Punktwolken im LAS-Format entwickelt. Programmbibliotheken können 3D-Punktwolken um zusätzliche Punkt-Attribute wie z. B. Klassifizierungscodes erweitern und damit den Informationsgehalt innerhalb einer Datenbank erhöhen.

Auch im Bereich proprietärer und kommerzieller Software gibt es Lösungen zur Integration von 3D-Punktwolken in Geoinformationssysteme und Geodatenbanken. In *ESRI's ArcGIS*-Produkten können Punktwolken im LAS-Format mit Hilfe von *LAS-Datasets* verwaltet, analysiert und angezeigt werden (ESRI 2016). In einem solchen Dataset sind Verweise auf mehrere, miteinander in Beziehung stehende LAS-Punktwolken in einem beliebigen Dateiverzeichnis gespeichert. Dadurch ist ein zentraler Zugriff auf Punktwolken, die verteilt in einzelnen Dateien vorliegen möglich. In Kombination mit Werkzeugen und Funktionspaletten zur Punktwolkengenerierung aus Luftbilddaten (den sog. *Ortho Mapping Tools*), deckt ArcGIS als ein einzelnes System bereits große Teile des Gesamtworkflows zur Erzeugung, Verarbeitung und GIS-basierten Analyse von z. B. UAV-Produkten ab. Die *LAStools*-Software-Suite von *Rapidlasso* umfasst eine große Auswahl von Punktwolken-Kommandozeilen-Tools, die auch als Plugin den Funktionsumfang von ArcGIS, *QGIS* und *Erdas Imagine* erweitern können und ebenfalls auf einzelne LAS-Dateien referenzieren (RAPIDLASSO 2018).

Ein ähnliches Produkt ist *OPALS* (<u>O</u>rientation and <u>P</u>rocessing of <u>A</u>irborne <u>L</u>aser <u>S</u>canning data), entwickelt von der TU Wien (OPALS 2017a). OPALS unterstützt verschiedenste Raster- und Vektorformate und fokussiert ebenfalls LiDAR-Punktwolken, auch wenn es nicht auf die Verarbeitung von solchen beschränkt ist. Die Software deckt mit ihren verschiedenen Modulen und Funktionen u. a. zur Georeferenzierung, Qualitätskontrolle, Klassifikation, Filterung und Geländemodellierung weitgehend die gesamte Punktwolken-Verarbeitungskette, dank des integrierten Datenmanagers auch für sehr große Datenmengen, ab. Für die direkte Anbindung an GIS-Software ist OPALS über Python nutzbar und für QGIS existiert die spezielle Erweiterung *QPALS* (OPALS 2017b). Schließlich unterstützt auch die kommerzielle Geodatenbank *Oracle Spatial* mit dem Packet *SDO_PC_PKG* ebenfalls die Speicherung und Verarbeitung von ndimensionalen Punktwolken. Wie bei Pointcloud für PostGIS werden Punktwolken in Blöcke benachbarter Punkte mit einer vorgegebenen Punkt-Kapazität gekachelt. Im Gegensatz zur PostGIS-Datenbank sind mit der Funktion *PC_Clip()* Selektionen von Teilmengen aus einer Punktwolke sowohl mit 2D- (Polygonen) als auch mit 3D-Geometrien (Quadern) möglich (ORACLE 2017).

5 Fazit & Ausblick

Punktwolken können als Objekte in Geodatenbanken integriert werden und es gibt viele Beispiele für den Nutzen von originalen 3D-Punktwolken in GIS. Auf Basis der räumlichen Nähe von Punktwolken-Punkten zu anderen Geoobjekten können durch räumliche Verbindung Informationen aus einer 3D-Punktwolke auf andere Geoobjekte übertragen werden und umgekehrt. Auch alltägliche Ingenieurprojekte, die nicht zwangsläufig klassischen GIS-Anwendungen zuzuordnen sind können durch das Management von Punktwolken in räumlichen Datenbanken optimiert werden. Projektabläufe profitieren von der kombinierten Speicherung und Analyse von großen 3D-Punktwolken mit anderen Projektdaten (ob administrativ oder fachlich mit und ohne Raumbezug) sowie den Vorzügen von zentraler Datenhaltung und Mehrbenutzerbetrieb.

Die PostGIS-Datenbank ermöglicht mit der Pointcloud–Erweiterung das Anlegen von Punktwolkentabellen. Bei der praktischen Arbeit im Hinblick auf 3D-Anwendungen hat PostGIS/Pointcloud jedoch die Einschränkung, dass Operationen mit Pointcloud-Objekten nur mit klassisch zweidimensionalen PostGIS-Features möglich sind. Außerdem ist für den Anwender die Frage nach einer geeigneten Punktwolken-Kachelung ein noch wenig beleuchtetes Thema. Es gibt gewisse Standardwerte die aber nicht zwangsläufig die beste Wahl für jede Art von Punktwolke sind. Die Punktwolken-Erweiterung Pointcloud wurde vor dem Hintergrund von LiDAR-Punktwolken aus Airborne Laserscanning für große Gebiete im kleinen Maßstab entwickelt und für solche Anwendungen sind grobe Kachelungen und Operationen in 2D ausreichend.

In diesem Beitrag wurde gezeigt, wie eine erweiterte Punktwolkenkachelung unter Berücksichtigung der Punktwolkenausdehnung entlang der Z-Achse realisiert werden kann, indem Funktionen der PDAL-Bibliothek kombiniert zum Einsatz kommen. Es lässt sich feststellen, dass eine solche 3D-Kachelung tendenziell positive Auswirkungen auf den Speicherplatzbedarf von Punktwolkentabellen in PostGIS hat und dass diese Art der Punktwolkenorganisation wesentlich präzisere Punktmengen in frühen Filterschritten bei mehrstufigen Abfrageprozeduren für 3D-Anwendungen liefert.

Im Kontext von Themen wie "Indoor-GIS" oder die Integration von *Building Information Modeling* (BIM) und GIS, kommt dem Thema "3D" in GIS eine ganz neue Bedeutung und auch Relevanz zu. Punktwolken haben individuelle Eigenschaften und werden für unterschiedliche Zwecke erfasst. Diese Umstände müssen bei der Wahl einer geeigneten Punktwolken-Kachelung und den in einem System zur Verfügung stehenden Operationen berücksichtigt werden.

Es gibt bereits umfangreiche Vergleiche von Datenbank-Lösungen, bei denen die Verwaltung von unterschiedlich großen LiDAR-Punktwolken in verschiedenen Systemen (PostGIS, Oracle Spatial u. MonetDB) getestet und einander gegenübergestellt wurden (VAN OOSTEROM et al. 2015). Noch eingehendere Untersuchungen der Auswirkungen von verschiedenen Punktwolken-Arten in verschiedenen Größen mit unterschiedlichen Kachelungen vor allem für "3D-relevante" Projektszenen im großen Maßstab (z. B. 3D-Punktwolken in einem "Campus-GIS") wären aufschlussreich.

6 Literaturverzeichnis

- ANDRAE, C., 2009: OpenGIS essentials: Die Geo-Standards von OGC und ISO im Überblick: Spatial Schema. Wichmann.
- ASPRS, 2013: LAS Specification Version 1.4 R13: The American Society for Photogrammetry & Remote Sensing.
- BRINKMANN, T., 2013: Geodatenbanksysteme in Theorie und Praxis. Wichmann.
- BURGER, S., ELFLEIN, A. & VÖLTER, U., 2015: Photo-based-scanning als Erweiterung des ingenieurgeodätischen Leistungsspektrums. ZfV Zeitschrift für Geodäsie, Geoinformation und Landmanagement, 140(4), 197-205.
- BOUNDLESS, 2015: Homepage Open Geo Suite 4.6: https://connect.boundlessgeo.com/docs/suite/4.6/index.html
- CLOUDCOMPARE, 2018: Projekthomepage: https://www.danielgm.net/cc/
- ESRI, 2016: Dokumentation ArcGIS Desktop: <u>http://desktop.arcgis.com/de/arcmap/10.3/manage-data/las-dataset/what-is-a-las-dataset-.htm</u>
- HAALA, N. & BRENNER, C., 1998: Interpretation of Urban Surface Models Using 2D Building Information. Computer Vision and Image Understanding, **72**, 204-214.
- HÖFLE, B., RUTZINGER, M., GEIST, T. & STÖTTER, J., 2006: Using airborne laser scanning data in urban data management - set up of a flexible information system with open source components. Proceedings UDMS 2006: Urban Data Management Symposium. Aalborg, Denmark, 7.11-7.23 (on CD-ROM).
- HÖFLE, B. & HOLLAUS, M., 2010: Urban Vegetation Detection using high Density Full-Waveform Airborne LIDAR Data – Combination of Object-based Image and Point Cloud Analysis. ISPRS TC VII Symposium, 2010.
- HÖFLE, B. & JOCHEM, A., 2012: 3D-Laserscanning-Punktwolken und GIS Aktuelle Entwicklungen. gis.SCIENCE, 2, 91-100.
- HUBER, D., 2011: The ASTM E57 File Format for 3D Imaging Data Exchange: Proceedings of the SPIE Vol. 7864A, Electronics Imaging Science and Technology Conference (IS&T), 3D Imaging Metrology, Vol. 7864A.
- JOCHEM, A., HÖFLE, B. & RUTZINGER, M., 2011: Extraction of Vertical Walls from Mobile Laser Scanning Data for Solar Potential Assessment. Remote Sensing, **3**, 650-667.
- LIBLAS, 2018: Projekthomepage: <u>https://liblas.org/</u>
- MUKUPA, W., ROBERST, G. W., HANCOCK, C. M., AL-MANASIR, K., 2016: A Review of the Use of Terrestrial Laser Scanning Application for Change Detection and Deformation Monitoring of Structures. Survey Review.
- MERWADE, V., COOK, A. & COONROD, J., 2008: GIS techniques for creating river terrain models for hydrodynamic modeling and flood inundation mapping. Environmental Modelling & Software, **23**, 1300-1311.
- OGC, 2018: Homepage der OGC Arbeitsgruppe "Point Cloud": http://www.opengeospatial.org/projects/groups/pointclouddwg
- OPALS, 2017a: Homepage von OPALS: https://geo.tuwien.ac.at/opals/html/index.html
- OPALS, 2017b: Homepage von OPALS und QPALS: https://geo.tuwien.ac.at/opals/html/usr_qpals.html

ORACLE, 2017: Homepage von Oracle:

https://docs.oracle.com/database/121/SPATL/sdo pc pkg-clip pc.htm#SPATL1190

- OTEPKA, J., GHUFFAR, S., WALDHAUSER, C., HOCHREITER, R. & PFEIFER, N., 2013: Georeferenced Point Clouds: A Survey of Features and Point Cloud Management. ISPRS International Journal of Geo-Information, **2**, 1038-1065.
- PCL, 2018: Projekthomepage: http://pointclouds.org/
- PDAL, 2018: Projekthomepage: https://pdal.io/index.html
- RAPIDLASSO, 2018: Homepage von Rapidlasso GmbH: https://rapidlasso.com/lastools/
- RUSU, R. & COUSINS, S., 2011: 3D is here: Point Cloud Library (PCL). IEEE International Conference on Robotics and Automation, 1-4.
- TANG, P., HUBER, D., AKINCI, B., LIPMAN, R. & LYTLE, A., 2010: Automatic reconstruction of asbuilt building information models from laser-scanned point clouds: A review of related techniques: Automation in Construction, 19, 829-843.
- VAN OOSTEROM, P., MARTINEZ-RUBI, O., IVANOVA, M., HORHAMMER, M., GERINGER, D., RAVADA, S., TIJSSEN, T., KODDE, M., GONCALVES, R., 2015. Massive point cloud data management: Design, implementation and execution of a point cloud benchmark: Computers & Graphics, 49, 92-125.
- VOEGTLE, T., STEINLE, E. & TOVARI, D., 2005: Airborne Laserscanning Data for Determination of suitable areas for Photovoltaics. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **36**(3/W19), 215-220.
- ZHOU, Q., PARK, J. & KOLTUN, V., 2018: Open3D: A Modern Library for 3D Data Processing. arXiv: 1801.09847.

Extraction of IndoorGML Model from an Occupancy Grid Map Constructed Using 2D LiDAR

LEALEM S. TESSEMA^{1,2}, REINER JÄGER² & UWE STILLA¹

Abstract: Topological semantic indoor spatial data is deemed important for efficient navigation of mobile robots and humans alike. In an effort to standardize and facilitate interoperability of indoor spatial data, the Open Geospatial Consortium has adopted the IndoorGML spatial data model. There has been a research gap identified in the use of such high-level, semantically rich spatial data (e.g. IndoorGML) in a Simultaneous Localization and Mapping framework. This paper presents an entry-point study towards addressing this research gap by presenting a method based on mathematical morphology as a means of extracting topological and semantic information from occupancy grid maps. The extracted semantic & topological information is translated into an IndoorGML compliant semantic Node-Relation-Graph and validated against the OGC IndoorGML schema.

1 Introduction

Along with the advent of SLAM (Simultaneous Localization and Mapping) algorithms, autonomous service robots have been gaining popularity which is justified by the availability of numerous consumer assistive robots such as the "Roomba" autonomous vacuum cleaner and the Care-O-bot from Fraunhofer IPA. Such service robots can be used in a range of scenarios ranging from home and personal assistive systems to industrial applications. In order to accomplish their task, especially those exclusively operating in an indoor environment, require the use of some sort of spatial representation of the environment.

Indoor map representation and navigation problems have been long addressed by the robotics community in light of autonomous navigation of robots. Different map representation models have been proposed, such as: feature based maps, semantic maps and topological maps. It is often the case that algorithms involved in the generation of maps are geared towards producing point clouds as end products. A significant challenge with point clouds is that, they do not provide high level understanding of the environment. High level information in a form of semantic or topology information is deemed valuable especially when it comes to applications in Building Information Modeling (BIM) and Location Based Services (LBS).

High-level map representation in SLAM, especially in the case of service robotics facilitates semantic understanding and human-robot interaction. Integration with the Open Geospatial Consortium (OGC) indoor spatial data model standard – IndoorGML could be realized as an implementation of semantically and topologically rich map information in a SLAM framework. The integration of semantic and topologic information for example in the graph based SLAM frame-

¹ Technische Universität München, Photogrammetrie und Fernerkundung, Arcisstraße 21, D-80333 München, E-Mail: [lealem.tessema, stilla]@tum.de

² Hochschule Karlsruhe, Center of Applied Research (CAR), Moltkestr.30, D-76133 Karlsruhe, E-Mail: [lealem_seyoum.tessema, reiner.jaeger]@hs-karlsruhe.de
work introduces new challenges and opportunities in implementing optimization algorithms that make use of metric and non-metric information.

Indoor LBS require spatial models that support attribution of semantic properties, contain accessibility information, store spatial relationships and serve for the function of multi-modal navigation. The complex nature of indoor scenarios added to the diversity of users and diverse navigation modalities asks for a flexible and efficient spatial data model that is useful for agents navigating in an indoor environment.

2 Problem statement

The robotics community has been highly invested on the use of point clouds or Truncated Signed Distance Functions (TSDF) to model 3D geometry. The disadvantage in using such representations is that they have large memory footprints and they do not inherently provide high level understanding of the geometric representation (CADENA et al. 2016). The most important aspect of high-level map representations relevant to this study is the capability of such representations to facilitate interaction between robotic maps, Geographic Information Systems (GIS) and Building Information Modeling (BIM) standards. IndoorGML is a standard with a potential to address this issue of interoperability. CADENA et al. (2016), after providing an elaborate discussion on state of the art in high-level map representations, beyond point clouds, mesh models, surface models and TSDFs." Based on this premises this paper presents an entry point study, which later on expands to the use of IndoorGML models in a SLAM framework.

Occupancy grid maps are primarily intended to be used for robot navigation related tasks, whereas with the introduction of the IndoorGML standard, the usability of occupancy grid maps could be further extended as sources of up to date indoor spatial information that is usable to applications in indoor GIS and pedestrian navigation.

The IndoorGML standard highlights the importance of indoor sub-spacing. Sub-spacing is important to represent the geometric properties of an indoor space. An example for a need for sub-spacing is a situation where an indoor space has a large and complicated corridor structure where the representation of such space as a single unit might hide meaningful geometry which could be useful for navigation tasks.

3 Related work

3.1 IndoorGML

SRIVASTAVA et al. (2018) developed a methodology to convert CAD drawings of indoor building data into IndoorGML. They extended the IndoorGML core module to handle semantic information. Their approach relies on the availability of CAD DXF floor plan. Similarly, PANG et al. (2018) proposed a method to extract a building's indoor space information based on simple indoor space boundary calculation on an already existing BIM and GIS models. On the other hand, DÍAZ-VILARIÑO et al. (2017) investigated a method to extract IndoorGML model from point cloud data acquired from a SLAM based 3D mapping system (laser scanner); their approach

made use of the sensor trajectory computed from the mapping system along with region growing and adjacency analysis to extract semantic information and reconstruct the scene.

DIAKITÉ et al. (2017) discuss a set of recommendations for the OGC IndoorGML standard with the intent of improving the standard in future versions. They primarily investigated problems related to the space subdivision concept in IndoorGML. It is pointed out that the current version (1.3) of the standard fails to address the context of furnished 3D indoor environments. The authors propose a multi-criteria approach (Geometry-driven, Topology-semantic-driven and navigation driven criteria) towards automatic subdivision of space cells.

3.2 Extraction of Topology from Occupancy Grid Maps

Joo et al. (2010) propose a method for automatic generation of topological maps from occupancy grid maps using virtual door detection. Their method is implemented in two phases where virtual doors are detected from the occupancy grid using adaptive curvature estimation of corner features in the first phase and the topological structure extracted from the first phase is optimized by the use of a genetic algorithm in the second phase. Despite the homogenous nature of the environment where their experiments were carried out, the authors claim that their approach has a high degree of accuracy.

Image processing techniques could be used to extract high-level information from occupancy grid maps (FERNÁNDEZ-MADRIGAL & BLANCO CLARACO 2013; ELFES 1989). Alternatively, a learning based approach using artificial neural networks and Bayesian integration has been successfully implemented by THRUN & BÜCKEN (1996) for the same purpose. They used an artificial neural network to interpret sonar sensor reading into an occupancy grid map whereas the topological map is generated by splitting the metric map into sub-regions by identifying critical points on a Voronoi diagram that are closest to an occupied grid cell within a given neighborhood and connecting these points to the corresponding occupied grid cell by critical lines which represent doorways. Even though this approach perfectly fulfills its purpose when it comes to navigation, it lacks the semantic labeling aspect where indoor spaces are labeled as transition spaces (corridors and doors) and rooms.

POTUGA & ROCHA (2012) implemented an image processing based approach similar to the methodology adopted in our study. Their work in general deals with obtaining a global topological abstraction from a preexisting occupancy grid map. The topological structure is basically extracted from the skeletonization of free space which results in a Voronoi diagram. Corner points on the Voronoi diagram are considered as nodes and the lines between such nodes constituted the edges of the graph. One major drawback of this approach when it comes to the Node Relation Graph concept of IndoorGML is that the graph constructed this way does not portray the actual semantic & topological relationship of the primal space. For instance, there is no means of telling which nodes are rooms, corridors or doors.

To summarize, research in the extraction of topological information from occupancy grid maps could be categorized in to two as machine learning based and image processing based approaches. Our approach belongs to the later and incorporates both topological and semantic indoor information as per the IndoorGML standard.

4 Method

4.1 Simulation & Data Acquisition

Our study presents a set of methods to extract 2D IndoorGML model from an occupancy grid map generated by a simulated indoor robot fitted with a 2D LiDAR sensor. In order to generate the occupancy map, a particle filter based open source SLAM algorithm known as "gmapping" (GRISETTI et al. 2007) is used as a black box. The simulation is carried out in the ROS-GAZEBO robot simulation environment where two scenarios, a simple hypothetical 3D floor plan and the popular "Willow garage" floor plan were used. The data was acquired by driving a robot fitted with a 2D LiDAR in the simulation environment.



Fig. 1: A simple hypothetical 3D floor plan (left) and the Willow garage dataset (right)

4.2 IndoorGML

The OGC IndoorGML standard is developed as an application schema of the Geography Markup Language (GML) with an intension to facilitate the representation and interoperability of indoor spatial data for the purpose of indoor navigation. One of the intended application area for IndoorGML is in the localization of features in indoor space. It is meant to address requirements for representing spatial components and constraints defined by architectural components such as rooms, corridors and doors and the respective relationships among these components. The standard defines indoor constraints based on the notions of cellular space, semantic representation, geometric representation, topological representation and multi-layered representation (LEE & LI 2012).

Since the main purpose of the standard is to provide a framework for indoor spatial data oriented towards navigation, the semantics aspect of indoor space is also dealt from the point of view of indoor navigation where indoor cells are classified into navigable and non-navigable cells (LEE & LI 2012). Although IndoorGML supports both geometric and topological information, we focus on the network representation of cellular space, which in our case is to be extracted from occupancy grid maps. By network representation, it implies topological relationship which also might contain some level of semantic information. IndoorGML proposes the use of a Node-Relation Graph (NRG) to represent topological relationships such as adjacency and connectivity.

In order to represent a cellular (geometric) indoor space in to a graph structure, the IndoorGML standard makes use of the theory of Poincaré duality. In Poincaré duality applied to indoor space, a room is to be represented as a point and the adjacency between rooms (shared 3D wall or 2D line) is to be represented by an arch connecting the two spaces. The standard identifies two sorts of adjacency properties, connectivity and accessibility.

The notion of Node-Relation Graph lends itself to the idea of graph data structure commonly known in mathematics and computer science. A graph generally represents relationships between pairs of objects commonly called nodes or vertices. This relationship is represented by a set of pairwise connections between nodes referred to as edges. A formal definition of a graph is given by MARCHAND-MAILLET & SHARAIHA (2000) as: A graph G = (V, E) is a set of vertices V with their inter-relationships given by the set of arcs E. If an orientation is associated with any arc, the graph is said to be directed otherwise G is an undirected graph. In this paper, when we only assume undirected graphs where the set of all possible relations between nodes is given by:

$$\boldsymbol{H} = \bigcup_{i}^{N} C(\boldsymbol{v}_{i}, 2) \; ; \; \boldsymbol{G} \subset \boldsymbol{H}$$
(1)

Where V represents the set of all vertices V_i and all possible edges are represented as pairwise combination of all vertices (nodes). An instance of **H** which represents a particular configuration space of an indoor environment - G is a set of nodes and edges in dual space.

4.3 Image Operations on Occupancy Grid Maps

The concept of occupancy grid maps as a probabilistic tessellated space representation of spatial information was first introduced by ELFES (1989). A formal definition of occupancy grid map is given as: "An occupancy field O(x) is a discrete-state stochastic process defined over a set of continuous spatial coordinates $x = (x_1, x_2 \dots x_n)$ while the occupancy grid is a lattice process, defined over a discrete spatial lattice". Each occupancy grid cell "c" is associated with a binary random variable s(c) with a Bernoulli distribution (FERNÁNDEZ-MADRIGAL & BLANCO CLARA-CO 2013; ELFES 1989). One advantage of occupancy grid maps is that they seamlessly fit into Bayesian particle filter based recursive estimation algorithms; on the other hand their huge storage requirement makes them infeasible options for mapping large scale environments (FERNÁNDEZ-MADRIGAL and BLANCO CLARACO 2013).

For the sake of simplicity, we describe the occupancy mapping approach in the case of mapping with a known pose (THRUN et al. 2005; FERNÁNDEZ-MADRIGAL & BLANCO CLARACO 2013). The posterior to be estimated under this assumption is the map given by the conditional probability $p(m|z_{1:t}, x_{1:t})$ for each pixel grid cell m_i and all sets of measurements and poses up to time t. The posterior becomes tractable if the individual distributions on m_i are estimated rather than on the whole joint probability. Assuming that the individual grid cells are independent from each other, the posterior could be simplified as a product of its marginal given by:

$$p(m|z_{1:t}, x_{1:t}) = p(\{m_i\}_i^N | z_{1:t} x_{1:t}) \approx \prod_i^N p(m_i | z_{1:t}, x_{1:t})$$
(2)

After successive application of Bayes rule, the conditional independence assumptions given by $z_t \perp z_{1:t-1} \mid x_t, m_i$ and $m_i \perp x_{1:t}$ enable the formulation of the log odds of the posterior $p(m_i \mid z_{1:t}, x_{1:t})$ as:

$$l_t(m_i) = \tau_t(m_i) - l_o(m_i) + l_{t-1}(m_i)$$
(3)

Where:

$$\begin{split} l_t(m_i) &= \ln \frac{p(m_i | Z_{1:t}, x_{1:t})}{p(\neg m_i | Z_{1:t}, x_{1:t})} \text{, the log odd of the posterior to be estimated;} \\ \tau_t(m_i) &= \ln \frac{p(m_i | Z_t, x_t)}{p(\neg m_i | Z_t, x_t)} \text{, the inverse sensor model of a 2D LiDAR for a given grid cell;} \\ l_0(m_i) &= \ln \frac{p(m_i)}{p(\neg m_i)} \text{, represents a-priori information about the map occupancy;} \\ l_{t-1}(m_i) &= \ln \frac{p(m_i | Z_{1:t-1}, x_{1:t-1})}{p(\neg m_i | Z_{1:t-1}, x_{1:t-1})} \text{, the previous occupancy state of the grid cell;} \end{split}$$

4.3.1 Mathematical Morphology

Occupancy grid maps can be transformed into binary images that represent occupied space as black pixels and free space as white pixels by thresholding. This enables the use of morphological operations to manipulate the geometric content based on the contents of neighboring pixels defined by a structuring element. In this sub-section we understand occupancy grid maps as simple binary images with values 1 as foreground pixels and 0 as background pixels. Onwards, we make use of formal definitions of morphological filters and operations given by BURGER & BURGE (2009).

Dilation: $I \oplus H \equiv \{(p+q) | \forall p \in I, q \in H\}$ **Erosion:** $I \ominus H \equiv \{p \in Z^2 | (p+q) \in I, \forall q \in H\}$ **Opening:** $I \circ H = (I \ominus H) \oplus H$ **Closing:** $I \bullet H = (I \oplus H) \ominus H$

Skeletonization is the process of converting foreground pixels in to strings of single pixels which capture the geometric essence of the foreground pixel components in the image. Skeletonizing an occupancy grid map leads to the representation of free space by a string of pixels which capture the geometric nature of the free space. In this regard, we have made use of Blum's Medial Axis Transform (MAT) (BLUM 1967) to construct skeletons of from occupancy grid maps. Such representation of free space is often referred to as generalized Voronoi graph (diagram).

One drawback of using morphological filters is the specification of the structure and dimension of the structuring element which needs to change as per the texture of the image to be used. The use of a suitable parameter depends on a prior knowledge of the environment such as width of doorways and corridors. In addition, certain morphological operators with a square structuring element work best on Manhattan like environments and perform weakly on other environments.

4.3.2 Morphological Segmentation (Watershed Transform)

The watershed transform borrows the notion of a watershed and catchment basin from physical geography. In geography, a watershed is an area of land that marks the boundary of a catchment basin. A catchment basin on the other hand represents the area of land where water drains off into a common pour point. Generation of a watershed or catchment is commonly performed from a digital elevation model in geography.

In image processing, the watershed transform is commonly used along with the distance transform. The result of a watershed transform in MATLAB is a label matrix which represents individual catchment basins where the watershed ridge pixels have a value of zero (GONZALEZ et al. 2009). One major drawback of watershed based segmentation is that without the use of interactively provided markers, the result could be over-segmented due to noise and other local irregularities.



Fig. 2: Watershed transform (c) on an occupancy grid map (a) and the corresponding distance transform (b).

4.3.3 Region Adjacency Graph (RAG)

Once the occupancy grid map is converted in to a symbolic image (segmented & labeled), the next step would be the extraction of the topological relationships among the labeled image components. The result of such operation is a region adjacency graph that stores spatial adjacency information. This final graph structure corresponds to IndoorGML's NRG and is later translated in to an IndoorGML file and was validated against the OGC IndoorGML schema online.

In order to extract a RAG from a symbolic image, horizontal and vertical adjacencies (4adjacency) between pixels with different labels are detected and these are added as new edges to the adjacency graph being constructed. In this study, we have made use of the algorithm proposed by SHAPIRO (1996) for the extraction of region adjacency from a labeled occupancy grid map.

5 Results

5.1 Experiment I

In the first experiment, we investigated the results of binary connected component labeling after elementary morphological operations were performed on an occupancy grid map. The method implemented for this experiment is presented in Tab. 1. In steps 1 and 2 the input occupancy map is binarized and the walls are further articulated by a dilation operation. Steps 3 to 6 deal with the extraction of doors. Finally, the region adjacency is computed by the method mentioned in section 4.3.3. As shown in Fig. , the algorithm performs well in the hypothetical (simple) dataset. Whereas in Fig. (b), it is shown that the applied method fails to cope with the scale and complexity of the Willow garage dataset. From the results, one could understand that the fundamental factor that contributes to the success or failure in this approach is the homogeneity and complexity of the environment.



Fig. 3: Occupancy grid maps and their corresponding NRG extracted using connected component labeling. (a) On a hypothetical environment and (b) On the Willow garage dataset.

Tab. 1: Method for the extraction of NRG based on connected component labeling.

Method: NRG from Connected Component Labeling

 Binarize occupancy grid map *I*_{bin} = binarize(*I*_{occ}, 0.85)

 Morphological erosion

 $I_{erd} = I_{bin} \ominus SE$

- 3. Morphological Opening $I_{opn} = I_{erd} \circ SE$
- 4. Connected component labeling
- $I_{label} = label(I_{opn})$
- Binary XOR to detect doors & label doors
 I_{door} = label(I_{opn} XOR I_{erd})
- 6. Reintroduce doors in the symbolic image (join the two labels)
 *I*_{label} = *I*_{door} ∪ *I*_{label}
- 7. Generate region adjacency graph $G_{RAG} = regionAdjacency(I_{label})$
- 8. Write G_{RAG} to IndoorGML as NRG

Fig. 6(b) (bottom) shows the region adjacency graph generated from the labeled symbolic image of the willow garage dataset. In such a graph the nodes actually represent the centroid of the corresponding labeled component. It is for this reason that all the edges appear to converge to the center of the largest component, which is the corridor in this case. Due to this phenomenon it is not possible to represent the actual topological & geometric structure of the environment using the method described in Tab. 1. Furthermore the door detection (step 5) which uses XOR operator on I_{opn} and I_{erd} results in incorrect regions due to the effect of the morphological thickening that also affects other regions of the image.

5.2 Experiment II

In the second experiment, we addressed the limitations discussed in experiment I. Furthermore, we made use of the watershed segmentation method to further subdivide the corridor so that the actual geometric structure of the environment could be preserved. For the detection of doors, we adopted a new method that makes use of the medial axis transform. The skeletonization of pixels representing empty space (white) results in what is known as the generalized Voronoi diagram. A generalized Voronoi diagram is the generalization of the ordinary Voronoi diagram of points. In this particular case the generalization would be in the use of the medial axes as seeds (instead of points). On the other hand, the complement of I_{opn} was also skeletonized and the intersection (step 3 on Tab. 2) of these and the Voronoi pixels were considered to be transition spaces (doorways).

Method: Door Detection using Medial Axis Transform (MAT)							

Tab. 2: Method for door detection using skeletonization

The detection of doors plays a major role in the semantic labeling of the whole grid map. After all the candidate doors were detected as per the method shown in Table 2, the NRG was initialized as a node only graph with the centroid of the door pixels. The connectivity information was added incrementally by introducing the labeled image from the watershed segmentation. In order to identify rooms and corridors, a rule was formulated where non-door nodes with a degree of two or more are labeled as a transition (corridor) spaces as shown in Table 3, step 6.



Fig. 4: Detection of door pixels using intersection of complementary skeleton pixels I_{skel} (yellow) and I_{vor} (gray)

Once the RAG was generated from the union of the labeled image which is a result of the watershed segmentation and the labeled door image, further refinement of the graph was carried out by introducing topological constraints based on a-prior knowledge. Cyclic sub-graphs which are caused by the over-segmentation during watershed transformation were removed by applying these constraints in a post processing stage. Cyclic corridor-corridor-corridor sub-graphs were modified in a similar fashion as described inTab. 4. Reflexive node relationships were also used as a constraint to avoid door-door and room-room adjacencies. Edges representing such connections were simply deleted from the NRG. We have also imposed a constraint to remove "dangling" (degree = 1) corridor nodes as this conflicts with our definition of a corridor, which is a node in the NRG with a degree of at least 2.

Tab. 3: Procedure for generating IndoorGML NRG from RAG obtained from watershed segmentation

Metho	d: IndoorGML NRG Generation from RAG
1.	Initialize G with door nodes
	$G_{NRG}(nodes) = unique (I_{door})$
	$G_{NRG}(edges) = \emptyset$
2.	Append space nodes from watershed label to G
	$G_{NRG} = addNode(unique(I_{label}))$
3.	Reintroduce doors in the symbolic image (join the two symbolic images)
	$I_{label} = I_{door} \cup I_{label}$
4.	Generate region adjacency graph
	$G_{RAG} = regionAdjacency(I_{label})$
5.	Add edges to graph
	$G_{NRG}(edges) = G_{RAG}(edges)$
6.	Find corridor nodes using degree of node
	$trans = degree(G_{NRG}(nodes)) \ge 2$
	$\neg doors = G_{NRG}. type \neq 'door'$
	$corridors = trans \cap \neg doors$
7.	Assign edge weights based on connectivity
	$edge(n_{door}, n_{door})$. weight = -1
	$edge(n_{room}, n_{room})$. weight = -1
	$edge(n_{corridor}, n_{corridor})$. $weight = 0$

A visual comparison of the NRG with and without the refinement is presented in Fig. 8. The refinement of the RAG extracted from watershed segmentation by modifying cyclic sub-graphs as per the method described inTab. 4 resulted in a simplification of the graph structure. Fig. 8(b) shows the final and simplified semantic-topological map of the Willow garage dataset where the ID of each graph node corresponds to the label of the respective region in the binary occupancy grid map.

Tab. 4: Procedure for refining the NRG by modifying cyclic sub-graphs

Method: Modify cyclic corridor-door-corridor sub-graphs	
for all $(n_i. type ==' door')$ in G_{NRG} do	
$n_{nh} = neighbors (G_{NRG}, n_i)$	
$e_{nh} = {}_{n_{nh}}C_2$	
for all e ⁱ nh G _{NRG} do	D
$if(e_{nh}^{i}.weight == 0)$	\sim
delete e_{nb}^i	(C)(C)
endif	\bigcirc \bigcirc
endfor	
endfor	



Fig. 5: Semantic topological map before (a) and after NRG refinement (b). In (b) spurious triangular cyclic sub-graphs are removed.

6 Discussion

In order to acquire an overall accuracy assessment, the final result on Fig. 8(b) was visually compared against a manually labeled image, which served as a ground truth. On the ground truth, there were a total of 84 doors and 65 rooms identified. A label was considered as accurate or not based on a prior definition of doors and rooms. For example, a structuring element of 15 by 15 was used in the process of door detection, which implies (with a pixel size of 0.1) doorways only under 1.5 meters were considered as true doors. And rooms are considered to be nodes with a degree of 1. Based on these assumptions, it was possible to achieve an accuracy of 90.5% for door detection. 62.5% of the errors (missed & wrong labels) are within the toilet area (highlighted in black rectangle on Fig. 8(b)), where the rooms are significantly small that the applied morphological operations resulted in loss of details. Finally, the room labeling accuracy was at 90.76% with 66.67% of the errors still lying within the toilet area. Outside the toilet area, there were only 2 wrong labels of rooms. Two room nodes that are accessible through more than one doorway (node degree > 1) were labeled as corridors and these are not considered as failure cases as they comply with the prior definitions of a rooms and corridors.

7 Summary

In this paper, we presented a semi-automatic method for the extraction of IndoorGML model from occupancy grid maps. Our approach laid out the foundations towards an automatic extraction of graph based semantic indoor data without the need for a learning based approach. The next step in this regard would be manipulation of the mentioned core image processing algorithms such as watershed transform and connected component labeling, in a way that is possible to handle arbitrarily scaled environments without manual adjustment of parameters, particularly of the structuring element in morphological operations.

The core of our approach lies on the detection of doors based on the results of medial axis transformation and refinement of the graph structure computed from region adjacency. The results presented could be considered as initial attempts to extend the functionalities and usability of indoor robot maps for the purpose of extracting meaningful information relevant to BIM and GIS. Next, our investigation continues to investigate the same problem in realistic occupancy grid maps constructed from real indoor environments. We assume working on a real dataset poses additional challenges especially when it comes to the detection of doors.

Further investigation in this line of research would focus on the use of hierarchical graph structures to elegantly handle the subdivision of complex indoor spaces, introduction of multi-layered grid maps with additional information layers (such as curvature from point clouds) and introduction of an IndoorGML model updating procedure based on change detection framework.

References

BLUM, H., 1967: A Transformation for Extracting New Descriptors of Shape. In: Wathen-Dunn, W. (ed.), Models for the Perception of Speech and Visual Form. MIT Press, Cambridge, 362-380.

- BURGER, W. & BURGE, M., 2009: Principles of Digital Image Processing: Fundamental Techniques. Springer, London.
- CADENA, C., CARLONE, L., CARRILLO, H., LATIF, Y., SCARAMUZZA, D., NEIRA, J., REID, I. & LEONARD, J.J., 2016: Past, Present, and Future of Simultaneous Localization and Mapping: Towards the Robust-Perception Age. IEEE Transactions on Robotics, **32**(6), 1309-1332.
- DIAKITÉ, A.A., ZLATANOVA, S. & LI, K.-J., 2017: About the Subdivision of Indoor Spaces in IndoorGML. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-4/W5, 41-48.
- DÍAZ-VILARIÑO, L., VERBREE, E., ZLATANOVA, S. & DIAKITÉ, A., 2017: Indoor Modeling from SLAM-Based Laser Scanner: Door Detection to Envelope Reconstruction. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(2/W7), 345-352.
- ELFES, A., 1989: Using Occupancy Grids for Mobile Robot Perception and Navigation. Computer, **22**(6), 46-57.
- FERNÁNDEZ-MADRIGAL, J.-A. & BLANCO CLARACO, J.L., 2013: Simultaneous Localization and Mapping for Mobile Robots: Introduction and Methods. Information Science Reference, Hershey, Pa.
- GONZALEZ, R.C., WOODS, R.E. & EDDINS, S.L., 2009: Digital Image Processing using MATLAB. Gatesmark Pub, S.I.
- GRISETTI, G., STACHNISS, C., AND BURGARD, W., 2007: Improved Techniques for Grid Mapping With Rao-Blackwellized Particle Filters. IEEE Transactions on Robotics, **23**(1), 34-46.
- JOO, K., LEE, T.-K., BAEK, S. & OH, S.-Y., 2010: Generating Topological Map from Occupancy Grid-Map using Virtual Door Detection. IEEE Congress on Evolutionary Computation, 1-6.
- LEE, J. & LI, K.-J., 2012: IndoorGML OGC Candidate Standard for Indoor Navigation. 58.
- MARCHAND-MAILLET, S. & SHARAIHA, Y.M., 2000: Binary Digital Image Processing: a Discrete Approach. Academic Press, San Diego.
- PANG, Y., ZHANG, C., ZHOU, L., LIN, B. & LV, G., 2018: Extracting Indoor Space Information in Complex Building Environments. ISPRS International Journal of Geo-Information, 7(8), 321.
- POTUGA, D. & ROCHA, R.P., 2012: Extracting Topological Information from Grid Maps for Robot Navigation: Proceedings of the 4th International Conference on Agents and Artificial Intelligence, SciTePress - Science and Technology Publications, 137-143.
- SHAPIRO, L.G., 1996: Connected Component Labeling and Adjacency Graph Construction. In: Topological Algorithms for Digital Image Processing. Elsevier, Amsterdam, 1-292.
- SRIVASTAVA, S., MAHESHWARI, N. & RAJAN, K.S., 2018: Towards Generating Semantically-Rich IndoorGML Data from Architectural Plans. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(4), 591-595.
- THRUN, S. & BÜCKEN, A., 1996: Integrating Grid-Based and Topological Maps for Mobile Robot Navigation. AAAI/IAAI, 2.
- THRUN, S., BURGARD, W. & FOX, D., 2005: Probabilistic Robotics. MIT press.

ALS Klassifizierung mit Submanifold Sparse Convolutional Networks

STEFAN SCHMOHL¹ & UWE SÖRGEL¹

Zusammenfassung: Die Klassifizierung von Punktwolken aus Airborne Laser Scanning (ALS) ist eine der Hauptkomponenten in deren Verarbeitung. Etablierte Methoden beruhen auf der aufwendigen Berechnung von punktweisen Merkmalen. Convolutional Neural Networks (CNNs) haben sich als mächtige Klassifikatoren etabliert, die zugleich auch die optimalen Merkmale selbstständig erlernen. Ihre Anwendung auf ALS-Daten ist jedoch nicht trivial und basiert bisher meist auf der Projektion der Punktwolke in zweidimensionale Abbildungen. Reine 3D-CNNs benötigen sehr viel Speicher und Rechenzeit. Wir verwenden daher Sparse Submanifold Convolutional Networks (SSCNs) für eine effiziente semantische Segmentierung von Voxelwolken in einer Ende-zu-Ende Encoder-Decoder-Architektur. Wir demonstrieren das Verfahren auf dem ISPRS Vaihingen 3D Semantic Labeling Benchmark und erreichen bis zu 85,0% Gesamtgenauigkeit bei zugleich geringen Inferenzzeiten von nur wenigen Sekunden.

1 Einleitung

Flugzeuglaserscanning (ALS) liefert Massendaten in Form von 3D-Punktwolken. Um hieraus semantische Information über Objekte zu gewinnen, wird als Zwischenschritt oftmals jedem 3D-Punkt eine Klasse aus einem gegebenen Katalog von Objektkategorien zugewiesen. Eine solche Klassifikation kann jedoch nicht isoliert für Einzelpunkte erfolgen. Vielmehr ist die Einbeziehung des räumlichen Kontextes erforderlich, der sich aus der Verteilung der Punkte in einer lokalen Nachbarschaft ergibt. Üblicherweise werden dazu aus der Umgebung jedes Punktes geometrische Merkmale abgeleitet. Im klassischen Ansatz erfolgt die Definition der betrachteten Nachbarschaften a priori durch den Menschen. Die Klassifikation der Punkte im Merkmalsraum erfolgt anschließend mit Standardverfahren wie Random Forests.

Convolutional Neural Networks (CNNs) haben sich in den letzten Jahren als State of the Art in der Bildanalyse etabliert. Um auch 3D-Daten mit diesen in erster Linie für 2D-Bilder entwickelten Methoden zu verarbeiten, erfolgt häufig eine Abbildung der 3D-Daten in eine Menge von 2D-Projektionen. Dies ist jedoch mit einem Verlust an Informationen verbunden und lässt sich nicht auf Daten anwenden, deren Dreidimensionalität erhalten bleiben soll.

Da Faltungsoperationen auf Rasterdaten mathematisch keiner Beschränkung bezüglich der Dimension des Raumes unterliegen, können CNNs theoretisch Rasterdaten mit beliebig vielen Dimensionen und beliebiger Größe verarbeiten. In der Praxis begrenzt jedoch der hohe Speicher- und Rechenbedarf die Datenmenge und damit die Auflösung von 3D-Inputs.

3D-Daten sind oftmals durch eine stark inhomogene räumliche Verteilungsdichte gekennzeichnet. Große Teile des (Voxel-) Raumes sind in der Regel nicht belegt. Um dies auszunutzen, adaptieren

¹ Universität Stuttgart, Institut für Photogrammetrie, Geschwister-Scholl-Str. 24D,

D-70174 Stuttgart, E-Mail: [stefan.schmohl, uwe.soergel]@ifp.uni-stuttgart.de

wir in dieser Arbeit Submanifold Sparse Convolutional Networks (SSCN) zur semantischen Segmentierung von ALS-Punktwolken.



Abb. 1: Schema der Prozesskette. Ein Punktwolken-Sample wird einem Voxelfilter unterzogen und anschließend durch ein SSCN in der Form eines U-Nets semantisch segmentiert. Anschließend werden die Voxellabel zurück auf die Punktwolke übertragen. Angedeutet ist die räumliche Auflösung des Samples innerhalb des Netzes: umso tiefer die Stufe, desto geringer die Auflösung.

2 Forschungsstand

2.1 Überwachte Klassifizierung von ALS Daten

Das übliche Vorgehen bei der semantischen Segmentierung einer Punktwolke, oft auch Punktwolkenklassifizierung genannt, besteht aus einem zweistufigen Verfahren. Als Erstes werden auf Basis von Expertenwissen entworfene, punktweise Merkmale berechnet. Neben echobasierten Merkmalen und normalisierten Höhen können eine Reihe von nachbarschafts-bezogenen Merkmalen verwendet werden, beispielsweise berechnet aus den Eigenwerten des Strukturtensors.

Im zweiten Schritt folgt die Klassifizierung der Punkte anhand dieser Merkmale. Typische Klassifikatoren sind Support Vector Machines (SVM) oder Random Forests (RF), z.B. in CHEHATA et al. (2009), BLOMLEY & WEINMANN (2017) und HACKEL et al. (2016). Solche Verfahren behandeln jeden Punkt einzeln, ohne Berücksichtigung semantischer Interaktionen zwischen den Klassen benachbarter Punkte, was zu feinkörnig verrauschten Resultaten führen kann. Zur Einbeziehung von räumlichem Kontext klassifizieren NIEMEYER et al. (2014; 2016) alle Punkte simultan in einem Conditional Random Field (CRF).

Die Berechnung der für diesen Ansatz notwendigen Merkmale beruht zu einem Großteil auf zeitaufwendigen Nachbarschaftsanfragen. Zudem muss für jede Anwendung manuell ein möglichst optimaler Satz dieser Merkmale gefunden werden.

2.2 Convolutional Neural Networks zur ALS Klassifizierung

Convolutional Neural Networks (CNNs) sind State of the Art in vielen Disziplinen wie der Computer Vision, insbesondere in der Bild-Klassifizierung. Sie lernen neben der Merkmalsklassifikation zusätzlich auch deren Extraktion in einer Ende-zu-Ende-Manier. Gewöhnliche CNNs benötigen rasterisierte, zweidimensionale Eingangsdaten. 3D Punktwolken sind jedoch unstrukturiert und weisen stark inhomogene Punktdichten auf. Die Anwendung auf ALS-Daten ist aus diesen Gründen nicht trivial. Die meisten vergleichbaren Arbeiten konzentrieren sich darauf, ALS-Punktwolken möglichst sinnvoll für die Verarbeitung mit CNNs in 2D bzw. 2,5D Raster umzuwandeln. So klassifizieren beispielsweise HU & YUAN (2016) ALS-Punkte, indem sie jeden Punkt durch eine vertikale Projektion ihrer Umgebung beschreiben. Jedes Pixel erhält dabei drei Werte: Z_{min} , Z_{mittel} und Z_{max} . Die vom neuronalen Netzwerk prädizierte Objektkategorie für ein solches Bild wird dann übertragen auf den ursprünglichen, in der Mitte liegenden 3D Punkt. Nachteilig bei diesem Ansatz sind die vielen redundanten Berechnungen, da für nahe beieinanderliegende Punkte mehrfach dieselben Merkmale berechnet und innerhalb des Netzes verarbeitet werden müssen. Zusätzlich ist das Ergebnis anfällig für verrauschte Ergebnisse, da die Punkte einzeln, ohne Berücksichtigung der semantischen Beziehungen benachbarter Punkte, klassifiziert werden.

Im Gegensatz dazu erlauben Encoder-Decoder-Architekturen die simultane Zuordnung aller Elemente (Pixel) der Eingabe (POLITZ & SESTER 2018; RIZALDY et al. 2018). Szenen können damit am Stück prozessiert werden. Das Problem des Informationsverlustes durch die Projektion in 2,5D bleibt jedoch bestehen, insbesondere bei Verdeckungen, Fassaden und Multi-Echo-Signalen. Zudem stellt die Berechnung der Abbildungen sowie die Rückprojektion einen zusätzlichen Aufwand dar.

Grundsätzlich können die Operationen eines CNN über beliebig viele Dimensionen definiert werden. Ebenso ist die Rasterisierung von Punktwolken auch im dreidimensional Raum möglich. Die daraus resultierenden Voxelgitter erfordern allerdings sehr viel Speicher und Rechenzeit bei der Verarbeitung in einem 3D-CNN. Dies ist besonders deshalb unverhältnismäßig aufwendig, da der Großteil des Raumes leere Voxel beinhaltet, also sehr *sparse* ist.

Um die geringe Dichte von 3D Daten zu nutzen, wurden verschiedene Ansätze entwickelt, um CNNs auch auf anderen 3D Datenstrukturen als Voxelgittern anzuwenden. GRAHAM et al. (2018) nutzen in ihren *Submanifold Sparse Convolutional Networks* (SSCNs) eine Implementierung der Faltungsschichten in Form von Matrix-Multiplikationen, um ausschließlich "belegte" Voxel zu betrachten. Dieses Verfahren erzielte die besten Resultate in der Segmentierung von Objektteilen (YI et al., 2017).

Die in der Computer Vision entwickelten 3D-CNNs wurden bisher meist nur für kleine, synthetische Datensätze oder räumlich stark begrenzte, terrestrische Aufnahmen und Innenraum-Scans verwendet. Die Anwendung auf großräumige, mittels ALS von realen Objekten aufgenommene, topographische Punktwolken ist nach unserem Wissen bisher nicht untersucht worden. In dieser Arbeit zeigen wir die Eignung von SSCNs für die Klassifizierung von ALS Punktwolken.

3 Methodik

3.1 Submanifold Sparse Convolutional Network

Wesentlicher und namensgebender Bestandteil von Convolutional Neural Networks sind die Faltungs-Schichten. In ihnen werden mehrere Faltungskerne mit gelernten Gewichten auf die Ergebnisse der jeweils vorherigen Schicht (*activation maps*) angewandt. Im 2D-Fall sind diese activation maps und Faltungskerne dreidimensional, wobei die dritte Dimension die Anzahl an Input-Kanälen bzw. an Filtern der vorherigen Schicht repräsentiert. Die Faltung hat die Form

$$Y_f^l = X^l * W_f^l \tag{1}$$

Dabei bezeichnet W_f^l einen 3D-Faltungskern f der aktuellen Schicht l und $X^l = h(Y^{l-1})$ das Ergebnis der vorherigen Schicht nach der Aktivierungsfunktion $h(\cdot)$.

Um diese Faltungen effizient auf GPUs zu berechnen, kann die Faltungsoperation als Matrixmultiplikation umgeschrieben werden (CHELLAPILLA et al. 2006; CHETLUR et al. 2014):

$$\boldsymbol{Y}^{l} = \boldsymbol{X}^{l} \cdot \boldsymbol{W}^{l} \tag{2}$$

Die Matrix $W^l \in \mathbb{R}^{k^2c \times |f|}$ enthält alle |f| Faltungskerne der Schicht, jeder mit der Größe $k \times k \times c$. Im Input $X^l \in \mathbb{R}^{|n| \times k^2c}$ bzw. Output $Y^l \in \mathbb{R}^{|n| \times |f|}$ steht |n| für die Anzahl an Positionen des Faltungskerns. Bei Bildern entspricht dies, ohne Stride und mit entsprechendem Padding, der Bild-Breite multipliziert mit der Bild-Höhe. Das Grundprinzip der Submanifold Sparse Convolution (SSC) ist es, nur diejenigen Zeilen n zu behalten, deren korrespondierende Stellen im ursprünglichen Input nicht leer sind. Dafür genügt es, die nicht leeren Stellen in Listenform zu speichern (*Voxelwolke*), ein voll aufgespanntes Voxelgitter wird nicht benötigt. Für weitere Details siehe GRAHAM (2015) und GRAHAM et al. (2018).

3.2 Netzwerk-Architektur

Wir adaptieren die U-Net Architektur (RONNEBERGER et al. 2015) aus GRAHAM et al. (2018) zur semantischen Segmentierung von voxelierten ALS Punktwolken (Abb. 1). Die Encoder-Decoder-Architektur erlaubt die Ende-zu-Ende Verarbeitung einer Voxelwolke. Pro Stufe des Encoders wird die Auflösung durch Unterabtastung halbiert. Das Netz setzt sich aus genug Schichten zusammen, um die räumliche Ausdehnung in der tiefsten Schicht auf ein Voxel zu reduzieren. Der Decoder ist symmetrisch zum Encoder und verwendet "Deconvolution"-Schichten, um die Auflösung stückweise wieder zu erhöhen. Diese Werte werden jeweils mit denen aus der korrespondieren Encoder-Stufe verbunden. Am Ende des Netzes besitzt jedes Voxel prädizierte Klassen-wahrscheinlichkeiten und wird der Klasse mit der höchsten Wahrscheinlichkeit zugewiesen. Die Label jedes Voxels werden schließlich, außerhalb des Netzes, auf die darin liegenden Punkte übertragen.

3.3 Zielfunktion

Ein Problem bei der semantischen Segmentierung mit CNNs sind Datensätze mit stark inhomogenen Klassenverteilungen. Neuronale Netze neigen generell dazu, beim Training häufig auftretende Klassen bei der Inferenz zu bevorzugen. Im Gegensatz zur üblichen Klassifizierung ist Unter- oder Übersampling hier aber nicht praktikabel, da die Klasseninstanzen nicht einzeln auftreten, sondern nur als Teil der Stichproben, z.B. als Pixel in einem Bild oder Voxel in einem (dünn besetzten) 3D-Gitter.

Alternativ zum Sampling kann auch die Zielfunktion angepasst werden. Wir verwenden in dieser Arbeit einen gewichteten, elementweisen Cross-Entropy-Loss, der mittels Stochastic Gradient Descent optimiert wird (LONG et al. 2015; RONNEBERGER et al. 2015; EIGEN & FERGUS 2015):

$$E = -\frac{1}{\sum_{n=1}^{N} \sum_{x \in \Omega^{(n)}} w(x)} \sum_{n=1}^{N} \sum_{x \in \Omega^{(n)}} \sum_{c=1}^{C} y_c(x) \log(\hat{y}_c(x)) w(x)$$
(3)

wobei *N* die Anzahl an Samples *n* im Mini-Batch bezeichnet, $\mathbf{x} \in \Omega^{(n)}$ die Voxelpositionen pro Sample, $\hat{y}_c(\mathbf{x})$ die prädizierte Wahrscheinlichkeit der Zugehörigkeit von \mathbf{x} zur Klasse *c*, y_c das gegebene, binär kodierte, Klassenlabel und C die Klassenanzahl. Es hat sich gezeigt, dass die Wurzel der inversen Klassenhäufigkeit hier bei diesem Datensatz ein guter Kompromiss zwischen Recall und Precision der seltenen Klassen ist. Wir verwenden daher

$$w(\mathbf{x}) = w(y(\mathbf{x})) = w_c = \frac{1}{\sqrt{f_c}}$$
(4)

mit f_c als relative Häufigkeit des wahren Labels y(x) bzw. der Klasse c im Trainingsdatensatz. Dieses *class balancing* führt dazu, dass seltene Klassen einen stärkeren Einfluss auf den Gradienten und damit den Lernfortschritt haben als häufigere Klassen. Bei feineren Voxelgrößen, die zu größerem Ungleichgewicht führen, sollte auf die Wurzel verzichtet werden.

4 Experimente

4.1 Datengrundlage

Wir demonstrieren die Eignung des Verfahrens auf dem Datensatz des ISPRS 3D Semantic Labeling Contest² (NIEMEYER et al. 2014). Er besteht aus zwei ALS Punktwolken, je eine für Training und Test, aufgenommen über Vaihingen an der Enz, Deutschland. Jedes Echo eines LiDAR-Sendepulses wurde darin als eigener Punkt aufgezeichnet mit den Merkmalen Intensität, Echonummer und -Anzahl. Zusätzlich liegt für jeden Punkt ein Label aus den Klassen *Powerline, Low vegetation, Impervious surfaces, Car, Fence/Hedge, Roof, Facade, Shrub* und *Tree* vor. Die nominelle Punktdichte pro Streifen beträgt 4 pts/m². Durch die 30 % Streifenüberlappung beträgt die effektive Punktdichte global etwa 8 pts/m². Zum Zeitpunkt dieser Arbeit ist der Contest bereits geschlossen, daher konnten wir unsere Ergebnisse nicht zur Evaluierung einsenden. Da nun aber auch die Label der Test-Punktwolke verfügbar sind, war trotzdem eine Auswertung zum Vergleich mit anderen Verfahren möglich.

Zusätzlich zu der Punktwolke steht ein True Orthophoto (TOP) desselben Gebiets aus dem entsprechenden 2D Contest zur Verfügung (CRAMER 2010). Dieses hat eine GSD von 9 cm und enthält die spektralen Kanäle Rot, Grün und nahes Infrarot (CIR). Bei einem Teil der Experimente wurde die Punktwolke zuvor anhand dieses TOP eingefärbt (Abb. 2).



(a) Trainings-Split



(c) Test-Split

Abb. 2: ISPRS Vaihingen 3D Semantic Labeling Datensatz. Angezeigt werden die anhand des TOP kolorierten Punktwolken.

² http://www2.isprs.org/commissions/comm3/wg4/3d-semantic-labeling.html

Zur laufenden Überwachung des Lernfortschritts separierten wir die Trainings-Punktwolke manuell in einen festen Trainings- bzw. Validierungs-Split (Abb. 2). Der Training-Split enthält 659.428 Punkte, der Validation-Split 94.448 und zum Testen stehen 411.722 Punkte zur Verfügung.

4.2 Vorverarbeitung

Im Gegensatz zu klassischen Verfahren ist hier keine separate Berechnung von Merkmalen erforderlich. Der einzige notwendige Vorverarbeitungsschritt ist das Voxelieren der Punktwolke. Dies dient zusätzlich der Homogenisierung der Punktdichte. Statt eines voll aufgespannten Voxelgitters bestimmen wir eine Liste nicht-leerer Voxel. Jedem Element dieser Voxelwolke werden die aus den innenliegenden Punkten gemittelten Echomerkmale zugewiesen. Die Klassen-Label werden durch Mehrheitsvotum bestimmt. Als Nebenprodukt des Voxelfilters fällt eine Indexliste an, mit der die prognostizierten Klassenwahrscheinlichkeiten von den Voxeln auf die ursprünglichen Punkte übertragen werden können. Wir untersuchen Voxelgrößen von 2 m, 1 m, 0,5 m, 0,25 m und 0,125 m. Zur künstlichen Erweiterung der Trainingsdaten werden

Trainings- und Validierung-Split vor der Voxelierung zwölfmal um je 30° um die Z-Achse rotiert. Um dem



Abb. 3: Voxeliertes Trainingsample mit Kantenlänge 32 × 32 × 64 m und Voxelgröße 1 m. Die Klassen sind farbcodiert gemäß in Abb. 5.

Netzwerk viele Lernbeispiele zeigen zu können und Überanpassung zu vermeiden, teilen wir die Punkt- bzw. Voxel-Wolken in kleinere Samples entlang eines horizontalen Gitternetzes auf. Die Samples müssen jedoch groß genug sein, um einen Aussagekräftigen räumlichen Kontext zu beherbergen. Wir experimentieren mit Samples von $16 \times 16 \times 64$ m, $32 \times 32 \times 64$ m und $64 \times 64 \times 64$ m Ausdehnung. Jedes Sample deckt so die volle vertikale Ausdehnung des Datensatzes ab. Der Überlapp der Training-Samples beträgt 30 %.

4.3 Training

Die Mini-Batch-Größe beim Training beträgt 128 für $16 \times 16 \times 64$ m große Samples. Da bei größerer Ausdehnung und gleicher Überlappung weniger Samples extrahiert werden können, die Anzahl an Parameterupdates pro Epoche jedoch vergleichbar bleiben soll, besteht ein Mini-Batch bei 32 m bzw. 64 m Kantenlänge aus 32 bzw. 8 Samples. Zur Optimierung verwenden wir Stochastic Gradient Descent mit Momentum und Weight Decay. Für jede Konfiguration werden 10 gleiche Netze unabhängig voneinander trainiert. Sie unterscheiden sich nur durch die zufällige Initialisierung ihrer Faltungskerne und die Reihenfolge der Training-Samples.

4.4 Inferenz

Standardmäßig werden die Validierungs- und Test-Sets auf die gleiche Weise in Samples aufgeteilt wie die Trainingsdaten, jedoch ohne Überlappung. Da bei der Inferenz von kleineren Samples an den Rändern möglicherweise wertvolle Nachbarschaftsinformation fehlen, nutzen wir die *fully convolutional* (LONG et al. 2015) Eigenschaft der Architektur, um die Genauigkeit von am Stück klassifizierten Punktwolken zu untersuchen. Für bessere und stabilere Ergebnisse verwenden wir zudem Ensembles aus zehn Netzen, deren prädizierte Klassenwahrscheinlichkeiten gemittelt werden. Zugewiesen wird einem Voxel die Klasse mit der höchsten Wahrscheinlichkeit.

4.5 Implementierungsdetails

Implementiert wurden die Experimente in Python 3.5 mit PyTorch³ 0.4 als Deep Learning Bibliothek. Das Framework für Submanifold Sparse Convolutional Networks von GRAHAM et al. (2018) ist öffentlich verfügbar⁴. Die Punktwolke wurde eingefärbt mittels OPALS (PFEIFER et al. 2014).

5 Ergebnisse und Diskussion

Tab. 1: Klassifizierungsgenauigkeit (Overall Accuracy, OA) für den Test-Set. Unter den Voxelgrößen ist jeweils die Anzahl an Voxeln im Test-Set sowie die reine Inferenz-Zeit für ein einzelnes Netz angegeben. Pro Konfiguration ausgewertet aus 10 Netzen.

r										
Voxel OA [%] Punkte OA [%] (Mittel ± StdAbw. Ensemble)		Voxel-Größe [m]								
		2,0 (26k in 0,2 s)	1,0 (85k in 0,3 s)	0,5 (210k in 0,6 s)	0,25 (320k in 0,8 s)	0,125 (374k in 1,2 s)				
Sample-Größe [m]	16 × 16 × 64	71,6 ± 0,7 74,1 76,3 ± 0,4 78,3	76,0 ± 1,3 78,0 80,2 ± 1,0 81,2	78,4 ± 1,8 80,6 80,3 ± 1,5 82,3	80,4 ± 1,3 82,9 80,7 ± 1,1 82,9	79,2 ± 1,3 82,0 79,2 ± 1,3 82,0				
	32 × 32 × 64	72,5 ± 0,6 75,9 76,7 ± 1,0 79,8	77,4 ± 0,6 79,5 81,4 ± 0,5 83,1	79,9 ± 0,7 81,9 81,6 ± 0,6 <u>83,5</u>	80,7 ± 0,7 83,4 81,0 ± 0,7 83,2	78,8 ± 2,0 82,5 78,8 ± 2,0 82,4				
	64 × 64 × 64	72,4 ± 1,2 75,0 77,0 ± 0,8 79,1	77,5 ± 0,8 79,7 81,4 ± 0,7 83,2	79,6 ± 0,7 81,7 81,4 ± 0,7 83,4	81,1 ± 1,0 83,4 81,5 ± 0,9 83,4	80,6 ± 1,4 83,6 80,5 ± 1,4 83,7				

In Tabelle 1 werden die Klassifizierungsgenauigkeiten bei verschiedenen Auflösungen und Sample-Größen gezeigt. Es ist klar zu erkennen, dass die Ensembles erwartungsgemäß die besten Ergebnisse liefern. Außerdem fällt die Auswertung auf den ursprünglichen Punkten bei gröberen Auflösungen besser aus als auf den Voxeln. Überschreitet die Auflösung die mittlere Punktdichte der ursprünglichen Punktwolke, ist kein signifikanter Unterschied mehr vorhanden. Das beste Ergebnis von **83,5 %** liefert eine Sample-Größe von $32 \times 32 \times 64$ m mit einer Voxelgröße von 0,5 m. Bei allen weiteren Untersuchungen bleibt dies die beste Konfiguration.

5.1 Inferenz via Samples vs. als Ganzes

Prädizierte man die Klassenlabel für den gesamten Test-Split am Stück, d.h. ohne zu samplen, verringerte sich die Klassifizierungsgenauigkeit bei Netzen, die auf $16 \times 16 \times 64$ m großen Sam-

³ https://pytorch.org/

⁴ https://github.com/facebookresearch/SparseConvNet

peln trainiert wurden, und zwar um durchschnittlich 1,8 Prozentpunkte. Bei 32 m bzw. 64 m Kantenlänge jedoch stieg sie um 0,8 bzw. 0,6. Das sich daraus ergebende beste Netzwerk hat dieselbe Konfiguration wie in Tabelle 1, erzielt jedoch **84,2 %** (Abb. 4a, 5b). Daraus lässt sich schließen, dass dieser Vorteil nur erzielt werden kann, wenn die Trainingsbeispiele ausreichend räumlichen Kontext besaßen.

5.2 Einfluss der Geometrie

Um den Einfluss der reinen Geometrie zu untersuchen, trainierten und testeten wir ein Netz der besten Konfiguration aus Tabelle 1, jedoch ohne die Echomerkmale. Jedes belegte Voxel besitzt dabei als einziges Merkmal den Wert 1. Die Klassifizierungsgenauigkeit beträgt **79,8 %** für ein Ensemble, bzw. etwa **75 %** für ein einzelnes Netz. Das größte Problem bereitet hier die Unterscheidung von Low Vegetation und Impervious Surface, beides sind Klassen mit sehr flacher und Bodennaher räumlicher Verteilung

5.3 Einfluss spektraler Merkmale

Das führende Verfahren in der Benchmark-Liste des ISPRS 3D Semantic Labeling Contest verwendet eine mit spektralen Informationen angereicherte Punktwolke. Zum Vergleich wiederholten wir die Versuche der Sample-Größe $32 \times 32 \times 64$ m ebenfalls mit zusätzlichen CIR-Werten aus dem zugehörigen TOP. Ein generelles Problem dabei sind die unterschiedlichen Aufnahmezeitpunkte von Bild und LiDAR-Punktwolke, was insbesondere bei Fahrzeugen zu falschen Einfärbungen führt. Des Weiteren werden Fassaden teilweiße wie die darüber liegenden Dächer eingefärbt.



Abb. 4: Fehlermatrizen der klassifizierten Test-Punktwolke.



(a) Ground Truth



(b) Klassifizierungsergebnis ohne CIR



(c) Klassifizierungsergebnis mit CIR

Abb. 5: Visueller Vergleich der klassifizierten Punktwolke, eingefärbt nach Klassen: Powerline (schwarz), Low. veg. (hellgrün), Imp. surf. (grau), Car (blau), Fence/Hedge (dunkelgrün), Roof (rot), Facade (weiß), Shrub (gelbgrün), Tree (grün)

Bei der sampleweisen Inferenz steigt die OA im Mittel um 1,9 Prozentpunkte, insbesondere aber bei einer Voxelgröße von 2 m. Die beste Konfiguration (siehe Tabelle 1) steigert sich von 83,5 % auf **84,6 %**. Wird die Test-Voxelwolke als Ganzes vom SSCN prozessiert, steigt die Genauigkeit im Schnitt um 1,5 Prozentpunkte, das beste Ergebnis um 0,8 auf **85,0 %** (Abb. 4b, 5c). Damit erreichen wir knapp das zum Zeitpunkt dieser Arbeit führende Ergebnis des Benchmarks, welches 85,2 % erzielt.

Fassaden werden zwar etwas häufiger als Dächer interpretiert, dafür aber seltener als Vegetation. Die Verwechslungen zwischen Straße und Fahrzeugen wird sogar geringfügig besser.

5.4 Rechenzeit und Speicher

Das Training dauert pro Netz für die verschiedenen Voxelgrößen ca. 6, 14, 30, 60 bzw. 100 Minuten. Die reine Inferenzzeit für das beste Ensemble (Voxelgröße 0,5 m) beträgt 11 s, hinzu kommen für Auswertung und Input/Output 19 s, weitere 0,1 s für die Voxelierung (plus 4 s für I/O), sowie gegebenenfalls 5 s fürs Sampling. Die letzteren Zeiten bieten implementierungs-bedingt noch einigen Raum zur Optimierung.

Bei größeren Datenmengen kann Parallelität der GPU besser genutzt werden. Wir vergrößerten den Testdatensatz durch Kopieren künstlich auf ca. 10 Millionen Punkte (Faktor 25). Die reine Inferenzzeit betrugt daraufhin bei ansonsten gleicher Konfiguration etwa 100 s, also nur das 8-fache.

Wenn geringe Einbußen bei der Genauigkeit akzeptabel sind, erlaubt die Einstellung der Voxelgröße und der Anzahl an Netzen im Ensemble eine simple Abwägung zwischen Rechenzeit und Genauigkeit.

Bei allen Berechnungen kam die folgende Hardware zum Einsatz: Intel-Core i7-6800K @ 6/12x 3,40 GHz mit 64 GB RAM und eine NVIDIA Titan X Pascal mit 12 GB Grafikspeicher.

6 Fazit und Ausblick

In dieser Arbeit konnten wir die Eignung von Submanifold Sparse Convolutional Networks zur semantischen Segmentierung von ALS-Punktwolken zeigen. Die erreichte Klassifizierungs-genauigkeit von 85,0 % ist das zum Zeitpunkt dieser Arbeit zweitbeste veröffentlichte Benchmark-Ergebnis.

Seltene Objektkategorien können durch eine gewichtete Zielfunktion trotzdem gut erkannt werden. In zukünftigen Arbeiten könnten weitere Ansätze wie die Dice Zielfunktion (MILLETARI et al. 2016) untersucht werden.

Die implizite Geometrie der Punktwolke hat sich als das primäre Merkmal erwiesen. Schwierige Klassen sind insbesondere Sträucher und Hecken bzw. Zäune, die oft als andere Vegetationsarten interpretiert werden. Niedrige Vegetation und versiegelte Flächen weisen aufgrund ihrer ähnlichen Geometrie starke Verwechslungen auf. Fassaden sind in diesem Datensatz nur schlecht repräsentiert und würden von einem stärker schräg messenden System profitieren.

Der in dieser Arbeit verwendete Datensatz ist, verglichen mit den Datensätzen, auf denen Deep Learning Methoden üblicherweise trainiert werden, sehr klein. Dies macht das Training instabil und Generalisierung schwierig. In weiteren Arbeiten werden wir die Anwendung auf größere Datensätze untersuchen.

7 Danksagungen

Der Vaihingen Datensatz wurde zur Verfügung gestellt von der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung (DGPF) [Cramer, 2010]: http://www.ifp.uni-stuttgart.de/dgpf/DKEP-Allg.html.

Die Titan X Pascal verwendet für diese Arbeit wurde gespendet von der NVIDIA Corporation. Der Autor dankt Philipp-Roman Hirt für seine Unterstützung bei der Voxel-Visualisierung.

8 Literaturverzeichnis

- BLOMLEY, R. & WEINMANN, M., 2017: Using Multi-Scale Features for the 3D Semantic Labeling of Airborne Laser Scanning. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-2/W4, 43-55.
- CHELLAPILLA, K., PURI, S. & SIMARD, P., 2006: High performance convolutional neural networks for document processing. Tenth International Workshop on Frontiers in Handwriting Recognition, Lorette, G. (Hrsg.), October, La Baule (France).
- CHEHATA, N., GUO, L. & MALLET, C., 2009: Airborne Lidar feature Selection for urban classification using random forests. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **38**(3/W8), 207-212.
- CHETLUR, S., WOOLLEY, C., VANDERMERSCH, P., COHEN, J., TRAN, J., CATANZARO, B. & SHEL-HAMER E., 2014: cuDNN: Efficient Primitives for Deep Learning. CoRR (arXiv:1410.0759).
- CRAMER, M., 2010: The DGPF-Test on Digital Airborne Camera Evaluation Overview and Test Design. Photogrammetrie-Fernerkundung-Geoinformation, **2010**(2), 73-82.
- EIGEN, D. & FERGUS, R., 2015: Predicting Depth, Surface Normals and Semantic Labels with a Common Multi-scale Convolutional Architecture. IEEE International Conference on Computer Vision, 2650-2658.
- GRAHAM, B., 2015: Sparse 3D convolutional neural networks. Proceedings of the British Machine Vision Conference (BMVC), Xie, X., Jones, M.W., Tam, G.K.L (Hrsg.), BMVA Press, September, 150.1-150.9.
- GRAHAM, B., ENGELCKE, M. & VAN DER MAATEN, L., 2018: Semantic Segmentation with Submanifold Sparse Convolutional Networks. IEEE Conference on Computer Vision and Pattern Recognition, 18-22.
- HACKEL, T., WEGNER, J.D. & SCHINDLER, K., 2016: Fast Semantic Segmentation of 3D Point Clouds with Strongly Varying Density. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, **III**-3, 177-184.
- HU, X. & YUAN, Y., 2016: Deep-Learning-Based Classification for DTM Extraction from ALS Point Cloud. Remote Sensing, **8**(9), Beitrag 730.
- LONG, J., SHELHAMER, E. & DARREL, T., 2015: Fully convolutional networks for semantic segmentation. IEEE Conference on Computer Vision and Pattern Recognition, 3431-3440.
- NIEMEYER, J., ROTTENSTEINER, F. & SÖRGEL, U., 2014: Contextual classification of lidar data and building object detection in urban areas. ISPRS Journal of Photogrammetry and Remote Sensing, **87**, 152-165.

- NIEMEYER, J., ROTTENSTEINER, F., SOERGEL, U. & HEIPKE, C., 2016: Hierarchical Higher Order CRF for the Classification of Airborne LiDAR Point Clouds in Urban Areas. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 41(B3), 655-662.
- MILLETARI, F., NAVAB, N. & AHMADI, S., 2016: V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. 2016 Fourth International Conference on 3D Vision (3DV), 565-571.
- PFEIFER, N., MANDLBURGER, G., OTEPKA, J. & KAREL, W., 2014: OPALS A framework for Airborne Laser Scanning data analysis. Computers, Environment and Urban Systems, 45(2014), 125-136.
- POLITZ, F. & SESTER, M., 2018: Exploring ALS and DIM Data for Semantic Segmentation using CNNs. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(1), 347-354.
- RIZALDY, A., PERSELLO, C., GEVAERT, C.M. & OUDE ELBERINK, S.J., 2018: Fully Convolutional Network for Ground Classification from LiDAR Point Clouds. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Science, IV-2, 231-238.
- RONNEBERGER, O., FISCHER, P. & BROX, T., 2015: U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MIC-CAI), 9351(LNCS), Springer Verlag, 234-241.
- YI, L., SHAO, L., SAVVA M., HUANG, H., ZHOU, Y. & WANG, Q. ET AL., 2017: Large-Scale 3D Shape Reconstruction and Segmentation from ShapeNet Core55. CoRR (ar-Xiv:1710.06104).

Accuracy Analysis of Digital Elevation Models from very High Resolution Satellite Imagery

ANA-MARIA LOGHIN¹, JOHANNES OTEPKA¹, WILFRIED KAREL¹, MARKUS PÖCHTRAGER¹ & NORBERT PFEIFER¹

Abstract: In recent years, the derivation of Digital Elevation Models (DEMs) from very high resolution stereo satellite imagery has become an interesting option and research topic, because the sub-meter GSD (Ground Sampling Distance) suggests that accuracy in the dm domain is possible. Current optical earth observation satellites are able to collect multi-view images of the same area during a single pass, and thus satellite imagery became an important source for 3D information extraction.

Photogrammetrically derived DEMs play a significant role in many engineering, land planning, geomorphology, forestry and infrastructure applications. To guarantee reasonable analyses in the multiple applications, surface models with high quality are needed. While previous research demonstrates the potential of Very High Resolution (VHR) stereo / tri-stereo satellite imagery for DEM reconstruction, questions about the accuracy of the derived models and their noise-induced roughness were not fully investigated. Therefore, the main focus of this contribution is to analyse and assess the potential in geometric quality of DEMs generated from high-resolution Pléiades and WorldView-3 stereo and tri-stereo scenes. We study the impact of the different acquisition geometries (stereo / tri-stereo, ground sample distance, viewing and incidence angles) on the estimated surface and its properties.

The study area is located in Allentsteig, Lower Austria, a hilly region covered by arable lands and coniferous forests stretching from 300 m to 690 m a.s.l. The entire photogrammetric workflow, comprising the satellite image triangulation, dense matching and 3D reconstruction were performed with the Match-AT and Match-T DSM modules of the Trimble Inpho software. The 3D reconstructed point clouds are interpolated into high resolution Digital Surface Models (DSMs) and their absolute vertical accuracy is evaluated against a LiDAR-derived Digital Terrain Model (DTM). We focus on the assessment analyses in free, open and smooth areas, without any vegetation or artificial structures like buildings or infrastructure. For this, a mask derived from available LiDAR DSM and DTM was used. The vertical quality of the reconstructed DEMs derived from the tri-stereo combination is analyzed with traditional and robust accuracy measurements, resulting in non-Gaussian distributions of errors, with a RMSE of 0.96 m (1.4 pixels) for Pléiades and of 0.37 m (1.2 pixels) for WorldView-3. When compared to a ground truth LiDAR DTM, the elevation differences show an undulation (~1.5 pixel), similar to waves that are visible in the along track direction. In order to minimize this effect and the vertical error caused by horizontal and vertical offsets, the photogrammetrically derived DEMs are aligned to the reference DTM by applying an affine 3D transformation determined with the least squares matching (LSM) techniques. The results show improvements in the vertical accuracy to 0.61 m (0.9 pixels) and 0.24 m (0.7 pixels) for Pléiades and WorldView-3 tri-stereo scenes, respectively, and a decrease of the "wave-effect" to less than one pixel.

¹ Technische Universität Wien, Department of Geodesy and Geoinformation (GEO), Gu
ßhausstra
ße 27-29/E120, A-1040 Wien, E-Mail: [ana-maria.loghin, Johannes.Otepka, Wilfried.Karel, markus.poechtrager, Norbert.Pfeifer]@geo.tuwien.ac.at

1 Introduction

Digital Elevation Models (DEMs), mathematically describing the terrain surface topography, play an essential role in GIS and environmental modelling, having applications in agriculture, geomorphology, forestry, natural resources, disaster management, urban planning and engineering. To guarantee reasonable analyses for multiple applications, surface models with high quality are needed. Airborne Laser Scanning (ALS) active systems are the primary data source for accurate DEM extraction of high spatial resolution (CONTRERAS et al. 2017; HABIB et al. 2005) with reported height accuracy between 15-25 cm (AGUILAR & MILLS 2008; GATZIOLIS & ANDERSEN 2008; LIU et al. 2007). The great advantage of this technology is the ability of capturing directly three-dimensional information in object space, based on the flight of a light signal between the sensor and the target (VOSSELMAN & MAAS 2010). Nevertheless, the major limitations of ALS data are related to access restrictions and low temporal resolution. On the other hand, multi view aerial imagery is the traditional and cost efficient method to acquire high resolution elevation data over large surface areas. Depending on the Ground Sample Distance (GSD) of the used sensor, the best vertical accuracies of airborne photogrammetrically derived DEMs vary between 0.44 and 2 GSD (BÜHLER et al. 2012; HOBI & GINZLER 2012; HOFFMANN & LEHMANN 2000; HU et al. 2008; RESSL et al. 2016). Since more than thirty years, the interest of the scientific community has turned to the potential of satellite imagery for DEM extraction in a timely and cost-effective manner, with the possibility of providing almost real time data. Compared with airborne remote sensing, the major advantages of satellite imagery is the large area coverage within a very short time (seconds), the worldwide availability without any limitations or access restrictions and the high temporal resolution (few days interval). The possibility of DEM extraction from satellite stereoscopic images started in 1986, with the launch of the first SPOT series satellites (TOUTIN & CHENG 2002). Since 1999, thanks to the continuous development of the remote sensing systems, spatial resolutions in the range of less than one meter can now be achieved by means of satellite-supported images. Very High Resolution (VHR) optical sensors are able to acquire images comparable with those provided by aerial imagery, in terms of high GSD values (0.30 m). Moreover, they are able to collect not only stereo, but tri-stereo images for the same area during a single flight pass. Due to their short revisit time and stereo/tri-stereo capability, today, the new generation of VHR pushbroom satellite sensors, e.g., QuickBird, SkySat, Ziyuan-3A, GeoEye-1, Pléiades 1A/1B, WorldView-1,-2,-3 and -4 are used for DEM extraction.

In this study, we consider two different VHR tri-stereo satellite systems, namely Pléiades and WorldView-3 for our analyses. The Pléiades constellation comprises two identical satellites, Pléiades-1A and Pléiades-1B, that were launched in December 2011 and December 2012, respectively. Both satellites are flying at 694 km altitude, in sun-synchronous orbits (98.2° inclination), with an offset of 180° from each other, which provides a daily revisit interval (ZHOU et al 2015). For the tri-stereo acquisitions, the B/H ratios (baseline/height) vary between 0.1 and 0.5 with corresponding stereo angles of ~6° to 28° (DE LUSSY et al. 2006). The sensor is able to collect panchromatic and multispectral images at nominal resolutions of 0.5 m and 2 m, respectively. The WorldView-3 high resolution commercial satellite was launched in August 2014. Operating at an altitude of 617 km, the sensor has an average revisit time of less than one day and is capable to collect up to 680 000 km² daily. WorldView-3 provides panchromatic, multispectral and short wave infrared imagery of 0.31 m, 1.24 m and 3.7 m resolutions, respectively, both in stereo (forward and back-ward view) and tri-stereo (forward, nadir and backward view) acquisition mode.

The potential of surface modelling from VHR satellite imagery and the advantages of using tristereo instead of stereo images for generating height models has been addressed in various publications. BERNARD et al. (2012) showed that by using tri-stereo configurations, reliable surface models can be obtained in urban areas. For a number of 295 ground control points, they report a Root Mean Square Error (RMSE) of 0.49 m in height. PERKO et al. (2014) assessed the mapping potential of Pléiades stereo/tri-stereo scenes, obtaining 3D geolocation accuracies of 0.5 m in planimetry and of 1 m in height. The results obtained by PANAGIOTAKIS et al. (2018) show that the DSM produced with tri-stereo analysis has better performances in terms of RMSEz (1.17 m), compared with the stereo-pair DSMs that perform worse (RMSEz of 1.48 m for forward-backward and of 1.64 m for forward-nadir). Additionally, DEMs derived by using Pléiades stereo/tri-stereo acquisitions were used for estimating lava flow volumes (BAGNARDI et al. 2016) and changes in height produced by earthquakes (ZHOU et al. 2015). After coregistration procedures based on crosscorrelation and iterative closest point, ZHOU et al. (2015) reported a 0.5 m standard deviation of heights, when comparing the Pléiades DEM with an airborne LiDAR DEM. The DSM extraction from WorldView-3 stereo-images is evaluated in Hu et al. (2016), where the elevation biases between the generated DEM and 7256 LiDAR check points are about 0.62 m.

In contrast to the previous research that demonstrates the capacity of Very High Resolution (VHR) satellite imagery for DEM reconstruction, in our study we investigate the potential of Pléiades and WorldView-3 stereo and tri-stereo scenes for DEM generation in open free, areas, where the DSM can be considered as DTM. Moreover, our study is, to the best to our knowledge, the first to assess the accuracy of DSMs derived from WorldView-3 tri-stereo scenes.

Specifically, our research study has as main purpose finding the answers to the folowing questions: (1) what is the impact of the different acquisition geometries (stereo / tri-stereo, ground sample distance, viewing and incidence angles) on DEM derivation, (2) which image combination produces the highest DEM accuracy and (3) what is the influence of spatial resolution and interpolation on the estimated surface roughness? To do so, we used Pléiades and WorldView-3 tri-stereo satellite imagery over a ~214 km² study area, located in Lower Austria. The absolute vertical accuracy of the photogrammetrically derived DEMs is evaluated against measured GCPs by means of Real Time Kinematic (RTK) and a LiDAR derived DTM. We focus on the assessment analyses in free, open and smooth areas, without any vegetation or artificial structures like buildings or infrastructure

2 Study area and Image Data Sets

The study area, located in Allentsteig, Lower Austria (48° 30' 30"N; 15° 08' 34"E; WGS84, 33N) is a hilly region with elevations between 300 m and 690 m a.s.l. The territory is covered mainly by arable lands and coniferous forests, the urban, suburban and rural areas being also present, but covering only a small area, compared with the entire extend. For the present study area, analyses were conducted using tri-stereo satellite images from two different VHR optical sensors: Pléiades-1B and WorldView-3 (Fig. 1). Each triple consists of three images, acquired from different along-track positions of the satellite: forward (F), close to nadir (N) and backward view (B).



Fig. 1: Study area and acquisition geometries of Pléiades (blue) and WorldView-3 sensors (orange) (view of the satellite positions and acquisition footprints in Google Earth)

The Pléiades tri-stereo data for Allentsteig area were acquired in the morning of 13 June 2017, around 10:10 a.m in North-South direction. Due to the high speed (~ 7.5 km/s) the sensor mounted on PHR1B platform was able to cover 159 km² within 23 seconds. The average viewing angles in along-track direction with respect to the nadir are 5.66°, 0.46° and -5.19°, respectively, for the forward, nadir and backward Pléiades images (Fig. 2). Based on the total travelled distance (167.8 km) and the orbit height (694 km), the baseline to height ratios (B/H) are of 0.13 (FN), 0.11 (NB) and 0.24 (FB) for each pairwise combination, with the corresponding convergence angles on the ground of 7.5° (FN), 6.3° (NB) and 13.8° (FB).

Sensor type			GSD	View	ing Angle/	es (°)	B / H Ra- tio	Conver	
& Acquisition date	Acquisition Time	View	(m) Along track	Across	Along	Overall		gence An- gle (°)	
Diáiadaa	10:09:51.5	Forward	0.71	3.15	5.66	6.47	0.13(FN)	7.5 (FN)	
	10:10:03.7	Nadir	0.71	3.62	0.46	6.32	0.11(NB)	6.3 (NB)	
13-00-2017	10:10:14.0	Backward	0.70	3.37	-5.19	3.41	0.24 (FB)	13.8 (FB)	
WorldView-	10:22:07.0	Forward	0.32	7.70	11.00	13.50	0.22(FN)	12.7 (FN)	
3	10:22:25.5	Nadir	0.31	7.23	-0.62	7.36	0.22(NB)	12.8 (NB)	
08-04-2018	10:22:44.1	Backward	0.32	6.72	-12.20	13.97	0.45 (FB)	25.5 (FB)	

Tab. 1 Acquisition properties for Pléiades and WorldView-3 data for the study area

To comparatively investigate the accuracy potential of DEM derived from VHR satellite imagery and the impact of the acquisition geometry, we tasked a new tri-stereo WorldView-3 dataset in the same area (100 km²), with an overlapping of 44.5 km² with the Pléiades images. According to the metadata, the acquisition was done on 8 April 2018, around 10:22, within 37 seconds. The alongtrack acquisition angles w.r.t. nadir are 11°, -0.6° and -12° for the forward, close to nadir and forward images, respectively (Fig. 2). Compared to Pléiades sensor, WorldView-3 platform flies at a lower altitude (617 km), a fact that leads to higher values for the B/H ratios, even if the satellite speed is almost the same. Thus, the B/H ratios for the FN and NB are both 0.22, whereas for FB is 0.45. The intersection angles of rays on the ground for each combination have higher values as well: 12.7° (FN), 12.8° (NB) and 25.5° (FB). The images were delivered in tiles: 8, 2 and 2 for nadir, forward and backward, respectively. The operational aspects of acquisition influence the radiometric characteristics of the images. The fast recording (less than one minute) guarantees same sun illumination conditions with no significant shadow changes in the scenes. From a visual inspection, some radiometric inconsistencies can be observed in the WorldView-3 forward image, i.e. saturation and spilling effects (Fig. 2), mainly caused by the reflective roof surface, in combination with the imaging incidence angle.

All optical satellite images were delivered as 16 bit pansharpened with four bands (Red, Green, Blue and Near-infrared). Their acquisition properties are given in Tab. 1. Depending on the viewing angle, the spatial resolutions (mean values for the GSD) for Pléiades images are varying between 0.70 and 0.71 m, whereas the WorldView-3 sensor provides images with higher resolutions, between 0.31 and 0.33 m (Fig. 2, magnified detail).



Fig. 2: Zoom in built-up area in the three Pléiades images (first row) and in the three WorldView-3 images (second row), acquired with forward-, nadir- and backward-viewing, with a magnified detail

3 Data Analyses and Processing

3.1 Reference Data

For the current analyses, we have used a number of 43 GCPs measured by means of Real Time Kinematic (RTK), with high accuracies (up to 1 cm). Additionally, reference data are based on a DTM and a DSM with 1 m spatial resolution, derived from an ALS flight measurement campaign in December 2015. These LiDAR models were used to compute a reference mask for the open, smooth surfaces. The LiDAR DTM vertical accuracy was checked against the RTK GCPs showing a σ_z of 0.12 m. The digital terrain model was used to compute the Pléiades and WorldView-3 nDSMs (normalized DSM) and to improve the absolute geolocation of the photogrammetrically derived DEMs. We used a digital orthophoto from 2017 at 0.20 m spatial resolution for defining the positions of new Ground Control and Check Points (CPs), whose corresponding heights were extracted from the LiDAR DTM, at the same location.

3.2 Satellite Image Processing and DEM derivation

3.2.1 Pléiades and WorldView-3 image processing

The workflow as applied in this study starts with satellite image triangulation and dense image matching, followed by 3D reconstruction, DEM interpolation, geolocation improvement and accuracy assessment (Fig. 3). The sensor model used for image orientation is based on eighty Rational Polynomial Coefficients (RPCs), available as metadata information. They are used as an alternative to the rigorous sensor model and allow the conversion between image and object space coordinates (POLI & TOUTIN 2012). The reported geo-location accuracy for the Pléiades RPCs is of 8.5 m CE90 (circular error at 90% confidence) corresponding to the nadir view (ASTRIUM 2012), whereas for the WorldView-3 is of 3.1 m CE90 (COMP & MULAWA 2015). For obtaining a submeter accuracy, the RPCs are refined by including GCPs in the workflow. In order to get a homogenous distribution over the entire area, besides the RTK GCPs, we additionally measured 7 and 14 GCPs in the reference orthophoto for Pléiades and WorldView-3 tri-stereo scenes, respectively. The target positions of these points correspond to fixed details on the ground, such as: corners of field and pavement boundaries, road surface changes and intersections. Their accuracy depends on the orthophoto planimetric accuracy (0.20 m) and on the DTM vertical accuracy (0.12 m). In total, we employed 50 GCPs for Pléiades and 36 GCPs for WorldView-3. During image orientation, Tie Points (TPs) are automatically extracted using a multi-ray image matching method.



Fig. 3: Workflow for the processing chain

Dense Image Matching was performed with the specialised Match-T DSM module of the Trimble Inpho software. During processing, a number of ten pyramid levels were generated: the higher seven levels adopt a Feature Based Matching (FBM), while the last three ones a Cost Based Matching (CBM) strategy. The CBM is similar to the Semi-Global Matching algorithm (HIRSCHMÜLLER

2008), that generates an object point for every pixel. The same processing chain was followed for both Pléiades and WorldView-3 triplets.

3.2.2 DEM derivation and geo-location improvement

Dense image matching was performed for all three images (forward-nadir-backward, FNB) of the triplet and for each stereo pair i.e., forward-backward (FB), forward-nadir (FN) and nadir-backward (NB). The object point coordinates were determined by forward intersections. In a last step, regular raster models of height values (i.e., DSM) at 0.5 m resolution are generated by using a robust moving planes interpolation. According to the input point density we have chosen this approach to be the optimal in terms of surface detail preservation.

After a quality checking, systematic errors were visible between the reconstructed DEMs and the LiDAR DTM. In order to reduce them, we applied a least squares matching (LSM) technique that estimates an affine 3D transformation between each pair of target DEM and the reference DTM over common open areas. We masked out the areas containing vegetation, rivers, lakes, buildings and other artificial objects on the ground, and only used smooth surfaces, which can be considered to be stable; here the differences in height are expected to be as low as possible. These areas were identified based on the absolute values of the reference nDSM LiDAR data that needed to be less than 10 cm. The percentage values of stable areas within each dataset are of approximately 37.8 % and of 32.8 % for Pléiades and WorldView-3 scenes. Subsequently, the resulted LSM parameters were used to transform each Pléiades and WorldView-3 point clouds, which were then reinterpolated into new DEMs. The DEM interpolation and geolocation improvement were conducted in the scientific software OPALS (Orientation and Processing of Airborne Laser Scanning) (PFEIFER et al. 2014).

3.3 Quality Assessment

For both tri-stereo scenes, in a first step, the orientation accuracy was assessed by considering the RTK GCPs and 50 CPs measured in the available orthophoto and LiDAR DTM.



Fig. 4: (a) Orthophoto visualised as true color RGB with overlaid RTK GCPs (red circles) and manually measured GCPs (red circles) (b) colour coded and (c) shaded reconstructed DEMs

Secondly, we evaluated the vertical quality of the resulted DEMs, from each image combination, against the available elevations of the ground checkpoints and the reference DTM in the open areas. Thus, the vertical accuracy in open areas was determined by (1) computing the RMSEs between reference and obtained Z coordinates in each generated DEM and (2) analysing the histogram of difference values that should be around zero. For this, the distribution histogram was investigated by deriving statistic parameters such as mean, standard deviation, a robust standard deviation (σ_{MAD}) and also RMSE.

4 Results and disscusion

4.1 Image orientation and matching results

For both scenes, the bundle adjustment was performed using all three images, RTK GCPs and automatically extracted Tie Points (TPs). To assess the image orientation result, we used the available RTK GCPs and 50 CPs homogenously distributed in each tri-stereo scene.

For Pléiades sensor, automatic tie point extraction identified 561, 552 and 582 TPs for Forward, Nadir and Backward images, respectively. The computed RMSE of the GCPs is at decimetre level, except the value in vertical direction which reaches 0.27 m (Tab. 2). For WorldView-3 images, 556, 585 and 554 TPs were automatically extracted in each image. The GCPs discrepancies are at sub-decimetre level in planimetry, whereas a larger value of 0.11 m was obtained in the vertical direction. The accuracy of the CPs for both sensors is similar, around 0.5 m. The standard deviations for the bundle adjustments are 0.54 and 0.46 pixels for Pléiades and WorldView-3, respectively.

Sensor type	No. of GCPs /	RM	IS Values (meters / pixe	els)
Sensor type	CPs	Х	Y	Z
Dláiadaa	43 GCPs	0.187 / 0.26	0.196 / 0.28	0.273 / 0.39
Pleiades	50 CPs	0.441 / 0.63	0.532 / 0.76	0.850 / 1.21
WorldView 2	22 GCPs	0.063 / 0.21	0.086 / 0.29	0.111 / 0.37
worldview-3	50 CPs	0.424 / 1.41	0.403 / 1.34	0.584 / 1.95

Tab. 2: RMSE values for GCPs and CPs after bundle block adjustment

Depending on the incidence angles combinations, four different dense 3D point clouds resulted for each dataset: FNB, FB, FN and NB (Tab. 3) in LAS file format. Image matching processes were performed using the same machine with a core of 3.50 GHz and 32 GB RAM.

	Pléia	ades		WorldView-3				
Scene	Matching	LAS file	No. of points	Scene	Matching	LAS file	No. of	
Comb.	Time (h)	(GB)	(mil.)	Comb.	Time (h)	(GB)	points (mil.)	
FNB	10	15	667	FNB	33	27	1049	
FB	9	15	582	FB	21	26	1048	
FN	10	15	584	FN	24	26	1048	
NB	10	15	597	NB	24	26	1048	

Tab. 3: Image matching results

In the case of Pléiades, the FNB triplet provides the largest point cloud (~667 million points), compared with the three stereo pairs having ~582 (FB), ~584 (FN) and ~597 (NB) million points. Approximately same number of 3D points are generated for all four WorldView-3 combinations

(~1048 million points). Overall, the reconstructed point clouds have a regular distribution, with one point per each image pixel and resulted densities of 4 points / m^2 for Pléiades and of 12 points / m^2 for WV-3.

The resulted photogrammetric point clouds are used as input for DEM derivation. Usually the transition from 3D points to 2.5D grid models is done through interpolation that tends to smooth the input elevation values. The accuracy of obtained models highly depends on the adopted interpolation strategy and the choice of grid resolution. We have used a robust moving planes interpolation that estimates a best fitting tilted plane for every point considering a search radius of 1 m. For minimizing the smoothness effect, we selected a small grid size of 0.5 m, 10 and 20 neighbours (within the 1 m circular neighbourhood) for Pléiades- and WorldView-3-derived point clouds, respectively. This interpolation strategy is actually a compromise between achieving fidelity to the true surface and respecting the limitations according to the density and accuracy of the data source. The nDSMs were derived by subtracting the reference LiDAR DTM from the Pléiades and WV-3 DEMs, respectively.

4.2 Influence of acquisition geometry on DEM quality

4.2.1 Quality assessment of Pléiades DEMs

For each Pléiades image combination, the vertical quality of the photogrammetrically derived DEMs is assessed using the available ground check points, before and after applying the LSM transformation (Tab.4). The RMSE values before LSM for GCPs vary between 0.27 m (FNB) and 0.35 (NB). The LSM transformation significantly reduced the RMSE values for both GCPs and CPs in each image combination, the smaller value being of 0.20 m (FNB).

No. of GCPs /				RMS	E (meters)			
CPs		Bef	ore LSM		After LSM			
	FNB	FB	FN	NB	FNB	FB	FN	NB
43 GCPs	0.27	0.28	0.30	0.35	0.19	0.20	0.23	0.25
50 CPs	0.31	0.31	0.36	0.38	0.25	0.26	0.32	0.34

Tab. 4: RMSE values in Z-direction for GCPs and CPs for Pléiades DEMs

Finally, the interpolated DEMs in open, free areas for each combination, were compared with the ALS reference data. From a visual inspection, after the LSM transformation we obtain a good agreement between DEMs and reference DTM (Fig.5).

Tab. 5 Accuracy assessment of Pléiades DEMs in open areas (values are in meters)

Scene		Befo	ore LSM		After LSM			
Comb.	Mean	Std	σμαρ	RMS	Mean	Std	σμαρ	RMS
FNB	0.80	0.53	0.51	0.96	0.15	0.60	0.50	0.61
FB	0.77	0.53	0.51	0.93	0.13	0.58	0.50	0.60
FN	0.72	0.65	0.68	0.97	0.13	0.75	0.72	0.76
NB	0.78	0.70	0.73	1.04	0.17	0.79	0.80	0.81

The statistic results show that the DIM terrain heights are systematically shifted by ~ 1 m, but they are reduced to a median close to zero after applying the LSM transformation. The FNB and FB image combinations show a slightly better accuracy than the FN and NB (Tab. 5).



Fig. 5: Pléiades nDSM statistics for open areas, before (top) and after (bottom) LSM. Left: color-coded height differences (masked areas are shown in grey). Right: nDSM height distribution for all combinations



Fig. 6: Analysis of a section profile in along-track direction before and after LSM between pairwise difference models from Pléiades satellite imagery

When comparing the resulted DEMs to the ground truth LiDAR DTM and the three stereo combinations among them, the elevation differences show an undulation similar to waves that are visible in the along-track direction (Fig. 6), with a wavelength of ~4500 m. In this case, LSM transformation slightly reduces the maximum difference elevation values. Except the NB-FN difference model, the wave amplitude values reach ~1.5 pixels.

4.2.2 Quality assessment of WorldView-3 DEMs

In a first phase, the WorldView-3 DEM accuracy is estimated by contrasting the GCPs and CPs elevations with those extracted from the computed DEM (Tab. 6), with improved results after LSM (best accuracy for GCPs in the FNB combination, of 0.10 m). The resulted statistics for CPs have higher values when compared with GCPs, due to the influence of DTM accuracy ($\sigma_z = 0.12$ m), since this served as base for CPs height extraction.



Tab. 6: RMSE values in Z-direction for GCPs and CPs for WorldView-3 DEMs

Fig. 7: WorldView-3 nDSM statistics for open areas before (top) and after (bottom) LSM. Left: colorcoded height differences (masked areas shown in grey). Right: nDSM height distributions
Scene	Scene Before LSM				After LSM					
Comb.	Mean	Std	σμαρ	RMSE	Mean	Std	σμαρ	RMSE		
FNB	-0.16	0.33	0.33	0.37	0.03	0.24	0.19	0.24		
FB	-0.17	0.34	0.33	0.38	0.02	0.24	0.20	0.24		
FN	-0.17	0.37	0.34	0.41	0.03	0.33	0.34	0.33		
NB	-0.11	0.46	0.48	0.47	-0.01	0.34	0.34	0.34		

Tab. 7 Accuracy assessment of WorldView-3 DEMs in open areas (values are in meters)

Finally, the interpolated DEMs in open, free areas were compared with the ALS reference data. Like for the Pléiades analysis, here also the pairwise comparisons allow the computation of statistic parameters. The resulting non-Gaussian distributions of the normalized elevations (nDSM) for the tri-stereo and stereo images are shown in Fig. 7 before and after LSM. Initially with two peaks and a RMSE of 0.47 m, the histogram corresponding to NB combination is corrected by applying the LSM transformation, achieving a lower value of 0.34 m. Better and similar results are obtained for the FNB and FB image combinations after LSM (with RMSEs of 0.24 m), with lower dispersions around zero.

Similar to Pléiades, by a visual inspection of the computed nDSMs, we could see the same wave effect in the along-track direction, with a wavelength of ~1500 m and a maximum amplitude of ~1.5 pixels for the FB-FN and NB-FB difference models (Fig. 8). By applying the 3D transformation, this effect is reduced to less than 1 pixel, except for the NB-FN difference model.



Fig. 8: Analysis of a section profile in along-track direction before and after LSM between pairwise difference models from WorldView-3 satellite imagery

4.3 Comparative analysis of Pléiades and WorldView-3 DEMs

To complete the analysis, the DEM generated from the Pléiades triplet was compared to the model obtained from WV-3 triplet in the overlap area of about 44.5 km². The distribution histogram in open areas before LSM shows a positive shift, with a mean of 1.35 m, a robust standard deviation of 0.47 m and a RMS of 1.43 m, whereas after transformation the statistic parameters decrease to 0.02 m, 0.43 and 1.15 m.



Fig. 9: Height differences between Pléiades and WV-3 tri-stereo DEMs with distribution histograms before and after LSM (masked areas shown in grey)

5 Conclusions

In this study, we focus on the potential of both Pléiades and WorldView-3 tri-stereo pansharpened images for surface reconstruction in open, smooth areas, where the DSM can be considered same with the DTM. Satellite image orientation and dense image matching were performed in Inpho Trimble software, which employes a high automation process. GCPs were used for improving the direct sensor orientation available as RPCs. Eight digital elevation models derived from different pairwise combinations were assessed with regard to their vertical accuracy using geodetic measurements (RTK GCPs), CPs and an ALS DTM as reference data. The quality analysis of photogrammetrically computed DEMs showed systematic errors, as well as non-normal distributions and a wavy effect in the along-track direction. Apparently the present solution in Match-AT is not able to fully exploit the benefit of given GCPs and a 3D LSM transformation needs to be appended. In future work this problem should be further investigated. The application of the affine 3D transformation brought improvements in the vertical accuracy of the tri-stereo DEMs from 0.96 m (1.4 pixels) to 0.61 m (0.9 pixels) for Pléiades and from 0.37 m (1.2 pixels) to 0.24 m (0.7 pixels) for WorldView-3. For both sensors, our computed RMSE values vary from 0.77 to 1.15 GSD after applying the LSM transformation. These are comparable with the reported vertical accuracies for airborne photogrammetrically derived DEMs, which are between 0.44 and 2 GSD (BÜHLER et al. 2012; HOBI & GINZLER 2012; HOFFMANN & LEHMANN 2000; HU et al. 2008; RESSL et al. 2016). When compared to 50 CPs, the elevation accuracies of the Pléiades DEMs (RMSEs of 0.31 m, 0.31 m, 0.36 m and 0.38 m corresponding to FNB, FB, FN and NB, respectively) are similar with the results obtained by BERNARD et al. (2012), who achieved a RMSE of 0.49 m. In accordance with the investigations reported by PIERMATTEI et al. (2018) and PANAGIOTAKIS et al. (2018), we obtain higher accuracies for the tri-stereo and forward-backward combination compared with the other two stereo pairs that perform worse. Compared with Pléiades, the small ground sampling distance (0.31 m) of the WorldView-3 images leads to higher accuracies, as the representation of details on ground surface is more reliable. Our vertical accuracy results agree with those reported by HU et al. (2016), of less than 0.5 m for 6001 ground LiDAR check points. The small acquisition convergence angles influence the vertical accuracy, as the images are very similar to each other, a fact that enables a better image matching, despite the decrease in geometric intersection quality. By careful georeferencing including LSM over stable areas,

change detection between Pléiades and WorldView-3 satellite images is limited by an accuracy of only 0.50 m. With accuracies comparable to airborne photogrammetrically derived DEMs, the VHR satellite images are an important alternative data source for high resolution digital elevation models derivation over large areas.

6 References

- AGUILAR, F.J. & MILLS, J., 2008: Accuracy Assessment of LiDAR-Derived Digital Elevation Models. The Photogrammetric Record, **23**, 148-169.
- ASTRIUM, 2012: Pléiades Imagery User Guide V 2.0.
- BAGNARDI, M., GONZÁLEZ, P.J. & HOOPER, A., 2016: High-resolution digital elevation model from tri-stereo Pleiades-1 satellite imagery for lava flow volume estimates at Fogo Volcano. Geophysical Research Letters, 43(12), 6267-6275.
- BERNARD, M., DECLUSEAU, D., GABET, L. & NONIN, P., 2012: 3D capabilities of Pleiades satellite. International archives of the photogrammetry, Remote Sens Spat Inf Sci, **39**, 553-557.
- BÜHLER, Y., MARTY, M. & GINZLER, C., 2012: High Resolution DEM Generation in High-Alpine Terrain Using Airborne Remote Sensing Techniques. Transactions in GIS, **16**(5), 635-647.
- COMP, C. & MULAWA, D., 2015: WorldView-3 Geometric Calibration. Digital Globe, JACIE Workshop, Tampa, Florida.
- CONTRERAS, M.A., STAATS, W., YIANG, J. & PARROTT, D., 2017: Quantifying the Accuracy of Li-DAR-Derived DEM in Deciduous Eastern Forests of the Cumberland Plateau. Journal of Geo-graphic Information System, 9, 339-353.
- DE LUSSY, F., GRESLOU, D., DECHOZ, C., AMBERG, V., DELVIT, J.M., LEBEGUE, L., BLANCHET, G. & FOUREST, S., 2012: Pleiades HR in flight geometrical calibration: Location and mapping of the focal plane. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci, **39**, 519-523.
- HABIB, A., GHANMA, M., MORGAN, M. & Al RUZOUQ, R., 2005: Photogrammetric and LiDAR data registration using linear features. Photogrammetric Engineering & Remote Sensing, 71(6), 699-707.
- HIRSCHMÜLLER, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. IEEE Transactionon on Pattern Analyses and machine Intelligence, **30**, 328-341.
- HOBI, M.L. & GINZLER, C., 2012: Accuracy assessment of digital surface models based on WorldView-2 and ADS80 stereo remote sensing data. Sensors, **12**(5), 6347-6368.
- HOFFMANN, A. & LEHMANN, F., 2000. Vom Mars zur Erde-die erste digitale Orthobildkarte Berlin mit Daten der Kamera HRSC-A. Kartographische Nachrichten, **50**(2), 61-71.
- HU, F., GAO, XM., LI, GY., LI, M., 2016: DEM Extraction from Worldview-3 Stereo-Images and Accuracy Evaluation. The International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 41, 327-332.
- HU, W., YANG, G. & YUAN, H., 2008: Application and accuracy evaluation of Leica ADS40 for large scale mapping. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci, **37**, 605-610.
- LIU, X., ZHANG, Z., PETERSON, J. & CHANDRA, S., 2007: LiDAR-derived high quality ground control information and DEM for image orthorectification. GeoInformatica, **11**(1), 37-53.

- 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.
- PIERMATTEI, L., MARTY, M., KAREL, W., RESSL, C., HOLLAUS, M., GINZLER, C. & PFEIFER, N., 2018: Impact of the Acquisition Geometry of Very High-Resolution Pléiades Imagery on the Accuracy of Canopy Height Models over Forested Alpine Regions. Remote Sensing, 10(10), 1542.
- POLI, D. & TOUTIN, T., 2012: Review of developments in geometric modelling for high resolution satellite pushbroom sensors. The Photogrammetric Record, **27**(137), 58-73.
- 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, **2016**(2), 57-73.
- TOUTIN, T. & CHENG, P., 2002: Comparison of automated digital elevation model extraction results using along-track ASTER and across track SPOT stereo images. Optical engineering, **41**(9), 2102-2107.
- VOSSELMAN, G. & Maas, H.G., 2010: Airborne and Terrestrial Laser Scanning. Whittles, Caithness, UK. 318.
- ZHOU, Y., PARSONS, B., ELLIOTT, J.R., BARISIN, I. & WALKER, R.T., 2015. Assessing the ability of Pleiades stereo imagery to determine height changes in earthquakes: A case study for the El Mayor-Cucapah epicentral area. Journal of Geophysical Research: Solid Earth, 120(12), 8793-8808.

Relevance Assessment of Spectral Bands for Land Cover and Land Use Classification: A Case Study Involving Multispectral Sentinel-2-like and Hyperspectral Data

MARTIN WEINMANN¹ & UWE WEIDNER¹

Abstract: In this paper, we comprehensively investigate the potential of both multispectral and hyperspectral data for land cover and land use classification and, in this context, we particularly focus on comparatively assessing the relevance of involved spectral bands for the considered classification task. We present a framework which comprises different data-driven techniques for assessing the relevance of spectral bands with respect to the given classification task. Based on the assessed feature relevance, our framework allows selecting relevant features as the basis for classification for which a Random Forest classifier is used. We also describe a transformation of given hyperspectral data to multispectral Sentinel-2-like data that are commonly used for large-scale land cover and land use classification. For performance evaluation, we provide classification results achieved for two standard benchmark datasets representing an urban area and an agricultural area, respectively.

1 Introduction

Land cover and land use classification is commonly performed on the basis of aerial or satellite imagery representing either multispectral or hyperspectral data. In the case of multispectral data, the number of spectral bands is relatively low and the spectral bands are wide. In contrast, the number of spectral bands is relatively high for the case of hyperspectral data, while the spectral bands themselves are narrow. Consequently, a more characteristic description of the spectral properties of the Earth's surface can be expected for the latter case in comparison to the use of multispectral data.

In contrast to multispectral data, where the neighboring spectral bands are well-separated by a sufficiently large margin across wavelengths, the spectral bands of hyperspectral data are directly next to each other so that the acquired reflectance values of neighboring spectral bands tend to be strongly correlated. This kind of redundancy typically has a negative impact on classification results, so that approaches for dimensionality reduction or band selection are commonly used. While dimensionality reduction techniques transform the given data to a new space, band selection techniques allow conclusions about relationships with respect to physical properties as they retain a subset of the original spectral bands which, in turn, can further be used for conclusions regarding a diversity of environmental applications.

In this paper, we comprehensively investigate the potential of both multispectral and hyperspectral data, and we thereby comparatively assess the relevance of involved spectral bands for the considered classification task. We present a framework which comprises different data-driven techniques for assessing the relevance of spectral bands with respect to the given classification

¹ Karlsruhe Institute of Technology, Institute of Photogrammetry and Remote Sensing, Englerstraße 7, D-76131 Karlsruhe, E-Mail: [Martin.Weinmann, Uwe.Weidner]@kit.edu

task. This includes both classifier-dependent techniques and classifier-independent techniques. The classifier-dependent techniques comprise 1) a sequential forward selection of spectral bands based on sequentially training a Linear Discriminant Analysis (LDA) classifier based on different sets of spectral bands and 2) the Mean Decrease in Permutation Accuracy (MDPA), a measure assessed during the training of a Random Forest classifier. The classifier-independent techniques comprise 1) a general relevance metric taking into account the relations between the given spectral bands and the defined classes to identify relevant spectral bands and 2) an approach taking into account both the relations between spectral bands and classes to identify relevant spectral bands and spectral bands. Based on the selected bands, we finally perform a pixel-based classification using a Random Forest classifier to quantify the effect on the achieved classification results.

As the available datasets contain either hyperspectral or multispectral data, we consider commonly used hyperspectral benchmark datasets for which a semantic labeling is available on a per-pixel basis, and we adequately transform these hyperspectral datasets to multispectral Sentinel-2-like data using the Sentinel-2 Spectral Response Functions (S2-SRFs). For performance evaluation, we consider two classification tasks, one focusing on the semantic interpretation of an urban area and one focusing on the semantic interpretation of an agricultural area. For both classification tasks, the data have been acquired from an airborne platform during a low-altitude flight campaign.

After briefly summarizing related work (Section 2), we present our framework in detail (Section 3) and we demonstrate its performance on two benchmark datasets (Section 4). This is followed by a discussion of the derived results (Section 5). Finally, we provide concluding remarks and suggestions for future work (Section 6).

2 Related Work

The semantic interpretation of hyperspectral imagery may easily be achieved via a pixel-wise classification relying on the reflectance values across all considered spectral bands. In this context, the reflectance values are defined as features and used as entries of a respective feature vector, while the classification may be based on well-known standard classifiers such as a Support Vector Machine (SVM) classifier (MELGANI & BRUZZONE 2004; CHI et al. 2008) or a Random Forest classifier (HAM et al. 2005; JOELSSON et al. 2005).

However, particularly for such high-dimensional classification tasks, it can often be observed that above a certain value a further increase of the number of involved features results in a (typically significant) decrease in predictive accuracy (MELGANI & BRUZZONE 2004; KELLER et al. 2016; BRADLEY et al. 2018). This effect is commonly referred to as the Hughes phenomenon (HUGHES 1968) and arises from the joint consideration of more or less relevant features and possibly even irrelevant or redundant features with respect to the considered classification task.

To address the Hughes phenomenon, either dimensionality reduction techniques or feature selection techniques are commonly applied. The latter have the advantage that they retain a meaningful subset of the original features with respect to the given task (i.e. the reflectance values corresponding to specific spectral bands with known wavelength, which may further be exploited for conclusions regarding feature engineering and for conclusions regarding environmental sciences), while discarding less relevant and/or redundant features. This, in turn, typically allows

gaining predictive accuracy and improving computational efficiency with respect to both time and memory consumption (GUYON & ELISSEEFF 2003; SAEYS et al. 2007; ZHAO et al. 2010). In the context of classifying hyperspectral data, such feature selection techniques resulting in the consideration of reflectance values across specific spectral bands have been used in several investigations (MELGANI & BRUZZONE 2004; LE BRIS et al. 2014; CHEHATA et al. 2014; KELLER et al. 2016; BRADLEY et al. 2018). Such techniques may also directly allow assessing the importance of single spectral bands for land cover and land use classification (LE BRIS et al. 2014; KELLER et al. 2016; WEINMANN & WEIDNER 2018). Typically, feature selection techniques are categorized with respect to filter-based methods, wrapper-based methods and embedded methods (GUYON & ELISSEEFF 2003; SAEYS et al. 2007; ZHAO et al. 2010; WEINMANN 2016).

The filter-based methods rely on evaluating relations between features and classes and possibly also among features. In this context, the relations are described via a score function which is directly applied to the given training data. When only focusing on relations between features and classes (univariate filter-based feature selection), the relations are quantified by comparing the values of a feature across all data points with the respective class labels, e.g. via the correlation coefficient, Gini index, Fisher score, or information gain. This allows ranking the features with respect to their relevance. When taking into account both feature–class relations and feature–feature relations (multivariate filter-based feature selection), it is also possible to remove redundancy to a certain degree. A respective approach is represented by Correlation-based Feature Selection (HALL 1999). In general, filter-based methods are classifier-independent and thus typically result in simplicity and efficiency.

The wrapper-based methods rely on the interaction with a classifier to select features based on their suitability for classification. This may be achieved via Sequential Forward Selection (SFS) where, beginning with an empty feature subset, it is successively tested which feature can be added to the feature subset so that the predictive accuracy of the classifier increases the most. Alternatively, a Sequential Backward Elimination (SBE) may be conducted where, beginning with the set of all features, it is successively tested which feature can be removed from the feature subset so that the predictive accuracy of the classifier increases the most. Alternatively, a Sequential Backward Elimination (SBE) may be conducted where, beginning with the set of all features, it is successively tested which feature can be removed from the feature subset so that the predictive accuracy of the classifier is decreased the least. In general, however, the exhaustive interaction with a classifier tends to cause a high computational burden.

The embedded methods rely on the interaction with a classifier which provides the capability to internally select the most relevant features during the training phase, e.g. a Random Forest classifier (BREIMAN 2001). In contrast to wrapper-based methods, the involved classifier has to be trained only once to allow concluding about the relevance of involved features and the computational effort is therefore still acceptable in comparison to wrapper-based methods.

Besides such feature selection techniques, a transformation of hyperspectral data to multispectral data may be applied, where particularly Sentinel-2-like data are of interest as they are systematically acquired for Earth observation with short revisit times (SPOTO et al. 2012). Such a transformation has been introduced for simulating Sentinel-2 and other multispectral imagery in general (THONFELD et al. 2012), for assessing agricultural land use based on simulated Sentinel-2 data (ELBERTZHAGEN et al. 2012) and for simulating Sentinel-2 products that are relevant for geological and soil analyses (VAN DER MEER et al. 2014). Furthermore, such a transformation allows reasoning about the potential of multispectral data in comparison to hyperspectral data for land cover and land use classification (WEINMANN et al. 2018).

3 Methodology

An overview of the applied methodology is shown in Fig. 1. Given a hyperspectral data cube, we first derive the original data representation and its transformation to multispectral Sentinel-2-like data. Subsequently, we focus on feature relevance assessment to draw conclusions about the relevance of considered spectral bands with respect to the classification task. In this context, we involve 1) a wrapper-based feature selection method relying on a Linear Discriminant Analysis (LDA) classifier, 2) an embedded feature selection method relying on the Mean Decrease in Permutation Accuracy (MDPA) assessed with a Random Forest (RF) classifier (BREIMAN 2001; LIAW & WIENER 2002), 3) a univariate filter-based feature selection method relying on a general relevance metric (GRM) proposed in (WEINMANN 2016) and 4) a multivariate filter-based feature selection (CFS) method (HALL 1999). Finally, a supervised classification based on a Random Forest (RF) classifier (BREIMAN 2001) is performed. In the following, we first focus on the two options for the data representation (Section 3.1). Subsequently, we explain the four applied data-driven techniques for feature relevance assessment (Section 3.2) and the used approach for supervised classification (Section 3.3) with more details.



Fig. 1: Overview of our framework with different components for data representation, feature relevance assessment and supervised classification.

3.1 Data Representation

In our work, we consider the original representation of hyperspectral data (Section 3.1.1) and its transformation to multispectral data in the form of Sentinel-2-like data (Section 3.1.2) as input for subsequent processing steps.

3.1.1 Original Data Representation

The straightforward approach consists in using the given representation of the data, where we simply concatenate the reflectance values across all given spectral bands to obtain the respective feature vectors on a per-pixel basis.

3.1.2 Transformation of Hyperspectral Data to Sentinel-2-like Data

We also transform the original representation of hyperspectral data to multispectral Sentinel-2-like data (WEINMANN et al. 2018). In this context, we first identify those spectral bands of the hyperspectral data which are within the 13 spectral bands corresponding to Sentinel-2 data. The reflectance values corresponding to these hyperspectral bands, in turn, are used to derive a reflectance value for each considered spectral band of the Sentinel-2 data. More specifically, we consider the weighted mean of the reflectance values, where the weights are determined via linear interpolation based on the Sentinel-2 Spectral Response Functions (S2-SRFs) which are depicted in Fig. 2.



Fig. 2: Visualization of the Sentinel-2 Spectral Response Functions (S2-SRFs), i.e. the measured spectral responses for each band of the Sentinel-2 MultiSpectral Instrument (MSI) (WEINMANN et al. 2018).

However, for a classification task focusing on the use of Sentinel-2 data to derive thematic maps with respect to a variety of land cover and land use classes, not all of the 13 given spectral bands contain valuable information (WEINMANN & WEIDNER 2018). The reflectance values corresponding to the spectral bands B₁ (center wavelength of 443 nm), B₉ (945 nm) and B₁₀ (1375 nm) are not considered, because they correspond to parts of the spectrum where the atmospheric transmission is low, e.g. due to ozone (O₃), oxygen (O₂) or water vapor (H₂O) which strongly affect the atmospheric transmissivity at specific wavelengths. Furthermore, the reflectance value corresponding to the spectral band B₈ (842 nm) is not considered, because it is overlapping with the spectral band B_{8a} but much wider and less characteristic. Consequently, we only use those reflectance values corresponding to the spectral bands B₂ (490 nm), B₃ (560 nm), B₄ (665 nm), B₅ (705 nm), B₆ (740 nm), B₇ (783 nm), B_{8a} (865 nm), B₁₁ (1610 nm) and B₁₂ (2190 nm).

3.2 Feature Relevance Assessment

Given the different data representations, we proceed with assessing the relevance of features (i.e. spectral bands) via an exhaustive interaction with a classifier (Section 3.2.1), via a classifier-internal metric (Section 3.2.2), and via classifier-independent schemes (Sections 3.2.3 and 3.2.4).

3.2.1 Wrapper-Based Sequential Forward Selection of Spectral Bands

A classifier-dependent feature relevance assessment may be achieved by sequential forward selection via interaction with a classifier. Thus, we derive a feature ranking by starting with an empty set of features and successively adding the feature that improves the predictive accuracy of the involved classifier the most when added to the given feature set. This procedure is repeated until all features have been added. Consequently, we obtain the rank r of a feature f_i with $i = 1, ..., N_f$ and we define the relevance $R \in [0,1]$ of the feature f_i according to

$$R(f_i) = 1 - \frac{r(f_i) - 1}{N_f - 1}$$

so that relevant features are characterized by values close to 1, while irrelevant features are characterized by values close to 0. In this context, we take into account that exhaustive comparisons of feature vectors via the Euclidean distance as performed with a Nearest Neighbor (NN) classifier result in a high computational burden, particularly when considering highdimensional feature vectors. Furthermore, we take into account that some classifiers such as a Support Vector Machine (SVM) classifier or a Random Forest (RF) classifier require an additional tuning of internal parameters during the training process (i.e. the kernel width and the cost parameter in case of the SVM classifier, or the number of decision trees, the maximum tree depth, etc. in case of the RF classifier), which also results in a high computational effort. Consequently, we involve a classifier which does neither require such a tuning nor rely on the evaluation of Euclidean distances between feature vectors. More specifically, the chosen classifier is represented by a Linear Discriminant Analysis (LDA) classifier relying on the principle of probabilistic learning. In the training stage, it is assumed that the instances of different classes follow a Gaussian distribution in the feature space. Accordingly, the training of the LDA classifier consists in fitting a multivariate Gaussian distribution to the given training data, where the parameters of a Gaussian distribution have to be estimated for each class. Due to a lack of knowledge about the behavior of single classes in the feature space, the same covariance matrix is assumed for each class so that only the means vary for the different classes. In the prediction stage, the class probabilities are evaluated for each feature vector to be classified and the label of the class with the maximum probability is selected.

3.2.2 Feature Ranking via an Embedded Method

To directly assess the relevance of the given spectral bands with respect to the considered classification task, we use a Random Forest classifier (BREIMAN 2001) which represents an embedded method for feature selection, since it allows assessing feature relevance via the consideration of the Mean Decrease in Permutation Accuracy (MDPA). In this regard, the main idea consists in training each decision tree of the Random Forest classifier on a randomly chosen subset of the training data, i.e. a bootstrap sample, and then performing the prediction for the data which are not in the bootstrap sample, i.e. the out-of-bag (OOB) data (LIAW & WIENER 2002). Subsequently, the OOB predictions of all trees are aggregated and an error rate is derived. To estimate the relevance of a feature, it is tested how much the prediction error increases if OOB data for that feature is randomly permuted, while all others are left unchanged. Using the MDPA as feature importance and sorting the features accordingly, we obtain the rank r of a feature f_i with

 $i = 1, ..., N_f$ and we define the relevance $R \in [0,1]$ of the feature f_i as done for the wrapper-based sequential forward selection method in the previous section.

3.2.3 Feature Ranking via a General Relevance Metric

To directly assess the relevance of the given spectral bands with respect to the considered classification task, we also apply a classifier-independent approach for feature ranking via a general relevance metric (GRM) (WEINMANN 2016). This metric is a compound metric defined on the basis of filter-based feature selection methods and makes use of the given training data. More specifically, each of the filter-based methods relies on a score function which evaluates relations between features and classes to distinguish between relevant and irrelevant features. For each feature, its values across all feature vectors given in the training data are concatenated to a vector whose "correlation" with the corresponding label vector is evaluated and represented by a realvalued score. In this context, different score functions may be used which address different intrinsic properties of the given training data (e.g. distance, information, dependency or consistency). To achieve a robust feature relevance assessment taking into account different intrinsic properties of the given training data, we involve seven score functions (Pearson correlation coefficient, Fisher score, Gini index, information gain, χ^2 -test, *t*-test and ReliefF) and derive a separate ranking with respect to each score function. For more details about these score functions, we refer to (WEINMANN 2016) and references therein. As a result, we get the rank r of a feature f_i with $i = 1, ..., N_f$ given the score function s_j . Averaging the ranks derived for the feature f_i across all $n_s = 7$ score functions yields the mean rank \bar{r} :

$$\bar{r}(f_i) = \frac{1}{n_s} \sum_{j=1}^{n_s} r(f_i | s_j)$$

On this basis, the relevance R of the feature f_i is defined according to

$$R(f_i) = 1 - \frac{\bar{r}(f_i) - 1}{N_f - 1}$$

with $R \in [0,1]$.

3.2.4 Filter-Based Selection of a Meaningful Subset of Spectral Bands

To directly derive a meaningful subset of the given spectral bands, we focus on feature subset selection. For this purpose, we use a filter-based feature selection method which is referred to as Correlation-based Feature Selection (CFS) (HALL 1999). This method takes into account 1) the correlation between features and classes to identify relevant features and 2) the correlation among features to identify and discard redundant features. More specifically, CFS exploits the average correlation $\bar{\rho}_{FF}$ between features and classes as well as the average correlation $\bar{\rho}_{FF}$ among classes to evaluate the relevance *R* of a feature subset comprising *n* features:

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

Here, the correlation metric is defined via the symmetrical uncertainty SU (HALL 1999) with

$$SU(X,C) = \frac{2 MI(X,C)}{E(X) + E(C)}$$

where X and C represent random variables for the features and classes, respectively. The term MI(X, C) represents the mutual information between X and C, while the terms E(X) and E(C) represent Shannon entropies indicating the distributions of feature values and classes, respectively. Deriving a suitable feature subset thus corresponds to maximizing the relevance R over the feature subset space in an iterative manner. This means that, given an initial feature subset, either a feature is added to the feature subset (forward selection) or a feature is removed from the feature subset (backward elimination) until the relevance R converges to a stable maximum.

3.3 Supervised Classification

For classification, we focus on a standard supervised classification based on a Random Forest (RF) classifier (BREIMAN 2001) which relies on the principle of ensemble learning. In the training stage, an ensemble of randomly trained decision trees is generated via bootstrap aggregating ("bagging") (BREIMAN 1996), i.e. various subsets of the training data are randomly drawn with replacement and an individual decision tree is trained for each subset. In the prediction stage, for a new feature vector to be classified, each decision tree casts a vote for one of the defined classes and the majority vote is selected to obtain the most probable class label. To select suitable values for the internal parameters of the Random Forest classifier (i.e. for the number of involved decision trees, the maximum tree depth, etc.), we conduct a grid search on a suitable subspace.

4 Experimental Results

In the following, we consider two classification tasks, one focusing on the semantic interpretation of an urban area and one focusing on the semantic interpretation of an agricultural area. First, we describe the used datasets (Section 4.1). Subsequently, we explain the conducted experiments and present the achieved results (Section 4.2).

4.1 Datasets

For our experiments, we use two benchmark datasets representing an urban area and an agricultural area. Both datasets are publicly available in a repository of hyperspectral remote sensing scenes: http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes

4.1.1 Pavia Centre Dataset

The Pavia Centre Dataset was acquired with the Reflective Optics System Imaging Spectrometer (ROSIS) during a low-altitude flight campaign over the city of Pavia, Italy. The considered scene corresponds to an urban area and the acquired data are represented in the form of two images with a size of 1096×223 pixels and 1096×492 pixels, respectively. Each pixel corresponds to an area of $1.3 \text{ m} \times 1.3 \text{ m}$ and comprises hyperspectral information on 102 spectral bands. In total, the Pavia Centre Dataset consists of about 784k pixels of which 7,456 have been labeled with respect to 9 semantic classes as shown in Fig. 3, while no labels are provided for the remaining pixels.



Fig. 3: Reference labels for the Pavia Centre Dataset: each pixel is characterized by reflectance values on 102 spectral bands. Unlabeled pixels are indicated in black.

4.1.2 Salinas Dataset

The Salinas Dataset was acquired with the Airborne Visible / Infrared Imaging Spectrometer (AVIRIS) during a low-altitude flight campaign over Salinas Valley in California, USA. The considered scene corresponds to an agricultural area which is mainly characterized by vegetables, corn, bare soils and vineyards, and the acquired data are represented as an image with a size of 512 \times 217 pixels. Each pixel corresponds to an area of 3.7 m \times 3.7 m and, after the removal of 20 water absorption bands, comprises hyperspectral information on 204 spectral bands. In total, the Salinas Dataset consists of about 111k pixels of which 54,129 have been labeled with respect to 16 semantic classes as shown in Fig. 4, while no labels are provided for the remaining pixels.





Fig. 4: Reference labels for the Salinas Dataset: each pixel is characterized by reflectance values on 204 spectral bands. Unlabeled pixels are indicated in black.

4.2 Experiments and Results

Our framework is tested on a standard laptop computer (Intel Core i7-6820HK, 2.7 GHz, 4 cores, 16 GB RAM, Matlab implementation). For the RF classifier, we use the implementation provided with (LIAW & WIENER 2002). This implementation also allows assessing the MDPA used for feature selection. For CFS and the score functions used for filter-based feature selection, we use implementations provided with (ZHAO et al. 2010), while we apply the GRM following the implementation used in (WEINMANN 2016).

First, we focus on the classification results derived with different configurations of our framework, i.e. with different feature (sub)sets provided as input to the RF classifier. In this context, we take into account that the wrapper-based feature selection method relying on a LDA classifier, the embedded feature selection method relying on the MDPA of a RF classifier and the GRM allow quantifying feature relevance and thus ranking the features according to their relevance with respect to the given classification task. However, unlike the CFS method, they do not retain a recommendation for the number of best-ranked features that should be used as the basis for classification. Hence, for the hyperspectral data, we follow (BRADLEY et al. 2018) and select the best-ranked features covering about 20% of all available features. For the Pavia Centre Dataset, we thus use subsets of the 20 best-ranked features, while we use subsets of the 40 best-ranked features for the Salinas Dataset. The quantitative classification results are provided in Tab. 1 in terms of overall accuracy (OA), κ -index (κ) and mean F1-score across all classes (mF1), and the qualitative classification results are provided in Fig. 5. With these results, we also provide results derived for the transformation of the hyperspectral data to multispectral Sentinel-2-like data.

Pavia Centre	OA (in %)	<i>к</i> (in %)	mF1 (in %)	Salinas	OA (in %)	κ (in %)	mF1 (in %)
All Features	96.50	95.05	90.50	All Features	87.13	85.69	92.19
LDA Wrapper	95.53	93.68	88.62	LDA Wrapper	86.38	84.87	91.37
RF-MDPA	96.22	94.65	89.59	RF-MDPA	86.95	85.49	92.16
GRM	86.68	81.69	76.03	GRM	85.72	84.16	91.13
CFS	96.48	95.01	90.60	CFS	87.14	85.71	92.32
Sentinel-2	96.37	94.86	90.01	Sentinel-2	85.84	84.26	91.08

Tab. 1: Classification results derived for different feature (sub)sets provided as input to a RF classifier.

In the next step, we focus on feature relevance assessment using the described methods, i.e. 1) the wrapper-based feature selection method relying on a LDA classifier, 2) the embedded feature selection method relying on the MDPA of a RF classifier, 3) the GRM and 4) the CFS. For the latter, we take into account that it retains a subset of features without a ranking. Hence, we assign the selected features a relevance of 1 and all others a relevance of 0. CFS retains 21 features for the Pavia Centre Dataset and 36 features for the Salinas dataset. For all involved methods, the derived relevance of the considered spectral bands is visualized in Fig. 6 for the Pavia Centre Dataset and in Fig. 7 for the Salinas Dataset. Furthermore, the feature relevance assessment results for the transformation to multispectral Sentinel-2-like data are depicted in Fig. 8.



Fig. 5: Classification results achieved for the Pavia Centre Dataset (first row and second row) and for the Salinas Dataset (third row) when using different feature (sub)sets: the color encoding follows the definitions provided in Figs. 3 and 4.



Fig. 6: Relevance of the spectral bands of the Pavia Centre Dataset when using the wrapper-based feature selection method relying on a LDA classifier (blue), the embedded feature selection method relying on the MDPA of a RF classifier (orange), the GRM (green) and the CFS (yellow). For the latter, selected features are assigned a relevance of 1 and all others a relevance of 0.



Fig. 7: Relevance of the spectral bands of the Salinas Dataset when using the wrapper-based feature selection method relying on a LDA classifier (blue), the embedded feature selection method relying on the MDPA of a RF classifier (orange), the GRM (green) and the CFS (yellow). For the latter, selected features are assigned a relevance of 1 and all others a relevance of 0.

Regarding feature relevance assessment, the involved wrapper-based method requires processing times of about 3 min 23 s and 41 min 24 s for the Pavia Centre Dataset and the Salinas Dataset, respectively. The MDPA is derived in 0.46 s and 2.20 s, the GRM in 4.95 s and 14.03 s, and the CFS in 0.69 s and 1.16 s for the Pavia Centre Dataset and the Salinas Dataset, respectively.



Fig. 8: Relevance of the Sentinel-2-like spectral bands derived from the Pavia Centre Dataset (left) and the Salinas Dataset (right) when using the wrapper-based feature selection method relying on a LDA classifier (blue), the embedded feature selection method relying on the MDPA of a RF classifier (orange), the GRM (green) and the CFS (yellow). For the latter, all features are selected and assigned a relevance of 1. The band IDs refer to the bands B₂, B₃, B₄, B₅, B₆ and B₇ for the Pavia Centre Dataset and to the bands B₂, B₃, B₄, B₅, B₆, B₇, B_{8a}, B₁₁ and B₁₂ for the Salinas Dataset.

5 Discussion

The derived classification results reveal that the classification of the Pavia Centre Dataset seems to be less challenging than the classification of the Salinas Dataset when considering the overall accuracy (Tab. 1). Intuitively, this might be due to the fact that there are only 9 classes of interest to be classified for the Pavia Centre Dataset and these classes are likely to be more distinctive, whereas the Salinas Dataset has 16 classes of interest which might have a higher similarity. However, the mean F₁-score across all classes reveals that instances of the different classes are on average better identified for the Salinas Dataset than for the Pavia Centre Dataset (Tab. 1). A look at the confusion matrices and class-wise evaluation metrics derived from these reveals that most of the classes of the Salinas Dataset are very well recognized and only the classes C08 ("Grapes_untrained") and C15 ("Vinyard_untrained") are not as well recognized. As instances of

these two classes occur frequently in the test data, the values for OA are worse. For the Pavia Centre Dataset, most of the classes can be well recognized and only the class C04 ("Self-Blocking Bricks") seems to be challenging due to the different materials which elements of this class might be composed of.

Furthermore, the derived classification results indicate that the different feature (sub)sets provided as input to the RF classifier tend to deliver results of similar quality (Tab. 1 and Fig. 5), even though very different strategies for feature selection are involved. In comparison to the use of all features, however, there is a significant reduction regarding the dimensionality of the considered feature vectors, as only about 20% of the given features are considered for all cases. A deeper analysis of the feature relevance assessment results (Figs. 6 and 7) reveals that the best-ranked and hence selected spectral bands are well-distributed across all available spectral bands. Interestingly, even the transformation of the given hyperspectral data to multispectral Sentinel-2-like data yields classification results of a similar, but slightly worse quality compared to the other approaches. This indicates that Sentinel-2-like data already provide a good source of information for land cover and land use classification, while the use of hyperspectral data seems to only slightly improve the results. Furthermore, the derived results indicate that multispectral Sentinel-2-like data seem to have a low degree of redundancy, so that the CFS method (which aims at reducing redundancy contained in the considered data) does not discard any of the given features (Fig. 8).

6 Conclusions

In this paper, we have comprehensively investigated the potential of both multispectral and hyperspectral data for land cover and land use classification and, in this context, we have also focused on comparatively assessing the relevance of involved spectral bands for the considered classification task. We have presented classification results achieved by using the original hyperspectral data as input to a Random Forest classifier, and we have demonstrated the potential of applying well-established approaches for feature selection: 1) a wrapper-based feature selection method relying on a Linear Discriminant Analysis classifier, 2) an embedded feature selection method relying on an internal metric of a Random Forest classifier, 3) a univariate filter-based feature selection method relying on a general relevance metric and 4) a multivariate filter-based feature selection method also allowing to reduce the degree of redundancy in the considered data. In addition, we have included a transformation of the given hyperspectral data to multispectral Sentinel-2-like data that are relevant for land cover and land use classification. The results derived for two benchmark datasets clearly reveal that Sentinel-2-like data already seem to provide a good source of information for land cover and land use classification, while the use of hyperspectral data (with or without well-established approaches for feature selection) seems to only provide a slight improvement for such a task. This is particularly interesting for large-scale land cover and land use mapping applications, where the use of Sentinel-2 data would correspond to a significantly less expensive data acquisition and a much faster repetition of data acquisition for the same area over time.

7 References

- BRADLEY, P. E., KELLER, S. & WEINMANN, M., 2018: Unsupervised feature selection based on ultrametricity and sparse training data: a case study for the classification of high-dimensional hyperspectral data. Remote Sensing, **10**(10), Article 1564.
- BREIMAN, L., 1996: Bagging predictors. Machine Learning, 24(2), 123-140.
- BREIMAN, L., 2001: Random forests. Machine Learning, 45(1), 5-32.
- CHEHATA, N., LE BRIS, A. & NAJJAR, S., 2014: Contribution of band selection and fusion for hyperspectral classification. Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 1-4.
- CHI, M., FENG, R. & BRUZZONE, L., 2008: Classification of hyperspectral remote-sensing data with primal SVM for small-sized training dataset problem. Advances in Space Research, **41**(11), 1793-1799.
- ELBERTZHAGEN, I., THONFELD, F. & MENZ, G., 2012: SVM-based agricultural land use assessment using Sentinel-2 simulation data. Proceedings of the Sentinel-2 Preparatory Symposium.
- GUYON, I. & ELISSEEFF, A., 2003: An introduction to variable and feature selection. Journal of Machine Learning Research, **3**, 1157-1182.
- HALL, M. A., 1999: Correlation-based feature subset selection for machine learning. PhD thesis, Department of Computer Science, University of Waikato, New Zealand.
- HAM, J., CHENG, Y., CRAWFORD, M. M. & GHOSH, J., 2005: Investigation of the random forest framework for classification of hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing, **43**(3), 492-501.
- HUGHES, G. F., 1968: On the mean accuracy of statistical pattern recognizers. IEEE Transactions on Information Theory, **14**(1), 55-63.
- JOELSSON, S. R., BENEDIKTSSON, J. A. & SVEINSSON, J. R., 2005: Random forest classifiers for hyperspectral data. IEEE Geoscience and Remote Sensing Symposium (IGARSS), 160-163.
- KELLER, S., BRAUN, A. C., HINZ, S. & WEINMANN, M., 2016: Investigation of the impact of dimensionality reduction and feature selection on the classification of hyperspectral EnMAP data. 8th IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 1-5.
- LE BRIS, A., CHEHATA, N., BRIOTTET, X. & PAPARODITIS, N., 2014: Use intermediate results of wrapper band selection methods: a first step toward the optimization of spectral configuration for land cover classifications. Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 1-4.
- LIAW, A. & WIENER, M., 2002: Classification and regression by randomForest. R News, 2/3, 18-22.
- MELGANI, F. & BRUZZONE, L., 2004: Classification of hyperspectral remote sensing images with support vector machines. IEEE Transactions on Geoscience and Remote Sensing, 42(8), 1778-1790.
- SAEYS, Y., INZA, I. & LARRANAGA, P., 2007: A review of feature selection techniques in bioinformatics. Bioinformatics, 23(19), 2507-2517.

- SPOTO, F., SY, O., LABERINTI, P., MARTIMORT, P., FERNANDEZ, V., COLIN, O., HOERSCH, B. & MEYGRET, A., 2012: Overview of Sentinel-2. IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 1707-1710.
- THONFELD, F., FEILHAUER, H. & MENZ, G., 2012: Simulation of Sentinel-2 images from hyperspectral data. Proceedings of the Sentinel-2 Preparatory Symposium.
- 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.
- WEINMANN, M., 2016: Reconstruction and analysis of 3D scenes From irregularly distributed 3D points to object classes. Springer, Cham, Switzerland.
- WEINMANN, M. & WEIDNER, U., 2018: Land-cover and land-use classification based on multitemporal Sentinel-2 data. IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 4946-4949.
- WEINMANN, M., MAIER, P. M., FLORATH, J. & WEIDNER, U., 2018: Investigations on the potential of hyperspectral and Sentinel-2 data for land-cover / land-use classification. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-1, 155-162.
- ZHAO, Z., MORSTATTER, F., SHARMA, S., ALELYANI, S., ANAND, A. & LIU, H., 2010: Advancing Feature Selection Research – ASU Feature Selection Repository. Technical Report, School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ, USA, 2010.

Techniken zur kombinierten Darstellung von 2D-Bodenradar und 3D-Punktwolken zur Analyse des Straßenraums

JOHANNES WOLF¹, SÖREN DISCHER¹ & JÜRGEN DÖLLNER¹

Zusammenfassung:Bei Mobile-Mapping-gestützter Erfassung von Straßenräumen sind sowohl unter- als auch oberirdische Daten von Relevanz. Erstere geben Auskunft über die strukturelle Beschaffenheit der Straßenoberfläche und ermöglichen die frühzeitige Erkennung von Anomalien wie Absackungen oder sich entwickelnde Schlaglöcher. Die oberirdischen Daten stellen Kontextinformationen bereit und erleichtern die visuelle Verortung. Eine kombinierte Darstellung beider Datenquellen ist somit ein hilfreiches Werkzeug bei der Auswertung digitaler Straßenraum-Repräsentationen. Dieser Beitrag präsentiert Techniken zur Interaktion in kombinierten Darstellungen von unterirdischen 2D-Bodenradarinformationen und oberirdischen 3D-Punktwolken. So werden 2D-Bodenradarinformationen auf die Messtrajektorie in zeitgleich erfassten 3D-Punktwolken projiziert, um die Orientierung zu erleichtern. Anomalien werden automatisiert erkannt und visuell hervorgehoben, wodurch die Menge der manuell auszuwertenden Daten signifikant reduziert wird. Evaluationsergebnisse auf Grundlage realweltlicher Datensätze stellen die Nützlichkeit des präsentierten Ansatzes unter Beweis.

1 Einleitung

Techniken des Mobile Mappings finden in unterschiedlichen Gebieten Anwendung (LI 1997) und werden immer häufiger zur Infrastrukturwartung eingesetzt (LI et al. 2017). Kommunen und private Unternehmen erfassen dreidimensionale Geodaten mit Systemen, die unterschiedliche Scanner kombinieren. Dadurch wird eine detaillierte Erfassung und Analyse urbaner Umgebungen, großer Fabrikanlagen und von Infrastrukturnetzwerken ermöglicht (AIRFIELD INVENTORY 2018). Fernerfassungsgeräte wie beispielsweise LiDAR-Scanner können auf beliebige Fahrzeuge montiert werden und ermöglichen eine hochpräzise Aufnahme der Straßenoberfläche, von Bordsteinkanten und von Objekten im Straßenraum (JAAKKOLA et al. 2008). Aus diesen Aufnahmen resultierende 3D-Punktwolken stellen eine detaillierte Repräsentation der erfassten Umgebung dar und werden als "digitale Zwillinge" in einer Reihe unterschiedlicher Anwendungsbereiche eingesetzt (OLSEN & KAYEN 2013; KUMAR et al. 2016). Abhängig von zusätzlich aufgenommenen Daten können 3D-Punktwolken mit RGB-Farben oder Intensitätswerten eingefärbt sein. Dadurch entstehen Datensätze, die auch durch ungeübte Nutzer intuitiv visuell exploriert werden können, wenn sie in GIS-Programmen und speziell für die Darstellung von 3D-Punktwolken entwickelten Programmen angezeigt werden (Abbildung 1).

Bodenradarmessungen (Ground penetrating radar, GPR) werden seit einigen Jahrzehnten für die Untergrundanalyse eingesetzt. GPR-Scanner können Materialeigenschaften mehrere Meter unter der Oberfläche messen und ermöglichen somit Einblicke in die nicht sichtbaren Fundamente von Straßen und Fußwegen (DAVIS & ANNAN 1989). Eine typische Visualisierung eines Bodenradar-

¹ Hasso-Plattner-Institut, Digital Engineering Fakultät, Universität Potsdam, Prof.-Dr.-Helmert-Str. 2-3, D-14482 Potsdam, E-Mail: [johannes.wolf, soeren.discher, juergen.doellner]@hpi.de

B-Scans ist in Abbildung 2 dargestellt. Diese beinhalten eine aufeinanderfolgende Sequenz einzelner Messungen (A-Scans) entlang der Fahrtrichtung. Die X-Achse gibt die gefahrene Strecke an, die Y-Achse zeigt die gemessenen Daten.

Scanner für Bodenradarinformationen können ebenfalls an Fahrzeuge montiert werden und bieten eine Datenquelle für Bereiche, die durch LiDAR nicht erfasst werden können (MOBILE GPR 2017).



Abb. 1: Visualisierung einer 3D-Punktwolke Abb. 2: Visualisierung eines GPR-B-Scans einer Straßenumgebung

Die Kombination von oberirdischen 3D-Punktwolken und unterirdischen 2D-Bodenradarinformationen ermöglicht eine umfassendere Analyse von Straßenumgebungen im Vergleich zur individuellen Auswertung beider Datenquellen. Ein häufiger Anwendungsfall für die Auswertung von Bodenradarinformationen ist das frühzeitige Erkennen möglicher Schlaglöcher (HUSTON et al. 2000). Durch die Betrachtung zugehöriger Informationen zur Straßenoberfläche aus der 3D-Punktwolke können in diesem Kontext beispielsweise Kanaldeckel sofort erkannt werden und müssen nicht mehr als mögliche Anomalien ausgewertet werden.

2 Verwandte Arbeiten

Benedetto et al. geben in ihrer Arbeit einen Überblick, wie Bodenradarinformationen zur Analyse des Straßenraums eingesetzt werden können (BENEDETTO et al. 2017). Sie diskutieren detailliert, welche Verarbeitungstechniken verwendet werden können, um die Beschaffenheit der Straßenoberfläche zu ermitteln. Evans et al. fassen Einsatzmöglichkeiten des Bodenradars zur allgemeinen Analyse von Straßen zusammen (EvANS et al. 2008). Saarenketo und Scullion listen die Erkennung von Schlaglochpositionen und Überwachung von Rissbildungen in ihrem Bericht zu Anwendungen von Bodenradar auf Straßen und Autobahnen auf (SAARENKETO & SCULLION 1994). Sie beschreiben weiterhin mögliche Analysen des Bodens und der Straßenschichten sowie die dafür benötigten Techniken zur Dateninterpretation (SAARENKETO & SCULLION 2000).

Giannopoulos beschreibt, wie Bodenradarinformationen visualisiert werden können (GIANNOPOU-LOS 2005). Ergänzend zu zweidimensionalen Profilen stellt er ein Beispiel für eine dreidimensionale Darstellung vor, in der die Daten in einen Quader eingefasst werden, für den die drei Hauptebenen (xy-, xz- und yz-Ebene) gerendert werden. Üblicherweise werden zweidimensionale GPR- Profile einzeln dargestellt, mehrere Scans können aber nebeneinander platziert werden, um ein räumliches Gefühl für die Daten zu vermitteln (GeoRadar: 3D and GPR 2005).

Puente et al. vergleichen unterschiedliche mobile Systeme für terrestrisches Laserscanning. Sie zeigen, dass LiDAR weit verbreitet ist und auf unterschiedlichsten Trägersystemen zur Datenerfassung eingesetzt wird. Nutzer können durch das Anzeigen und Bearbeiten dichter Punktwolken realweltliche Bedingungen einfach nachvollziehen (PUENTE et al. 2013). Discher et al. und Eitel et al. beschreiben die Relevanz von 3D-Punktwolken für viele Anwendungen im Geodatenbereich (DISCHER et al. 2019; EITEL et al. 2016). Biasion et al. beschreiben Anwendungen mobiler Laserscanner zur Umgebungsanalyse für Katastrophenmanagment (BIASION et al. 2005). Mehrere Autoren diskutieren die automatisierte Analyse von 3D-Punktwolken: Oberflächenkategorien pro Punkt können aus der Topologie einer 3D-Punktwolke abgeleitet (CHEN et al. 2017) oder durch Ansätze des Deep Learnings ermittelt werden (BOULCH et al. 2017). Ermittelte Oberflächenkategorien können wiederum genutzt werden, um dreidimensionale Modelle von Gebäuden oder Infrastrukturobjekten zu rekonstruieren (TEIZER et al. 2005).

Einen allgemeinen Überblick zu Renderingtechniken für 3D-Punktwolken geben Gross und Pfister (GROSS & PFISTER 2011). Während *fotorealistische* Ansätze (SCHÜTZ & WIMMER 2015; PREINER et al. 2012) darauf abzielen, typische Artefakte (z.B. visuelle Überlagerungen oder lückenhafte Oberflächendarstellungen) durch das Rendering der Punkte mit geeigneter Größe und Ausrichtung zu reduzieren, konzentrieren sich *nicht-fotorealistische* Ansätze (SIMONS et al. 2014; ZHANG et al. 2014) auf die gezielte Hervorhebung von Kanten und Strukturen innerhalb der erfassten Daten. All diese Renderingtechniken können in die Visualisierung integriert werden, wie am Beispiel des *Eye Dome Lighting* von Boucheny (BOUCHENY 2009) demonstriert. Fokus-und-Kontext-Visualisierungen wurden im Bereich von Mesh-basierten Modellen diskutiert (VAARANIEMI et al. 2013; ELMQVIST & TSIGSAS 2008) – von statischen Sichtbarkeitsmasken (SIGG et al. 2012) bis zu interaktiven Linsen (TRAPP et al. 2008). Discher et al. wenden solche Techniken auf 3D-Punktwolken an, sodass Nutzer für ihre Arbeit relevante, aber normalerweise verdeckte, Bereiche hervorheben können (DISCHER et al. 2017). Ähnlich zum hier gezeigten Ansatz verwenden sie ein Multi-Pass-Rendering basierend auf G-Buffern, allerdings kombinieren Sie 3D-Punktwolken nicht mit weiteren Geodaten.

Um 3D-Punktwolken beliebiger Größe rendern zu können, trennen aktuelle Out-of-core-Ansätze das Rendering vom Datenmanagement. Beispielhafte Systeme (MARTINEZ-RUBI et al. 2016; GOS-WANI et al. 2013; RUSINKIEWICZ & LEVOY 2000) unterteilen 3D-Punktwolken in kleinere, repräsentative Bereiche, die in Echtzeit gerendert werden können.

3 Datencharakteristik und Anforderungen

Aktuelle Technik aus dem Bereich der Fernerkundung ermöglicht es, digitale Zwillinge der realen Welt mit sehr geringem zeitlichem und finanziellem Aufwand zu erfassen. Technische Grundlage sind (1) aktive Sensoren (z.B. LiDAR), bei der elektromagnetische Strahlung ausgesendet wird, um Entfernungen zwischen Sensor und Oberfläche zu messen (EITEL et al. 2016) und (2) passive Sensoren (z.B. Digitalkameras), die natürliche Strahlung nutzen, um aus einer fortlaufenden Reihe von Bildern mit Hilfe von *Dense Image Matching* 3D-Informationen abzuleiten (REMONDINO et al. 2013).

Mit passiven Sensoren können Rohdaten schneller erfasst werden, andererseits erfordern sie eine rechenintensive Nachbearbeitung, während aktive Sensoren 3D-Punktwolken direkt zum Zeitpunkt der Aufnahme generieren. Mit beiden Ansätzen können detaillierte Datensätze erzeugt werden, die hohe Punktdichten liefern. 3D-Punktwolken können für ganze Gebäudekomplexe, Städte oder Länder in großem Maßstab effizient generiert werden, indem Sensoren an ausgewählten Standorten platziert (lokale terrestrische Erfassung), auf mobile Trägerplattformen wie Autos oder Züge montiert (Mobile Mapping) oder an UAVs und Flugzeugen (luftgestützte Erfassung) befestigt werden. So können Datensätze mit Milliarden von Punkten und hunderten Gigabyte an Daten entstehen.

Bodenradar wird für die Analyse unterirdischer Bereiche genutzt. Dabei werden elektromagnetische Wellen in den Untergrund abgestrahlt, deren Reflektionen an unterschiedlichen Schichten der Straße und des Bodens Aufschlüsse über deren Beschaffenheit geben. Radarantennen zur Straßenanalyse werden üblicherweise an Messfahrzeugen angebracht, die im fließenden Verkehr mitfahren können. Die Antennen sind ein Stück über dem Boden montiert und während der Fahrt kann der Abstand zur Fahrbahn leicht schwanken. Um dies zu berücksichtigen, wird üblicherweise eine Time-Zero-Korrektur zur Sensorkalibrierung durchgeführt. Nach der Erfassung werden die Daten zumeist in der Form von B-Scans ausgewertet. In der in Abbildung 2 gezeigten Zweifarbendarstellung in rot und blau wird die Ausrichtung der Amplitude des empfangenen Signals durch die Farbe dargestellt und deren Größe durch die Intensität. Mobile-Mapping-Fahrzeuge nutzen das *Global Positioning System (GPS)* zur Positionserfassung. Dessen Messdaten werden in Paaren von Zeitstempeln und Positionen vorgehalten, die zusammen die Trajektorie des zurückgelegten Weges angeben. Unter Nutzung dieser Angaben und dem Wissen über die fixen Positionen der Li-DAR-Scanner und Bodenradarantennen am Fahrzeug können die erfassten Daten aller Quellen präzise im gleichen räumlichen Koordinatensystem platziert werden.

Unter Berücksichtigung der Datencharakteristiken der oben erwähnten Datenquellen wurden folgende Anforderungen für eine kombinierte Darstellung von 2D-Bodenradar und 3D-Punktwolken abgeleitet:

- A1: Keine Beschränkung auf Aufnahmemethoden, Anzahl, räumliche Ausdehnung oder Größe der Datensätze. Letzteres gilt insbesondere für 3D-Punktwolken, weil diese häufig hunderte Gigabyte Rohdaten umfassen können.
- A2: Korrekte Positionierung von Bodenradardaten und 3D-Punktwolken in einem homogenen dreidimensionalen Referenzsystem.
- A3: Verdeckungsfreie Darstellung individueller B-Scans eines Bodenradar-Datensatzes.
- A4: Visuelles Filtern und Hervorheben bestimmter Gebiete in den Daten, die von besonderem Interesse für die Auswertung sind.

4 Systemüberblick

Die oben genannten Anforderungen sind in den Entwurf und die Implementierung der Anwendung eingeflossen. Das System besteht aus folgenden Hauptbestandteilen:

4.1 Punktwolkenmanager

Alle 3D-Punktwolken werden in einem einzelnen, homogenen räumlichen Datenmodell gesammelt. Auf dieses Modell wird über den Punktwolkenmanager zugegriffen, der die räumlichen Informationen mit allen verfügbaren Pro-Punkt-Attributen (z.B. Farbwerte) vorhält. Um A1 gerecht zu werden, müssen die Daten in räumlichen Datenstrukturen vorgehalten werden, die Level-ofdetail-Repräsentationen (LOD-Repräsentationen) ermöglichen, wie zum Beispiel Quadtrees (GAO et al. 2014), Octrees (ELSEBERG et al. 2013) oder kd-Trees (RICHTER et al. 2015). Diese LOD-Repräsentationen werden durch die Prozessierungs-Engine generiert.

4.2 Bodenradarmanager

In ähnlicher Weise werden Bodenradarinformationen effizient durch den Bodenradarmanager zur Verfügung gestellt. Daten der B-Scans werden zusammen mit den Daten der Trajektorie verarbeitet und können dadurch im gleichen dreidimensionalen System wie die Punktwolke platziert werden (A2). B-Scans können individuell angefragt werden (A3). Wegen des deutlich geringeren Datenaufkommens werden für das Bodenradar keine LOD-Repräsentationen generiert. Die in den Testdaten verarbeiteten Datensätze umfassten beispielsweise nur höchstens 25,7 MB Rohdaten pro B-Scan.

4.3 Prozessierungs-Engine

Mithilfe der Prozessierungs-Engine können die erfassten Daten vorverarbeitet werden. So wird die Georeferenzierung der Bodenradardaten anhand der Trajektorie vorgenommen, die 3D-Punktwolken werden um Ausreißer bereinigt und es werden LOD-Repräsentationen für sie erzeugt. Diese Verarbeitung kann parallel für mehrere Datensätze ausgeführt werden und nutzt die jeweils vorhandenen Systemressourcen optimal aus. Die verarbeiteten Daten stehen anschließend über den Punktwolkenmanager und Bodenradarmanager zur Verfügung.

4.4 Visualisierungs-Engine

Die Visualisierungs-Engine dient der Bereitstellung einer interaktiven, kombinierten Darstellung der vorhandenen Datensätze. Dabei findet ein Multi-Pass-Rendering mit mehreren G-Buffern (SAITO & TAKAHASHI 1990) Anwendung:

Für jeden Datentyp wird der zugehörige Manager angefragt, der nur die Ausschnitte der Daten zurückgibt, die für die aktuelle Ansicht relevant sind. So wird beispielsweise View-Frustum-Culling und Detail-Culling eingesetzt, um die Menge der anzuzeigenden Punkte an die vorhandenen CPU- und GPU-Kapazitäten anzupassen. Die geladenen Daten werden in separate G-Buffer gerendert, spezialisierte Frame-Buffer-Objekte mit kombinierten 2D-Texturen für Farbe, Tiefeninformation oder Normalenvektorausrichtung. Die G-Buffer werden anschließend zusammengesetzt und ermöglichen es, bildbasierte Nachverarbeitungsschritte durchzuführen, die der visuellen Filterung und Hervorhebung dienen, wie zum Beispiel beim Hervorheben von Tiefenunterschieden in 3D-Punktwolken mittels Eye Dome Lighting (BOUCHENY et al. 2009) oder interaktiven Linsen. Die Nachverarbeitung kann zur Laufzeit über die Nutzeroberfläche konfiguriert werden.

4.5 Nutzeroberfläche

Basierend auf den Aktionen der Nutzer werden auf der Nutzeroberfläche die angefragten Daten dargestellt. Insbesondere können Nutzer (1) Position und Betrachtungswinkel ändern, (2) auswählen welche B-Scans angezeigt werden sollen, (3) diese B-Scans horizontal und vertikal anschneiden, (4) Nachverarbeitungsschritte für die Darstellung auswählen und konfigurieren sowie (5) Bereiche von besonderem Interesse hervorheben (A4).

5 Visualisierungstechniken

Die kombinierte Darstellung von 2D-Bodenradar und 3D-Punktwolken basiert auf zwei Hauptkomponenten in der Nutzeroberfläche: Einer 3D-Szenendarstellung und einem 2D-Übersichtsfenster.

5.1 3D-Szenendarstellung

Zunächst werden die B-Scans des Bodenradars mittels der Trajektorie und der Information über die genaue Lage der Antennen in den dreidimensionalen Raum platziert (A2). Nutzer können sich gegenseitig verdeckende B-Scans bei Bedarf gezielt ausblenden (A3). Ein kompletter Bodenradar-Datensatz wird durch eine quaderförmige Struktur repräsentiert, die den gesamten vom Bodenradar erfassten Bereich überdeckt (Abbildung 3). Der Raum zwischen den B-Scans wird dabei durch interpolierte Werte aufgefüllt. Dadurch kann der dargestellte Ausschnitt je nach Bedarf vertikal und horizontal angeschnitten werden, um Bereiche von besonderem Interesse hervorzuheben (A4).



Abb. 3: Auf die GPS-Trajektorie platzierter Quader (links). Vertikal und horizontal angeschnittener Quader (rechts).

Die Sichtbarkeit der Bodenradarinformationen wird verbessert, indem der Quader nicht unter der Straßenoberfläche dargestellt, sondern über diese gehoben wird (A3). Um den räumlichen Kontext zur 3D-Punktwolke beizubehalten, werden die über dem Quader befindlichen Punkte gleichermaßen angehoben. Zusätzlich werden sie eingefärbt, um den Unterschied zu den anderen Punkten deutlich zu machen (Abbildung 4). Eine interaktive Linse ermöglicht es, Punkte über dem Quader in der Umgebung des Mauszeigers auszublenden, um einen direkten Blick auf die Daten des Bodenradars zu erhalten. Alternativ kann der Modus auch umgekehrt eingestellt werden, sodass über dem Quader befindliche Punkte ausschließlich in der Umgebung des Mauszeigers dargestellt werden, (A4).



Abb. 4: Über den Quader gehobene Punkte sind hervorgehoben. Eine interaktive Linse ermöglicht einen direkten Blick auf die Daten.

5.2 2D-Übersichtsfenster

In diesem Fenster wird ein Gesamtüberblick zu den Bodenradarinformationen eines Datensatzes gegeben (Abbildung 5). Mit Hilfe des Übersichtsfensters kann die Darstellung des Quaders in der 3D-Szenendarstellung angepasst werden. Dies umfasst (1) das Einstellen der Höhe über der Fahrbahnoberfläche, (2) das Zuschneiden in horizontaler und vertikaler Richtung sowie (3) das Einund Ausblenden spezifischer B-Scans. Beim Wählen einer bestimmten Breite für einen horizontalen Ausschnitt kann dieses "Sichtfenster" entlang der Trajektorie verschoben werden. Dadurch wird gleichzeitig auch die Kameraposition innerhalb der 3D-Szene verschoben, sodass immer die jeweilige Umgebung in der 3D-Punktwolke sichtbar und auf den Quader zentriert ist (A4).



Abb. 5: 2D-Übersichtsfenster mit Bodenradarinformationen und Optionen zum Anschneiden und Hervorheben.

Nutzer können weiterhin einstellen, wie Texturen aus den Bodenradardaten generiert werden. Die Eingabedaten enthalten Rohinformationen zum reflektierten Signal, die generierten Texturen zeigen die Amplituden dieser Messwerte in rot und blau für die positiven und negativen Werte. Da die Werte stark schwanken, können Nutzer einstellen, wie sie skaliert werden sollen. Beispielsweise ist eine hohe Skalierung sinnvoll für Datenbereiche, in denen es nur geringe Unterschiede in den Messwerten gibt. Zusätzlich können Bereiche mit großen Unterschieden zu den jeweiligen umgebenden Werten durch das Einstellen von Schwellwerten hervorgehoben werden, wodurch Anomalien in den Bodenradarinformationen interaktiv lokalisiert werden können.

6 Anomalie-Erkennung

Bei der großflächigen Aufnahme von Bodenradardaten fallen große Datenmengen an, deren manuelle Auswertung erheblichen Zeitaufwand verursacht. Große Teile der Daten enthalten aber keine Informationen, die im weiteren Verlauf der Auswertung relevant sind. Insbesondere sind über lange Strecken gleichmäßig verlaufende Straßenschichten ein Zeichen für einen guten Straßenzustand, bei dem kein Anlass zur detaillierten Begutachtung besteht. Hingegen sind alle vom Normalzustand abweichenden Bereiche von Interesse. Zur Unterstützung bei der Auswertung der Daten werden die Bodenradarinformationen automatisiert auf Anomalien hin untersucht, um diese für die Nutzer hervorzuheben. Bei der manuellen Analyse müssen anschließend nur noch diejenigen Bereiche berücksichtigt werden, in denen Abweichungen existieren.



Abb. 6: Hervorhebung der Position gefundener Anomalien in der 3D-Szene.

Dazu werden etablierte Machine-Learning-Verfahren aus dem Bereich der Bildanalyse eingesetzt. Nach der Umwandlung der Rohdaten des Bodenradars in Texturen können diese als Bilder interpretiert werden. Die oben beschriebene Darstellung der Messwerte mit zwei Farben erzeugt Bilder, in denen der Linienverlauf interpretiert werden kann. In der prototypischen Implementierung wurde ein Netzwerk zur Handschrifterkennung verwendet (PEDREGOSA et al., 2011), das sich auch für diesen Anwendungszweck als geeignet herausgestellt hat. In einem Testdatensatz wurden Bereiche markiert, in denen Anomalien in den Daten enthalten sind. Das mit diesen Daten trainierte Netzwerk kann anschließend auf unbekannten Datensätzen angewendet werden, um Bereiche mit Anomalien zu erkennen. Diese werden in der Anwendung sowohl in der 3D-Szene als auch im Übersichtsfenster hervorgehoben (Abbildung 6). Nutzer können durch diese Darstellung schnell zu den für sie interessanten Gebieten innerhalb des Datensatzes springen.

7 Evaluation

Das hier präsentierte System wurde in C++ entwickelt und erweitert das von Discher et al. vorgestellte System (DISCHER et al. 2017). Die Testdaten umfassen vier parallel erfasste B-Scans des Bodenradars, die kontinuierlich in Fahrtrichtung erfasst wurden (Abbildung 7) und 3D-Punktwolkendaten aus einer Mobile-Mapping-Erfassung mit LiDAR. Zwei Antennen des Bodenradars haben mit einer Frequenz von 2.000 MHz gemessen, die anderen beiden mit 1.000 MHz. Die Signale reichen bis zu 0,45 m bzw. 0,90 m in den Boden. Die Antennen wurden im Abstand von jeweils 0,42m zueinander positioniert. Jeder B-Scan enthält 650 m Straßendaten mit 13.146 Datenpunkten, die jeweils 512 mal 4 Byte umfassen.



Abb. 7: An das Messfahrzeug montierte Bodenradarantennen.

Die 3D-Punktwolke ist eine Kombination aus zwei LiDAR-Scans, die nach links und rechts vorne oberhalb des Fahrzeugs aufgenommen wurden. Neben den dreidimensionalen Koordinaten hat jeder der 58,7 Millionen Messpunkte Informationen über die Intensität des reflektierten Strahles, die als Grauwerte dargestellt werden. In dem vom Bodenradar abgedeckten Bereich wurden mehr als 1.000 Punkte pro Quadratmeter erfasst.

7.1 Nutzbarkeit

Die Anordnung der Antennen sorgt für eine Hürde in der Darstellung der Daten: Die erste und dritte Antenne messen mit einer Frequenz, die zweite und vierte mit einer anderen. Diese Anordnung ist hilfreich, um den annähernd gleichen Bereich mit zwei unterschiedlichen Frequenzen und damit unterschiedlichen Tiefen und Genauigkeiten zu messen. Allerdings sind benachbarte B-Scans dadurch nicht mehr direkt vergleichbar. Eine Interpolation der Zwischenwerte kann aus diesem Grund zu unerwarteten Ergebnissen führen, die individuelle Auswahl der gewünschten B-Scans hilft dabei, das Problem zu umgehen. Weil die B-Scans dicht nebeneinanderliegen, verdecken sie sich gegenseitig. Da die exakte räumliche Platzierung innerhalb der 3D-Punktwolke essentieller Bestandteil der Darstellung ist, kann der Abstand untereinander nicht einfach vergrößert werden. Die Sichtbarkeit innerer B-Scans kann wiederum durch die individuelle Auswahl der anzuzeigenden B-Scans hergestellt werden.

7.2 Performanz

Wie von Discher et al. gezeigt, ist das verwendete System in der Lage, interaktive Frameraten für beliebig große 3D-Punktwolken zu generieren (DISCHER et al. 2017). Die Bodenradarinformationen werden als Rohdaten von dem hier vorgestellten System eingelesen und in Texturen umgewandelt, die in der Anwendung angezeigt werden können. Die Performanz des Systems beruht deshalb auf der Dauer des initialen Ladevorgangs sowie – während der Laufzeit – auf der benötigten Zeit zur Aktualisierung und zum Rendern der B-Scan-Texturen. In den Testdaten haben die Texturen eine Auflösung von 13.146 x 512 Pixeln mit einer Tiefe von 4 Byte für die Rohdaten und einem Byte in den resultierenden Texturen. Sowohl die Aktualisierung als auch das Rendern der Texturen ist vollständig GPU-beschleunigt und damit im Vergleich zur Verarbeitung der 3D-Punktwolke vernachlässigbar. Für größere Datensätze könnten Speicherverwaltung und Level-of-Detail-Verfahren eine größere Rolle spielen, um der Begrenzung des verfügbaren GPU-Speichers zu begegnen.

8 Fazit und Ausblick

Eine kombinierte Darstellung von 2D-Bodenradar und 3D-Punktwolken ermöglicht Vergleiche zwischen beiden Datensätzen und somit eine bessere Datenauswertung, zum Beispiel im Kontext der Analyse des Straßenzustandes. Anomalien in den Bodenradarinformationen können mit der 3D-Punktwolke abgeglichen werden, um auch an der Straßenoberfläche sichtbare Abweichungen zu erkennen. Kanaldeckel können in der 3D-Punktwolke sowohl bei manueller Auswertung als auch durch automatisierte Verfahren einfach identifiziert werden und deren Position kann bei der Anomalie-Erkennung in den Bodenradarinformationen berücksichtigt werden. Durch das Zuschneiden der B-Scans auf bestimmte Bereiche können sich Nutzer auf relevante Details fokussieren und störende verdeckende Informationen ausblenden. Einzelne B-Scans können komplett ausgeblendet werden, um die visuelle Komplexität weiter zu reduzieren, insbesondere im Bezug auf die unterschiedlichen Messfrequenzen. Ferner könnte die Implementierung von Level-of-Detail-Verfahren für größere Datensätze hilfreich sein.

Die Darstellung mehrerer Bodenradar-Scans im Bereich von Kreuzungen und bei parallel befahrenen Spuren bieten Potenzial für weitere Entwicklungen bezüglich der Auflösung sich in diesen Bereichen überlagernder Daten.

Die über das Bodenniveau gehobene Darstellung von B-Scans und gleichzeitiges Anheben der darüberliegenden Punkte aus der 3D-Punktwolke ermöglichen eine fokussierte Sicht auf die 2D-Bodenradarinformationen. Das Ausblenden von Punkten im Bereich einer durch den Mauszeiger gesteuerten Linse erlaubt die Inspektion der direkt darunterliegenden Daten im jeweiligen Kontext. Das Einstellen von Schwellwerten und Skalierungsfaktoren für die Messwerte des Bodenradars unterstützt die Auswertung in Bezug auf Anomalien. Die automatisierte Hervorhebung potenziell relevanter Bereiche beschleunigt das Auswerten der Daten, weil nur noch Teilmengen der aufgenommenen Information betrachtet werden müssen.

Anomalien in den Bodenradarinformationen an Positionen, die in der 3D-Punktwolke als Kanaldeckel identifiziert wurden, können entsprechend gefiltert werden. Dadurch kann die Menge der manuell auszuwertenden Bereiche weiter eingeschränkt werden.

9 Danksagung

Wir danken dem Amt für Geoinformation, Vermessung und Kataster der Stadt Essen für die Bereitstellung von Daten, die im Rahmen der hier vorgestellten Arbeit verwendet wurden.

10 Literaturverzeichnis

- AIRFIELD INVENTORY, 2018. https://www.xyht.com/lidarimaging/huge-airport-one-vehicle/, letz-ter Zugriff 27.10.2018.
- BENEDETTO, A., TOSTI, F., CIAMPOLI, L. B. & D'AMICO, F., 2017: An overview of ground-penetrating radar signal processing techniques for road inspections. Signal Processing, 132, 201-209.
- BIASION, A., BORNAZ, L. & RINAUDO, F., 2005: Laser scanning applications on disaster management. Geo-information for Disaster Management, Springer, Berlin/Heidelberg, 19-33.
- BOUCHENY, C., 2009: Interactive Scientific Visualization of Large Datasets: Towards a Perceptive-Based Approach. Doktorarbeit, Université Joseph Fourier, Grenoble.
- BOULCH, A., SAUX, B. L. & AUDEBERT, N., 2017: Unstructured point cloud semantic labeling using deep segmentation networks. Eurographics Workshop on 3D Object Retrieval, 2017.
- CHEN, D., WANG, R. & PEETHAMBARAN, J., 2017: Topologically aware building rooftop reconstruction from airborne laser scanning point clouds. IEEE TGRS, **55**(12), 7032-7052.
- DAVIS, J. L. & ANNAN, A. P., 1989: Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy. Geophysical prospecting, **37**(5), 531-551.
- DISCHER, S., RICHTER, R. & DÖLLNER, J., 2018: A scalable webGL-based approach for visualizing massive 3d point clouds using semantics-dependent rendering techniques. Proceedings of Web3D, 19:1-19:9.
- DISCHER, S., RICHTER, R. & DÖLLNER, J., 2017: Interactive and view-dependent see-through lenses for massive 3d point clouds. Advances in 3D Geoinformation, 49-62.
- DISCHER, S., RICHTER, R., TRAPP, M. & DÖLLNER, J., 2019: Service-oriented processing and analysis of massive point clouds in geoinformation management. Service-Oriented Mapping, Springer, Berlin/Heidelberg, 43-61.
- EITEL, J. U., HÖFLE, B., VIERLING, L. A., ABELLÁN, A., ASNER, G. P., DEEMS, J. S., GLENNIE, C. L., JOERG, P. C., LEWINTER, A. L., MAGNEY, T. S., MANDLBURGER G., MORTON, D. C., MÜLLER, J. & VIERLING, K. T., 2016: Beyond 3-D: The new spectrum of lidar applications for earth and ecological sciences. Remote Sensing of Environment, **186**, 372-392.
- ELMQVIST, N. & TSIGAS, P., 2008: A taxonomy of 3d occlusion management for visualization. IEEE TVCG, 14(5), 1095-1109.
- ELSEBERG, J., BORRMANN, D. & NÜCHTER, A., 2013: One billion points in the cloud an octree for efficient processing of 3D laser scans. ISPRS Journal of Photogrammetry and Remote Sensing, **76**, 76-88.
- EVANS, R. D., FROST, M. W., STONECLIFFE-JONES, M. & DIXON, N., 2008: A review of pavement assessment using ground penetrating radar (GPR). 12th International Conference on Ground Penetrating Radar.
- GAO, Z., NOCERA, L., WANG, M. & NEUMANN, U., 2014: Visualizing aerial lidar cities with hierarchical hybrid point-polygon structures. Graphics Interface, 137-144.
- Geo Radar: 3D and GPR, 2005. http://www.geo-radar.pl/en/methods/georadar/3d/, letzter Zugriff 25.10.2018.
- GIANNOPOULOS, A., 2005: Modelling ground penetrating radar by gprmax. Construction and building materials, **19**(10), 755-762.

GROSS, M. & PFISTER, H., 2011: Point-based graphics. Morgan Kaufmann, Burlington.

- HUSTON, D. R., PELCZARSKI, N. V., ESSER, B. & MASER, K. R., 2000: Damage detection in roadways with ground penetrating radar. Eighth International Conference on Ground Penetrating Radar, **4084**, 91-95.
- JAAKKOLA, A., HYYPPÄ, J., HYYPPÄ, H. & KUKKO, A., 2008: Retrieval algorithms for road surface modelling using laser-based mobile mapping. Sensors, **8**(9), 5238-5249.
- KUMAR, S., DESHPANDE, A., HO, S. S., KU, J. S. & SARMA, S. E., 2016: Urban street lighting infrastructure monitoring using a mobile sensor platform. IEEE Sensors Journal, 16(12), 4981-4994.
- LI, F., ELBERINK, S. O. & VOSSELMAN, G., 2017: Semantic labelling of road furniture in mobile laser scanning data. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, **42**(2/W7), 247-254.
- LI, R., 1997: Mobile mapping: An emerging technology for spatial data acquisition. Photogrammetric Engineering and Remote Sensing, **63**(9), 1085-1092.
- MARTINEZ-RUBI, O., DE KLEIJN, M., VERHOEVEN, S., DROST, N., ATTEMA, J., VAN MEERSBERGEN, M., VAN NIEUWPOORT, R., DE HOND, R., DIAS, E. & SVETACHOV, P., 2016: Using modular 3D digital earth applications based on point clouds for the study of complex sites. International Journal of Digital Earth, 9(12), 1135-1152.
- Mobile GPR, 2017: https://www.catsurveys.com/Services/MGPR, letzter Zugriff 26.06.2018.
- OLSEN, M. J. & KAYEN, R., 2013: Post-earthquake and tsunami 3D laser scanning forensic investigations. Forensic Engineering 2012: Gateway to a Safer Tomorrow, 477-486.
- PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A., COURNA-PEAU, D., BRUCHER, M., PERROT, M. & DUCHESNAY, E., 2011: Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12,2825-2830.
- PREINER, R., JESCHKE, S. & WIMMER, M., 2012: Auto splats: Dynamic point cloud visualization on the GPU. EGPGV, 139-148.
- PUENTE, I., GONZÁLEZ-JORGE, H., MARTÍNEZ-SÁNCHEZ, J. & ARIAS, P., 2013: Review of mobile mapping and surveying technologies. Measurement, **46**(7), 2127-2145.
- REMONDINO, F., SPERA, M. G., NOCERINO, E., MENNA, F., NEX, F. & GONIZZI-BARSANTI, S., 2013: Dense Image Matching: Comparisons and Analyses. Digital Heritage International Congress, 47-54.
- RICHTER, R., DISCHER, S. & DÖLLNER, J., 2015: Out-of-core visualization of classified 3D point clouds. 3D Geoinformation Science, Springer, Berlin/Heidelberg, 227-242.
- RUSINKIEWICZ, S. & LEVOY, M., 2000: Qsplat: A multiresolution point rendering system for large meshes. SIGGRAPH, 343-352.
- SAARENKETO, T. & SCULLION, T., 1994: Ground penetrating radar applications on roads and highways. Research Report 1923-2F, Texas Transportation Institute, College Station, Texas.
- SAARENKETO, T. & SCULLION, T., 2000: Road evaluation with ground penetrating radar. Journal of applied geophysics, **43**(2-4), 119-138.
- SAITO, T. & TAKAHASHI, T., 1990: Comprehensible rendering of 3-D shapes. ACM SIGGRAPH Computer Graphics, 24, 197-206.

- SCHÜTZ, M. & WIMMER, M., 2015: High-quality point-based rendering using fast single-pass interpolation. Digital Heritage, 369-372.
- SCHÜTZ, M. & WIMMER, M., 2015: Rendering large point clouds in web browsers. Proceedings of CESCG, 83-90.
- SIGG, S., FUCHS, R., CARNECKY, R. & PEIKERT, R., 2012: Intelligent cutaway illustrations. Proceedings of PacificVis, 185-192.
- SIMONS, L., HE, S., TITTMAN, P. & AMENTA, N., 2014: Pointbased rendering of forest lidar. Proceedings of EnvirVis, 19-23.
- TEIZER, J., KIM, C., HAAS, C., LIAPI, K. & CALDAS, C., 2005: Framework for real-time three-dimensional modeling of infrastructure. TRR Journal, **1913**, 177-186.
- TRAPP, M., GLANDER, T., BUCHHOLZ, H. & DÖLLNER, J., 2008: 3D generalization lenses for interactive focus + context visualization of virtual city models. International Conference Information Visualisation, 356-361.
- VAARANIEMI, M., FREIDANK, M. & WESTERMANN, R., 2013: Enhancing the visibility of labels in 3D navigation maps. Progress and new trends in 3D geoinformation sciences, 23-40.
- ZHANG, L., SUN, Q. & HE, Y., 2014: Splatting lines: An efficient method for illustrating 3D surfaces and volumes. 18th Meeting of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games, 135-142.

UAV Remote Sensing Data Handling: A Transition from Testing to Long-Term Data Acquisition for Intensive Forest Monitoring

STUART KRAUSE^{1,2,3}, MAXIMILIAN STRER¹, JAN-PETER MUND² & TANJA GM SANDERS¹

Abstract: Unmanned Aerial Vehicles (UAVs) and their application in intensive forest monitoring is opening up new possibilities for forest research. Recent trials have shown promising results for the extraction of tree parameters and the quantification of phenological observations. Long-term data acquisition strategies require the handling of large amounts of data as well as increased computing performance requirements. This study explores the challenges in transitioning from the testing phase of UAV-based remote sensing technology to a long-term data acquisition scheme capable of handling Big Data by means of robust archiving and processing workflows as well as the incorporation of cloud computing.

1 Introduction

For the past two years, the use of Unmanned Aerial Vehicles (UAVs) as a nearfield remote sensing platform were tested at the Britz forest research station in Brandenburg, Germany. The possibility to acquire high temporal, spatial and spectral resolution datasets by means of self-determined guidelines rather than adapting to secondary sources is of interest (TORRESAN et al. 2016). During this testing phase, selected datasets were acquired, processed and analyzed on local computers and in-house servers. The next stage in the progression of this rapidly developing technology is now however, the establishment of long-term data collection schemes which require the capacity to handle the storage and processing of "Big Data" (LI et al. 2016).

Recent studies at the Britz research station showed that UAV mounted multispectral sensors can obtain accurate geometric features such as tree heights and crown diameter and promising results have been demonstrated in the quantification of phenological phases. Remote sensing datasets acquired in the field for experimental studies were successfully processed and analyzed on localized Linux and Windows workstations. The transition to a more permanent and long-term data collection and processing scheme does however pose the challenge in terms of data storage and processing capabilities.

The concept of Big Data has been in discussion for a number of years now and it is estimated that the volume of global scientific data doubles every two to three years (Guo et al. 2014). Due to an increase in the interest of remote sensing data for environmental research, the issue of acquiring, storing, analysing, visualizing and sharing such data is becoming a challenge (LIU 2015) and solutions are required.

¹ Thünen Institute of Forest Ecosystems, Alfred-Möller-Str. 1, Haus 41/42, D-16225 Eberswalde, E-Mail: [stuart.krause, maximilian.strer, tanja.sanders]@thuenen.de

² Eberswalde University for Sustainable Development, Alfred-Möller-Str. 1, D-16225 Eberswalde, E-Mail: Jan-Peter.Mund@hnee.de

³ University of Bonn, Meckenheimer Allee 166, 53115 Bonn, Germany

Big Data is not only attributable to the volume and velocity of data as in storage and computing capabilities, however also the variety and complexity of remote sensing data (MA et al. 2015). In terms of long-term UAV image data collection by a research institution, the amount of data would not necessarily be deemed as Big Data especially when compared to a typical satellite ground station which could be receiving up to a terabyte of data per day. Relatively speaking however, a low altitude (flying height < 100 m) UAV employed by a forest research institution during phenological phases can acquire up to several terabytes of raw data per year. The storage of such data as well as photogrammetric products could potentially carry on for decades. Furthermore, the processing, analysis and visualization of such data requires costly high-performance computing hardware.

Cloud computing, also known as "Serverless Computing" (STIGLER 2018), allows the implementation of on-demand self-service computing resources without having to invest in expensive localized high performance hardware. Users can access an online control panel and configure "instances" in the cloud depending on current computing requirements. An instance can be created in a "pay-as-you-go" manner where users are charged only for the time an instance is running. Popular cloud computing services such Amazon Web Services (AWS), Microsoft Azure and Google Cloud can run Linux and Windows multiple instances configured to a user's processing requirements.

In this study, the concepts of Big Data and cloud computing relative to UAV multispectral data acquisition for long-term intensive forest monitoring will be explored. Hereby, proposed protocols for the archiving and processing of acquired UAV-based remote sensing datasets as well as resulting photogrammetric products will be presented.

1.1 Study Site

The Britz research station is an intensive forest monitoring site located near the town of Eberswalde, Germany (Fig. 1) and under administration of the Thünen-Institute for Forest Ecosystems. Since the mid-1970s the site has been implemented for the continuous acquisition of weather and forest hydrology data, as well as individual tree parameters in for example incremental tree growth by dendrometers and tree health with sap-flow sensors. Recently, the site is also being used as a testing ground for the development of UAV-based remote sensing methods to modernize and digitize typical national forest inventories and intensive forest monitoring (i.e. Level II and ICP Forests). Currently, the extraction of individual tree attributes in terms of tree height and crown diameter as well as the quantification of phenological observations provided promising results.



Fig. 1: Britz Research Station in Brandenburg, Germany. Nine tree stands totaling an area of approximately 4 ha.

1.2 High Resolution Datasets

The benefits of employing UAVs and their ability to be mounted with various lightweight active and passive sensors at a high temporal, spatial and angular resolution is of interest for forestry research. Cost-effective consumer-grade optical sensors mounted on UAV Photogrammetry (UAVP) platforms can acquire multi- and hyperspectral imagery for the purpose of creating very dense photogrammetric point clouds as well as very high resolution (VHR) Orthomosaics with Structure of Motion (SfM) software (Fig. 2).


Fig. 2: (a) Aerial Image (b) Point Cloud (c) Orthomosaic (d) Digital Surface Model

For the purpose of assessing changes in forests over time, there is an interest in the capability to acquire data at a high temporal resolution. UAVs can be flown daily depending on weather conditions whereas open access passive satellite-based sensors such as Landsat-8 and Sentinel-2ab have revisit times of 16 and 5 days respectively. UAVs will be limited to a much smaller area (< 5 ha), however can acquire VHR (< 2 cm) imagery which will not be affected by cloud cover in terms of occlusion of the earth's surface. An upscaling of UAV datasets for inventory plots could cover large areas with the incorporation of multispectral satellite data and terrestrial measurements and observations.

In terms of spatial resolution, UAVP can provide a ground sampling distance (GSD) of typically 0.08 to 0.1 meters depending on flying height and the camera's intrinsic parameters (Tab. 1). At such high resolutions, not just individual trees can be detected, however also the apical meristems and individual leaves. Furthermore, high angular resolution refers to the possibility to capture imagery at various angles as opposed to only at nadir (90°). Low oblique imagery can be implemented to parts of trees normally subjected to occlusion from nadir in order to create more complete 3D information, as well as account for the ground area at the base of tree stems. Additionally, terrestrial imagery taken horizontally from the ground floor can account for areas below the forest canopy and aid in the extraction of tree stem diameter. Terrestrial imagery can also be fused with aerial nadir and oblique imagery for the purpose of creating highly dense point clouds depicting an entire forest stand. Table 1 shows the parameters for the Sony A7r RGB and Near-Infrared (NIR) cameras typically used at the Britz research station.

Sensor	Sensor Type	Focal Length	Sensor Size	Bands	GSD at 80 m
Sony A7r	Full Format	35.3 mm	7360 x 4912 pixels	RGB/NIR	0.01 m

Tab. 1: Parameters for the Sony A7r camera.

1.3 Data Acquisition and Processing

During the testing phase, various datasets were acquired from six flight campaigns in early 2018 for the purpose of creating 3D models of forest stands comprised of imagery acquired at the nadir, oblique and horizontal angles. The flight campaigns covered a Scots Pine stand and a European Beech stand of an area of 0.4 ha and 0.5 ha respectively (Fig. 1). During the phenological phases of 2018, weekly flight campaigns capturing multispectral imagery were carried out for the entire Britz research station covering nine tree stands totaling an area of approximately 4 ha. Flight plans were calculated to capture overlapping multispectral imagery for all plots within two missions.

After each flight mission, RAW imagery is copied to a portable hard drive alongside GPX logfiles and flight mission metadata. Later in the office, the data acquired in the field is archived on a local server awaiting further processing. Processing begins with the direct georeferencing of RAW imagery with the GPX logfile. To follow, georeferenced RAW imagery are corrected for vignetting if necessary, and histograms are matched for exposure and converted to 8-bit or 16-bit Tagged Image File Format (TIFF). Images are then loaded into the photogrammetry software and GNSS RTK Ground Control Points (GCPs) are marked within imagery. The SfM photogrammetry software, once set up, processes for the most part semi-automatically depending on the software package. Photogrammetric products such as point clouds, orthomosaics and Digital Surface Models (DSMs) are then further processed for analysis (i.e. Canopy Height Models).

Tab. 2: Data size distribution of a typical nadir flight mission with reference to the processing stage.

Sensor	Format	RAW (Image)	RAW Count	GeoTIFF	GeoTIFF (Selected)	Point Cloud	Orthomosaic
Sony A7r	ARW	35 MB	350	300 MB	300	2 GB	2 GB

2 Proposed Protocols

In order to plan for future flight campaigns and long-term UAV data acquisition, protocols were developed based on the example of a weekly flight campaign for phenological observations carried out with an Octocopter equipped with RGB and NIR sensors. Ideally each flight mission is radiometrically calibrated with a calibration target and flown in winds under 5 m/s to limit tree movement and aircraft instability. Each mission is required to be accompanied with metadata which is updated during the processing chain:

- Average windspeed throughout mission
- Flying height, estimated side- and frontal-overlap, focal length, sensor size, and GSD.
- After SfM processing: sensor intrinsic parameters, camera positions, actual GSD, RMSE, point cloud density, image list, and number of GCP markings.

Table 3 displays a description of the proposed processing levels alongside estimated processing times. RAW images are also converted to Adobe Digital Negative (DNG) format for the purpose

of insuring prolonged support. Depending on storage capabilities, Level 2a-c can be deleted when no longer required. Important to note is that the Level 2a processing stage is only carried out in the case that GCPs and/or radiometric calibration panels were not available.

Product Level	Description	Est. Processing Time
Level 0	RAW imagery is recorded directly from the sensor. Imagery is stored on a portable hard drive alongside GPX log file. All images are included, including those acquired before and after the automatic flight grid.	Directly after flight mission (10 minutes)
Level 1a	RAW data converted to DNG and archived on local server alongside GPX log file and metadata containing flight plan parameters including wind speed and quality rating.	Upon returning to the office (60 minutes)
Level 1b	Copies of original images are georeferenced with the log file, selected for specific plot coverage and uploaded to the S3 (AWS) cloud storage. Plot selection added to ID in database as well as folder name.	30 minutes plus upload time
Level 1c	Georeferenced RAW imagery is copied to an EC2 instance (AWS) and corrected for Vignetting and exposure histograms are matched. Images are converted to 16-bit GeoTIFF and RAW imagery is deleted from the cloud (S3).	60 minutes (mostly automated)
Level 2a	Images are processed with SfM software without GCPs or radiometric calibration. Outputs are point clouds and orthomosaics.	2 hours
Level 2b	Images are processed with SfM software including GCPs.	2-3 hours
Level 2c	Images are processed with SfM software with GCPs and radiometric calibration.	3 – 3.5 hours
Level 3	Point clouds (LAS) and orthomosaics (GeoTIF) are clipped to a selected forest stand, downloaded from cloud and stored on a local server. EC2 instance(s) is (are) terminated.	30 minutes
Level 4a	Canopy Height Model is calculated.	< 10 minutes
Level 4b	Canopy Height Model and indices are calculated.	< 20 minutes
Level 5	Product levels 2a-c, and 3 are made available to the public (Levels 4ab optional).	To be determined

Tab. 3: Proposed processing levels for UAV acquired multispectral imagery (adapted from BIESEMANS & EVERAERTS 2006).

3 Discussion and Outlook

The linking of the archived imagery to metadata is essential for the selection of homogeneous datasets suitable for time-series analysis as well as linkage to satellite imagery for resampling and upscaling. The use of RAW imagery is important for the flexibility in image processing for scientific applications (VERHOEVEN 2010) however for camera specific RAW formats (i.e. Sony ARW) it is unknown whether such formats will be supported in the future. The decision to archive original datasets in DNG format in addition to the Sony ARW format, is based on the concept of "future proofing" (KODDE 2014) with the assumption that Adobe and other software packages will provide DNG support well into the future.

The 16-bit TIFFs created during the processing stages of Level 2a-c can be deleted in order to minimize storage space if necessary. In the event of retroactive processing, TIFFs can be recreated according to defined protocols, where the image list and exposure settings will require to be accessed in the metadata in order to insure identical processing parameters.

With Level 5, the goal would be to provide open-access data to the public in the form of orthoimagery and point clouds for research and educational purposes, and should be in accordance with the European Spatial Data Infrastructure (INSPIRE) initiative (SEIFERT 2011). This could be accomplished by a Web Mapping Service where users search and visualize datasets in a web browser and specific datasets can be downloaded directly or ordered via an Email request.

4 Conclusion

Remote sensing with UAVs for the purpose of intensive forest monitoring can accumulate large amounts of data that require adequate storage capacity and high-performance computing. The possibility to link acquired UAV remote sensing data to terrestrial measurements and observations as well as satellite data is of interest. With an understanding of the concept of Big Data in terms of the volume and complexity of data acquired through UAV remote sensing, robust processing and archiving protocols can be created in order to insure the future integrity and utility of such data. Furthermore, the use of cloud computing can provide high-performance computing capabilities eliminating the necessity to invest in costly hardware.

5 References

- BIESEMANS, J. & EVERAERTS, J., 2006: Image processing workflow for the PEGASUS HALE UAV payload. 2nd International Workshop "The Future of Remote Sensing", ISPRS Inter-Commission Working Group I/V Autonomous Navigation.
- GUO, H., WANG, L., CHEN, F. & LIANG, D., 2014: Scientific big data and Digital Earth. Chinese Science Bulletin, **59**(35), 5066-5073.
- KODDE, M., 2014: Future-proof Data Storage. https://www.giminternational.com/content/article/future-proof-data-storage, accessed 10.12.2018.
- LI, S., DRAGICEVIC, S., CASTRO, F. A., SESTER, M., WINTER, S., COLTEKIN, A., PETTIT, C., JIANG, B., HAWORTH, J., STEIN, A. & CHENG, T., 2016: Geospatial big data handling theory and

methods: A review and research challenges. ISPRS Journal of Photogrammetry and Remote Sensing, **115**, 119-133.

- LIU, P., 2015: A survey of remote-sensing big data. Frontiers in Environmental Science, 3.
- MA, Y., WU, H., WANG, L., HUANG, B., RANJAN, R., ZOMAYA, A. & JIE, W., 2015: Remote sensing big data computing: Challenges and opportunities. Future Generation Computer Systems, 51, 47-60.
- SEIFERT, M., 2011: INSPIRE Data Specification for harmonised Orthoimageries. Photogrammetrie Fernerkundung Geoinformation, **2011**(6), 447-455.
- STIGLER, M., 2018: Beginning Serverless Computing Developing with Amazon Web Services, Microsoft Azure, and Google Cloud. 1st edition, Apress, 216 p.
- TORRESAN, C., BERTON, A., CAROTENUTO, F., DI GENNARO, S. F., GIOLI, B., MATESE, A., MIGLIETTA, F., VAGNOLI, C., ZALDEI, A. & WALLACE, L., 2016: Forestry applications of UAVs in Europe: a review. International Journal of Remote Sensing, 38(8-10), 2029-2036.
- 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.

UHD Mapping von Teststrecken für automatisiertes Fahren

RICHARD LADSTÄDTER¹, PATRICK LULEY¹, STEFAN LADSTÄTTER¹ & HEINZ MAYER¹

Zusammenfassung: Für die Validierung automatisierter Fahrfunktionen auf öffentlichen Straßen (Real-Test), in der Simulation als auch am Prüfstand werden hochgenaue, richtige, vollständige und hochauflösende Referenzkarten der Verkehrsinfrastruktur, sogenannte "Ultra High Definition Karten" (UHDmaps) benötigt. In diesen Karten sollen relevante Objekte wie Verkehrszeichen, Fahrbahnmarkierungen, Betonleitwände etc., aber auch die Fahrbahn selbst als sogenannte "Ground-Truth" repräsentiert sein. JOANNEUM RESEARCH betreibt am Standort in Graz ein Mobile Mapping System, mit dessen Hilfe für die Österreichische Testregion für automatisiertes Fahren (ALP.Lab) im Zeitraum Juli-November 2018 über 400km Autobahnen und Schnellstraßen aufgenommen wurden. In diesem Beitrag werden sowohl die Aufnahme der Teststrecken als auch die weitere Verarbeitung der Daten bis zur hochgenauen Karte beschrieben. Anhand einiger Beispiele soll der hohe Informationsgehalt einer solchen Karte gezeigt werden. Als Ausblick wird skizziert mit welchen Methoden die UHDmaps im Sinne automatischer Kartenupdates aktuell gehalten werden sollen.

1 Einleitung

1.1 Forschungslabor für Hochautomatisiertes Fahren

JOANNEUM RESEARCH betreibt am Institut DIGITAL seit einigen Jahren Forschung zum Thema hochautomatisiertes Fahren (HAF), und hat für diesen Schwerpunkt 2017 begonnen ein eigenes Forschungslabor aufzubauen. Durch eine Förderung aus dem Zukunftsfonds des Landes Steiermark konnte ein Mobile Mapping System der neuesten Generation als zentraler Bestandteil des Labors angekauft werden. Dieses ermöglicht die Erfassung hochgenauer 3D-Daten der Straßeninfrastruktur als Grundlage zur Validierung von HAF-Sensorik bzw. darauf aufbauender Algorithmen.

1.2 ALP.Lab

Die ALP.Lab GmbH (Austrian Light Vehicle Proving Region for Automated Driving) baut in der Steiermark die notwendige Infrastruktur für das Testen von automatisierten Fahrzeugen auf öffentlichen und privaten Straßen als auch in der Simulation und auf Prüfständen. Der Aufbau der Testregion wird im Rahmen des Programms Mobilität der Zukunft der FFG gefördert. Als wissenschaftlicher Partner beschäftigt sich die JOANNEUM RESEARCH innerhalb von ALP.Lab mit der Aufgabenstellung hochgenaue Karten von öffentlichen oder privaten Testtrecken zu erstellen. Die Karten werden zunächst für das öffentliche Streckennetz auf definierten Abschnitten der Autobahnen A2/A9 in der Steiermark erstellt. Kunden der ALP.Lab, wie beispielsweise Hersteller von hochautomatisierten Fahrzeugen, können zur Validierung von Fahrfunktionen auf öffentlichen Straßen auf die Referenzkarten (UHDmaps) selbst oder auch auf Validierungs-Cloud-Services der ALP.Lab zurückgreifen. Während die Referenzkarte eine Validierung der Umfeldwahrnehmung

¹ Joanneum Research Forschungsgesellschaft mbH, Institut DIGITAL, Steyrergasse 17, A-8010 Graz, E-Mail: [richard.ladstaedter, patrick.luley, stefan.ladstaetter]@joanneum.at

von Fahrzeugen hinsichtlich der statischen Objekte ermöglicht, wird bei der Validierung von dynamischen Objekten - wie andere Verkehrsteilnehmern - auf Referenzdaten der digitalen Infrastruktur (Sensoren entlang der Fahrbahn) als auch auf ein dachmontiertes Referenz-Sensor-System zurückgegriffen.

1.3 UHD Karten zur Validierung automatisierter Fahrfunktionen

Da die hohen Anforderungen an Genauigkeit, Detailreichtum und Aktualität von verfügbaren Karten bzw. Vermessungsgrundlagen nicht abgedeckt werden konnte wurde das gesamte ALP.Lab Streckennetz mittels Mobile Mapping im Zeitraum Juli-November 2018 neu aufgenommen. In Anlehnung an den Begriff "HDmaps" (High Definition Navigationskarten der neuersten Generation, die autonomes Fahren unterstützen sollen) ist daher der Begriff "UHDmaps" gerechtfertigt, da hier Absolutgenauigkeiten im niedrigen Zentimeterbereich, höchster Detailierungsgrad durch den eingesetzten Dual-Scanner und hochaktuelles Kartenmaterial angestrebt werden. Während HDmaps von kommerziellen Anbietern in einem globalen Maßstab für den automatisierten Fahrbetrieb selbst erstellt werden, erfüllen UHDmaps die hohen Anforderungen aus Test, Entwicklung und Validierung von automatisierten Fahrfunktionen auf einem regionalen Maßstab für Testregionen.

Die in der Karte benötigten Objekte und deren Erfassung bzw. Attributierung wurden in Zusammenarbeit mit dem ALP.Lab Konsortium festgelegt und lehnen sich stark an den von der DLR entwickelten Leitfaden "Road2Simulation" (RICHTER & SCHOLZ 2018) an. Aufbauend auf diesem Leitfaden wurde von der JOANNEUM RESEARCH eine eigene Digitalisierungsrichtlinie ausgearbeitet, die zunächst die manuelle bzw. semi-automatische Auswertung der erhobenen Mobile Mapping Daten beschreibt. Aufgrund des hohen Personalaufwands für die manuelle bzw. semiautomatische Kartenerstellung müssen große Teile der Karte (und alle weiteren Aktualisierungen) möglichst vollautomatisiert erstellt werden. Der manuell erstellte Teil der Referenzkarte kann dabei zur Evaluierung der Qualität der eingesetzten automatisierten Auswerteverfahren herangezogen werden.

1.4 Das ALP.Lab Streckennetz

Der Kernbereich der Teststrecken (Abschnitt 1) bildet die Autobahn A2 von km170 (Laßnitzhöhe) bis km186 (Knoten Graz West), siehe Abbildung 1. In diesem Bereich wurden zusätzlich insgesamt fünf Passpunktnester mit 40x40cm großen Zieltafeln errichtet. Weiters enthält der ebene, durchgehend dreispurige Abschnitt zwischen Graz Ost (A2 km179) und Graz West zahlreiche stationäre Verkehrssensoren auf Überkopfbrücken, die zur Beobachtung der Verkehrssituation bzw. autonomer Testfahrzeuge eingesetzt werden können. Nach einer Generalsanierung wurde der letztgenannte Abschnitt übrigens erst im Oktober 2018 wieder frei befahrbar.

Neben den beiden Richtungsfahrbahnen Wien bzw. Klagenfurt wurden auch die zahlreichen Rampen der beiden Autobahnknoten Graz Ost bzw. West mitaufgenommen, um Auf- und Abfahrszenarien ebenfalls abbilden zu können. Beim Knoten Graz Ost wurde zusätzlich der Zubringer A22 bis zur Abfahrt Raaba ergänzt, damit ist ein direkter Anschluss an das Magna Firmengelände vorhanden.Der Abschnitt 1 wird auf der A2 Richtung Westen durch einen zweiten Abschnitt bis Steinberg ergänzt. Dieser enthält einen kurvigen Abschnitt mit starker Steigung (ab Mooskirchen) und einen kurzen Tunnel (300m). Die beiden Abschnitte 1 & 2 ergeben zusammen eine Länge von 76km (für beide Richtungsfahrbahnen).



Abb. 1: Übersicht über die im Projekt ALP.Lab aufgenommenen Streckenabschnitte

Als dritter Abschnitt wurde die A9 vom Knoten Graz West bis zum Knoten St. Michael definiert. Dieser Abschnitt enthält neben zwei langen doppelröhrigen Autobahntunneln (Plabutschtunnel, 10km und Gleinalmtunnel, 8km) und drei kürzeren Tunnel auch die Mautstation Gleinalm. Auch dieser Abschnitt weist starke Steigungen/Gefälle und kurvige Abschnitte auf. Ab dem Gleinalmtunnel Nordportal ist zudem öfters mit äußerst winterlichen Bedingungen zu rechnen.

Der vierte Abschnitt des ALP.Lab Streckennetzes schließt beim Knoten Graz West an und führt auf der A9 bis zur österreichischen Staatsgrenze bei Spielfeld und noch ca. 500m weiter bis Sentilj auf der slowenischen Seite. Damit enthält dieser Abschnitt einen Grenzübertritt, was wieder ein eigenes Testszenario darstellt. Dieser 40km lange, sehr flache Autobahnabschnitt ist im Spätherbst sehr anfällig für Bodennebel, was ebenfalls ein wichtiges Testszenario ergibt.

Der fünfte Abschnitt wird durch die Schnellstraßen S35 und S6 von Peggau über Bruck bis St. Michael gebildet, wodurch sich zusammen mit dem Gleinalmabschnitt auf der A9 eine nördliche "Testschleife" ergibt. Die Schnellstraßen haben einen durchgehend sehr kurvigen Verlauf und weisen teilweise sehr enge Kurvenradien auf. Zusätzlich befinden sich in diesem Abschnitt sieben kurze bzw. mittellange Tunnel (bis 4km Länge).

Als sechster Abschnitt wurde schließlich ein ca. 10km langer Nord-Süd Transit durch das Grazer Stadtgebiet definiert (A9/Abfahrt Graz Nord bis Verteilerkreis Webling) der das Streckennetz somit um ein Stadtszenario ergänzt. Insgesamt weist das ALP.Lab Streckennetz somit ca. 425km Länge auf und liegt zwischen 300 und 800 m.ü.A. (Spielfeld bzw. Gleinalm). Das Streckennetz liegt in einem Umkreis von nur 45km zur Landeshauptstadt Graz und hält dennoch wie beschrieben zahlreiche unterschiedliche Testszenarien bereit.

1.5 Definition des Koordinatensystems für ALP.Lab

Für die UHD-Karten der ALP.Lab Teststrecken wird eine Absolutgenauigkeit im niedrigen Zentimeterbereich gefordert. Damit eine derart hohe Genauigkeit bei der Punktbestimmung in 3D überhaupt erreicht werden kann benötigt man ein entsprechend gut definiertes bzw. realisiertes Referenzkoordinatensystem. Eine gute Übersicht über die verfügbaren Koordinatensysteme in Österreich gibt das Bundesamt für Eich- und Vermessungswesen (OTTER et al. 2015).

Nur das System ETRS89 / ETRF2000 / Austria 2002 erfüllt alle gestellten Anforderungen:

- Definitionsgenauigkeit +/- 5mm in der Lage, +/-1cm in der Höhe
- International anerkannte Realisierung des europäischen Referenzsystems ETRS89
- Homogene Genauigkeit daher auch grenzübergreifend gegeben

Die Realisierung erfolgt durch das Netz an APOS-Referenzstationen, in das auch die Station Graz-Lustbühel und Stationen im benachbarten Ausland (Maribor, Zalaegerszeg) eingebunden sind. APOS bietet auch einen Echtzeit-Korrekturdienst (RTK) an, mit dem ebenfalls cm-Genauigkeit direkt im System ETRS89 erreicht werden kann (z.B. für die Positionierung von Fahrzeugen während einer Testfahrt).

Für die Verebnung in der UHD-Karte wurde das internationale System UTM gewählt, im speziellen UTM33 / ETRS89 (EPSG:25833) für die Steiermark (Ostösterreich). Für die Höhenangabe wird direkt die ellipsoidische Höhe im System ETRS89 herangezogen.

2 Datenerfassung mittels Mobile Mapping

2.1 Beschreibung des Leica Pegasus Mobile Mapping Systems

Dem Forschungslabor für hochautomatisiertes Fahren steht seit Mai 2018 ein Mobile Mapping System (MMS) der neuersten Generation zur Verfügung (Pegasus:Two Ultimate Dualhead² von Leica Geosystems, siehe Abbildung 2). Das MMS ist mit folgender Sensorik ausgestattet:

- Zwei Stück Z+F LIDAR Profiler 9012 (+/- 1mm Messgenauigkeit, Rotation bis 200Hz, 1 Million Punkte pro Sekunde, Reichweite 100m)
- Vier seitlich angeordnete hochauflösende RGB-Kameras (12 MPixel, bis 8fps)
- Eine RGB-Panoramakamera (24MPixel, 360x180° FOV)

² https://leica-geosystems.com/products/mobile-sensor-platforms/capture-platforms/leica-pegasus_twoultimate

- Optionale "Pavement"-Kameras zur hochauflösenden Aufnahme der Fahrbahn
 - Novatel SPAN System zur präzisen Positionierung/Orientierung:
 - o GNSS Receiver unterstützt GPS, GLONASS, Galileo und Beidou
 - o IMU-FSAS Inertialmesssystem (200Hz)
 - o Kistler Odometer (externer Radsensor) mit ca. 2mm Auflösung



Abb. 2: Installation des Mobile Mapping Systems auf einem JR Firmenfahrzeug (Foto: B. Bergmann)

Mit Hilfe dieser Sensorik können hochauflösende 3D Punktwolken erzeugt und unter optimalen Bedingungen mit einer Absolutgenauigkeit von +/- 2cm georeferenziert werden. Durch die beiden ca. 45° gegenüber der Vertikalen geneigten und um 90° gegeneinander versetzten Profiler werden Scanabschattungen minimiert, wodurch bereits eine Aufnahmefahrt pro Richtungsfahrbahn ausreichend ist, um dichte und vollständige Scans der Straßeninfrastruktur inklusive der Fahrbahn zu generieren. Bei einem Trigger-Intervall von 4-5m werden pro Stunde ca. 150GByte an Rohdaten erzeugt, mit dem verfügbaren Onboard Speicher können so 6-7 Stunden durchgehend Daten aufgezeichnet werden.

Nach der Lieferung des Pegasus Systems im Mai 2018 wurde nach einer intensiven Einschulungsund Testphase im Juli 2018 mit der Aufnahme der Teststrecken begonnen. In diesem Zusammenhang sei noch erwähnt, dass hier durch das Forschungslabor schon wichtige organisatorische Vorarbeiten geleistet wurden wie z.B. das Einholen von Genehmigungen der Datenschutzbehörde, Zulassung des externen Radsensors durch die Kfz-Prüfstelle etc.

2.2 Aufnahme des ALP.Lab Streckennetzes

Den konkreten Befahrungen mit dem MMS ging eine ausführliche Planungsphase voraus um eine reibungslose Aufnahme unter optimalen Bedingungen garantieren zu können. Da die Autobahnen und Schnellstraßen aufgrund des sehr hohen Verkehrsaufkommens im Großraum Graz für dieses Projekt nicht gesperrt werden konnten war eine enge Koordination mit dem Autobahnbetreiber notwendig. Nur durch ein Rückhalten des Verkehrs mit Begleitfahrzeugen können Scanabschattungen durch überholende Fahrzeuge verhindert werden, daher wurde folgende Befahrungstaktik angewandt:

- Befahrung vorzugsweise am Sonntagvormittag (LKW Fahrverbot, geringeres Verkehrsaufkommen)
- Begleitschutz durch mindestens zwei Fahrzeuge (siehe Abbildung 3)
- Aufnahme von maximal 35km langen Teilabschnitten um zu großen Rückstau zu vermeiden



Abb. 3: Aufnahme der Autobahn A9 im Großraum Graz (Foto: G. Greiner, ALP.Lab GmbH)

Als zusätzliche Rahmenbedingungen sind zu nennen:

- Aufnahme erfolgt mit 80km/h um den Scanlinienabstand auf maximal 15cm zu begrenzen
- Eine erfolgreiche Aufnahme ist nur bei guten Wetterbedingungen möglich (trockene Fahrbahn, gute Lichtverhältnisse), daher vorzugsweise in den Monaten Mai-Oktober
- Aufgrund zahlreicher Baustellen (z.B. Generalsanierung Graz Ost-West) muss deren Fertigstellung abgewartet werden
- Für Spezialaufnahmen der Tunnel (Aufnahme mit 30km/h, Spezialbeleuchtung) müssen Tunnelsperren (meistens in der Nacht) ausgenutzt werden

Dank der ausgezeichneten Zusammenarbeit mit den betroffenen Autobahnmeistereien konnten im Zeitraum Juli – November 2018 in insgesamt zehn Mobile Mapping Missionen qualitativ äußerst hochwertige Daten vom gesamten Streckennetz erfasst werden, in denen praktisch keine Scanabschattungen bzw. "LIDAR-Ghosts" durch vorbeifahrende Fahrzeuge enthalten sind. Dies erleichtert ungemein die weitere Verarbeitung der Daten und ist ein wesentlicher Faktor zur Erreichung einer hohen Qualität der UHDmap.

2.3 Aufbereitung der Mobile Mapping Daten

Um von den aufgenommenen Rohdaten zu den georeferenzierten Bild- und Laserscandaten zu kommen bedarf es vor allem einer sorgfältigen Berechnung der Fahrzeugtrajektorie. Diese wird für jede Mission aus den aufgezeichneten GNSS, IMU und Odometerdaten sowie den zusätzlichen beschafften GNSS-Basisstationsdaten der im Aufnahmegebiet liegenden APOS Referenzstationen abgeleitet. Die Positionen werden daher mittels "Differential GNSS" aus den Basislinienmessungen im System ETRS89 abgeleitet und die Trajektorie schließlich als Fusion aller verfügbarer Messdaten (IMU, Odometer) mittels "Tightly Coupled" Ansatz mit Inertial Explorer berechnet (siehe Abbildung 4). Die Stützung durch das Odometer wirkt sich speziell bei den langen Tunnels positiv auf die berechnete Trajektorie aus. Da diese im Postprocessing zeitlich gesehen als "Forward" und "Reverse" Lösung berechnet werden kann ergeben sich zusätzliche Möglichkeiten zur Qualitätsverbesserung.



Abb. 4: Berechnung der Fahrzeugtrajektorie mit Inertial Explorer (Abschnitt A2 Graz West bis Laßnitzhöhe)

Die erzielbaren Lage-Genauigkeiten können wie folgt abgeschätzt werden:

- +/- 2cm unter optimalen Bedingungen (Open-Sky, kein Multipath)
- <= +/- 5cm unter erschwerten Bedingungen (verringerte Sat.-Anzahl durch GNSS Abschattungen, Multipath)
- +/- 5-10cm bei kurzeitigem totaler GNSS Ausfall z.B. in kurzen Tunnels bzw. leichte Abschattungen/Multipath in bebauten Gebieten
- > +/-10cm bei längerem schlechten GNSS Empfang (z.B. Stop & Go Verkehr in der Stadt mit GNSS Abschattung bzw. Multipath durch hohe Gebäude, totaler GNSS Ausfall in längeren Tunnels)
- In sehr langen Tunnels (>5km) erreicht der Fehler in Tunnelmitte die Größenordnung von einigen Metern

Eine tatsächliche Überprüfung der Lage- bzw. Höhengenauigkeit kann allerdings nur durch unabhängig eingemessene Kontrollpunkte erfolgen. Zu diesem Zweck wurden durch ein Vermessungsbüro auf der A2 im Bereich von km170.4 bis km184.8 insgesamt fünf Passpunktnester mit einer Genauigkeit von +/- 1cm bzw. +/-2cm in Lage bzw. Höhe eingemessen. Ein Vergleich mit den in den Mobile Mapping Daten gemessenen Zieltafeln ergab eine sehr gute Übereinstimmung, die Standardabweichung aus insgesamt 21 Punkten betrug +/- 2.7cm in der Lage und +/- 3.6cm in der Höhe.

3 Erstellung der UHDmap

Abbildung 5 stellt den Verarbeitungsablauf von Mobile Mapping Daten über die semiautomatische Digitalisierung in eine Geodatenbank bis hin zum Export der digitalen Karte in für Simulation oder Analytik verwendbare Formate dar:



Abb. 5: Workflow der UHD Kartenerstellung

Dieses Blockdiagramm beinhaltet ebenfalls einen Verweis auf die laufende Forschung mit dem Ziel einer vollautomatischen Digitalisierung basierend auf Mobile Mapping Daten, die ohne manuelle (Hilfs-)Eingriffe Features ableitet und direkt in die Geodatenbank schreibt. Für die Entwicklung des automatischen Workflows werden manuell annotierte Ground Truth Daten aus der Geodatenbank herangezogen. Im Folgenden werden die Komponenten "Mobile Mapping Daten", "Semiautomatische Digitalisierung", "Geodatenbank", "Konsistenzprüfung" und "UHD Kartenexport" genauer beschrieben.

3.1 Mobile Mapping Daten

Nach der Aufbereitung der Fahrzeugtrajektorie werden der Laserscanpunktwolke absolute Koordinaten im System ETRS89/UTM33 zugewiesen. Neben den 3D Koordinaten in Metern X (Rechtswert), Y (Hochwert) und Z (Höhe über Referenzellipsoid) sind zudem der Reflexionswert und ein dazu assoziierter RGB Farbwert aus der kalibrierten Panoramakamera verfügbar. Dadurch ist eine 3D Datenannotation in der Punktwolke mit optionaler Überlagerung der Bilder der Panorama- und Planarkameras möglich.



Abb. 6: Punktwolke mit überlagertem Panoramabild und Annotationen

3.2 Beschreibung der benötigten 3D Objekte / Geodatenbank

Zum Aufbau der digitalen Karte wird ein Datenbankstrukturschema aufbauend auf Road2Simulation (R2S), einem Leitfaden zur Erhebung von Straßendaten für Simulation und Entwicklung vom Deutschen Institut für Luft- und Raumfahrt (RICHTER & SCHOLZ 2018) benutzt. Hierbei werden Straßenattribute über OGC Simple Feature³ konforme Datentypen (3D Linien und Polygone) mit Metainformationen extrahiert. Die Speicherung der Daten erfolgt direkt in eine auf PostgreSQL 10.5 mit Postgis 2.4 basierende Geodatenbank, die neben optimierten Datenstrukturen für Speicherung von 3D Koordinaten außerdem interne Funktionalität zu deren Prozessierung zur Verfügung stellt (3D Linienschnitt, Distanzberechnung, Datenbereinigung, etc.). Eine Übersicht der in die Datenbank gespeicherten Features stellt Tab. 1 dar.

³ https://www.opengeospatial.org/standards/sfa

Feature	R2S Pendant	Datentyp	Metainformation
Straßenverlauf	r2s_referenceline	LINESTRING Z	Straßenname,
			Richtung, Katego-
			rie
Abbiegebeziehungen	r2s_referenceline	LINESTRING Z	Vorgänger, Nach-
			folger
Fahrspurbegrenzungen	r2s_laneborder	LINESTRING Z	Fahrspurtyp
Fahrbahnmarkierungen	r2s_mark	MULTILINESTRING Z	Typ, Farbe,
			Stärke
Verkehrszeichen und Signal-	r2s_pointobject	POLYGON Z	Typ, Inhalt
anlagen			
Bodenmarkierungen	r2s_pointobject	POLYGON Z	Тур
Überkopfkonstruktionen	r2s_linearobject	MULTILINESTRING Z	Тур
Leitplanken und Betontren-	r2s_linearobject	MULTILINESTRING Z	Тур
ner			
Lärmschutzwände	r2s_linearobject	MULTILINESTRING Z	Тур

Tab. 1: Übersicht erhobene S	straßenfeatures
------------------------------	-----------------

Der Unterschied zu Road2Simulation im Detail ergibt sich durch zusätzliche Speicherung des 3D Umrisses von Punktobjekten zu der 3D Zentrumskoordinate, wie etwa Rechtecke bei Hinweisoder Kreise bei Verbotsschildern (siehe auch Abbildung 7b). Zusätzlich werden Metadaten der Bearbeitungshistorie hinterlegt um eine Versionierung zu ermöglichen.



Abb. 7: Beispiele für die Digitalisierung verschiedener Objekte in Orbit

3.3 Semiautomatische Digitalisierung

Die semiautomatische Auswertung der Punktwolke erfolgt in der Software ORBIT GT Feature Extractor, welche in der Lage ist Mobile-Mapping Trajektorien mit dazugehörigen Punktwolken und Messbildern darzustellen.

In Abbildung 8 ist die Software Orbit mit geladener Punktwolke und einer 3D Linienmessung eines Teilsegments einer unterbrochenen Bodenmarkierung dargestellt. Der obere Teil (Map 2D) zeigt den Grundriss der Punktwolke eingefärbt mit RGB Information, der untere Teil (Mobile Mapping) die dem Panoramabild überlagerte Punktwolke eingefärbt mit Intensitätsinformation.

Rechts ist das aktuell ausgewählte Messwerkzeug zu finden. Die aktuelle Messung ist gelb dargestellt; das Messergebnis mit 6.215 Metern in der Lage und 0.025 Metern in der Höhe kann rechts abgelesen werden.



Abb. 8: ORBIT Feature Extractor Desktop

Zur Extraktion von Features stehen neben manuelle Selektionstools ebenfalls semiautomatische Annotationsworkflows zur Verfügung. So ist in Bereichen mit hoher Punktwolkendichte und gutem Kontrastverhältnis in der Reflexionsintensität (z.B. bei Bodenmarkierungen) bei Spezifikation eines Startpunktes eine automatische Vervollständigung eines Liniensegments bis hin zu mehreren Kilometern möglich (siehe auch Abbildung 7a).

3.4 Herstellen der Datenkonsistenz

Um den Road2Simulation Spezifikationen zu entsprechen ist nach einem (semi-)automatischen Digitalisierungsvorgang eine entsprechende Konsistenzüberprüfung der Kartenfeatures durchzuführen. Dabei werden neben der Sicherstellung von Punkt- und Tangentenstetigkeit in Übergängen logisch zusammenhängender Liniensegmente aufgenommene Daten auf Plausibilität über-prüft und Redundanzen entfernt. Zusätzlich werden Daten, die sich aus erhobenen Features ableiten lassen (z.B.: Fahrspuren und Mittelachsen) generiert.

Die automatische Prozessierung ist darauf ausgelegt, ihre Ergebnisse direkt in die Geodatenbank zurück einzufügen bzw. zu ergänzen. Die Nachbearbeitungsschritte können ebenfalls iterativ ausgeführt werden.

3.5 UHD Kartenexport

Im Rahmen des Projektes wurde eine vollständige manuelle Annotierung folgend der Road2Simulation Spezifikation im Abschnitt Graz West – Graz Ost – Laßnitzhöhe durchgeführt (siehe Ausschnitt in Abbildung 9).



Abb. 9: Beispiel Digitalisierung Knoten Graz West

Das für die Geodatenbank verwendetet Road2Simulation Datenbankschema kann in mehrere Containerformate (z.B.: Geopackage, Spatiallite, Postgres-Dump, GeoJson, Shapefiles) exportiert und mit GIS Bearbeitungssoftware für Analyse und Vergleichszwecke geöffnet werden. Es wird darüber hinaus in weiterer Folge möglich sein, in die Simulationsformate OpenDrive und road5 für IPG Carmaker zu exportieren.

4 Fazit & Ausblick

Die JOANNEUM RESEARCH hat sich mit dem jungen Forschungsbereich für hochautomatisiertes Fahren primär dem drängenden Thema der Validierung von automatisierten Fahrzeugen gewidmet. Zur Sicherstellung der Wirkung der investierten Forschungsgelder werden wichtige Österreichische Akteure aus Wissenschaft und Industrie, welche auch hinter der Österreichischen Testregion für automatisiertes Fahren (ALP.Lab GmbH) stehen, direkt in die Anforderungsdefinition und Evaluierung der Ergebnisse mit eingebunden. Im speziellen lauten die Forschungsfragen (i) welche Inhalte und qualitativen Eigenschaften eine Referenzkarte für die Validierung von automatisierten Fahrzeugen haben muss als auch (ii) wie hochgenaue, richtige und vollständige Referenzkarten (UHDmaps) automatisiert auf Basis von Mobile Mapping Daten erstellt werden können. Schlussendlich wird erforscht wie (iii) eine solche Kartenbasis aktuell gehalten werden kann. Die gegenständige Abhandlung zeigt die Erarbeitung der Grundlagen für die Bearbeitung der oben genannten Forschungsfragen. Ein Verständnis für den Stand der Technik im Bereich Mobile Mapping und semi-automatischer Erstellung von Referenzkarten sowie die Evaluierung von verfügbaren Werkzeugen hat aufgezeigt, wo die Forschung ansetzen muss um mittelfristig eine vollständige Automatisierung der Kartenerstellung erreichen zu können.

Neben der automatisierten Kartenerstellung, welche derzeit bereits in einer ersten Version evaluiert wird, sollen im Rahmen des 2019 startenden strategischen Forschungsprojekts DERIVE vollständige 3D Modelle als auch Material Eigenschaften direkt aus Mobile Mapping Daten abgeleitet werden. Damit werden die Grundlagen für eine realitätsnahe Simulation von virtuellen Sensorsystemen als auch für die externe Stimulation von Sensoren am Prüfstand geschaffen.

Kommerzielle Anbieter von Karten setzen derzeit auf einen manuellen Workflow durch massiven Personaleinsatz in Billig-Lohn-Ländern, weshalb eine Automatisierung basierend auf Methoden der Sensorfusion, Signalverarbeitung und künstlicher Intelligenz neben der wissenschaftlichen Herausforderung auch ein hohes wirtschaftliches Potential darstellt.

Die Absicherung von automatisierten Fahrzeugen zur Gewährleistung des sicheren Betriebs auf öffentlichen Straßen stellt Wissenschaft und Industrie vor große Herausforderungen deren Bewältigung schlussendlich entscheidend für den Erfolg einer automatisierten Mobilität sein wird.

5 Literaturverzeichnis

- OTTER, J., HÖGGERL, N., IMREK, E., STANGL, G. & ZAHN, E., 2015: 3-D Referenzsysteme in Österreich. <u>http://www.bev.gv.at/pls/portal/docs/page/BEV_Portal_Content_Allgemein/</u> <u>0200_Produkte/Schnittstellenbeschreibungen/Systeme_Landesvermessung_2015.pdf</u>, letzter Zugriff 20.12.2018.
- RICHTER, A. & SCHOLZ, M., 2018: Road2Simulation Guidelines Leitfaden zur Erhebung von Straßendaten für Simulation und Entwicklung. Version 1.2, <u>https://www.dlr.de/ts/Portaldata/16/Resources/projekte/road2simulation/TD_TS_R2S_Guidelines_v1.2_DE.pdf</u>, letzter Zugriff 20.12.2018.

Creating Models of Custom Image Classification Workflows Using Machine Learning Techniques

JASON D. WOLFE¹, THOMAS BAHR² & NICOLAI HOLZER²

Abstract: Machine learning algorithms are becoming increasingly popular for image classification due to their robustness and accuracy. This study demonstrates a streamlined process for comparing the accuracy of different machine learning classifiers using a visual modelling tool. The classifiers included ensemble learning methods such as AdaBoost, Random Forest, and Extra Trees; along with traditional methods such as Maximum Likelihood and Minimum Distance. Classification models were built and run using Sentinel-2A imagery of a study area with vegetation land-cover classes that were spectrally similar. The Random Forest and Extra Trees classifiers yielded overall accuracies of 100% for 145,514 training samples. The Ada-Boost classifier yielded an overall accuracy of 93% for 210,235 samples. The traditional Maximum Likelihood and Minimum Distance yielded lower accuracies: 84% and 45% for 147,601 samples.

1 Introduction

With the continued advancement of new machine learning algorithms, the options that are available for image classification have increased significantly. Comparing the performance and accuracy of multiple classification algorithms can be time-consuming, particularly if analysts choose to write programming (API) code. This study shows how a visual modelling tool can provide a streamlined process to building custom image classification workflows.

Visual models allow for easy customization by interchanging different trainers and classifiers. Inputs, outputs, data management operations, and processing tasks can be linked with a drag-anddrop user interface instead of learning API code.

A case study demonstrates the use of simple models to run multiple supervised classification workflows with Sentinel-2A multispectral imagery. The workflows use three machine learning algorithms from the Python "scikit-learn" package (PEDREGOSA et al. 2011):

- AdaBoost
- Extra Trees
- Random Forest

Two additional workflows were built and run using Minimum Distance and Maximum Likelihood classifiers, which are traditional remote sensing algorithms.

¹ Harris Geospatial Solutions Inc, 385 Interlocken Crescent Suite 300, Broomfield, CO 80021, USA, E-Mail: jwolfe05@harris.com

² Harris Geospatial Solutions GmbH, Talhofstraße 32a, D-82205 Gilching, E-Mail: [Thomas.Bahr, Nicolai.Holzer]@harris.com

2 Methodology

A case study was designed to test the effectiveness of ensemble-based machine learning classifiers in discriminating among land-cover classes that were spectrally similar.

2.1 Study Area

The study area is in eastern Texas, USA. The site was chosen because it contains a mix of land cover classes that are predominantly vegetative such as wetlands, forest, agriculture, and range-land. The area of interest is bounded by coordinates $30.96 - 31.04^{\circ}N$, $96.74 - 96.84^{\circ}W$ (Figure 1).



Fig. 1: U.S. Geological Survey orthophoto mosaic of study area (left) and overview map (right) (WOLFE et al. 2018).

2.2 Data Processing

A Sentinel-2A Level-1C image of the study area was used for image classification. The image was acquired on 8 August 2017. Level-1C images are calibrated to top-of-atmosphere reflectance (ESA 2015a). The following processing steps were used with ENVI[®] 5.5 software to prepare the image for analysis:

- Creating a layer stack of the visible and near-infrared bands, while resampling the 20-meter bands to 10 meters.
- Defining a spatial subset around the area of interest.
- Using the Quick Atmospheric Correction (QUAC) tool to correct the image for atmospheric effects. Pixels represent surface reflectance.
- Creating and applying a mask of all non-vegetation features such as roads, water, and buildings.

The resulting dataset contained 10 spectral bands, listed in Table 1.

Band	Wavelength (nm)	avelength (nm) Band	
B2 – Blue	492.4	B6 – Vegetation Red Edge	740.5
B3 – Green	559.8	B7 – Vegetation Red Edge	782.8
B4 – Red	664.6	B8A – Narrow NIR	864.7
B8 – NIR	835.1	B11 – SWIR	1613.7
B5 – Vegetation Red Edge	704.1	B12 – SWIR	2202.4

Tab.	1: Sentinel-2A	spectral	bands and	associated	central	wavelengths	(ESA 2015b)).
						0	、	

Field studies to validate ground truth were not feasible, so training data consisted of samples of image pixels that belonged to five feature types: "Nonforested Wetland", "Rangeland", "Forest", "Grass / Pasture", and "Cropland" (Figure 2). Distinguishing between these land-use types using remote sensing imagery can be challenging since many areas contain a mix of land-use types. In seldom cases the categories are defined by abrupt boundaries, except for possibly "Cropland".



Fig. 2: U.S. Geological Survey orthophoto mosaic of study area with regions of interest (ROI) training samples.

Data-processing steps are described more detailed in WOLFE et al. (2018).

2.3 Machine Learning Classifiers

Machine learning algorithms offer a relatively accurate and efficient way to perform image classification. They can be trained using an initial image and training data, and then used to classify other similar images. For this study, three supervised classifiers were selected from the Python "Scikit-learn" package (PEDREGOSA et al. 2011). These are ensemble learning methods, which train individual decision trees and combine the results to yield the most accurate prediction (SAINI & CHOSH 2017).

The AdaBoost algorithm (FREUND & SCHAPIRE 1995; ZHU et al. 2009) fits a classifier on an original dataset, then fits more copies of the classifier on the same dataset using adjusted weights. This focuses the training on the incorrectly classified pixels.

The Random Forest algorithm has only recently been applied in land-use classification (e. g. KUL-KARNI & LOWE 2016). It is a decision tree-based classifier whereby several classifiers are trained and their results are combined based on the votes given by each tree. Random Forest is effective at estimating the importance of individual variables in multi-source classification.

The Extra Trees, or extremely randomized, algorithm (GEURTS et al. 2006) is a decision tree-based classifier. When looking for the optimal split to separate samples of a node into two groups, it draws random splits for each of the randomly selected features. It then chooses the best split. This method is known for its computational efficiency. It is a more randomized version of Random Forest.

These classifiers have significantly less user-defined parameters than traditional machine learning classifiers such as Support Vector Machine (SVM). This makes them ideal candidates for users who are new to remote sensing.

The Python algorithms were imported into IDL[®] 8.7 and converted to ENVITasks so that the classifiers could run within the ENVI[®] environment.

2.4 Classification Models

Five classification models were built using the ENVI[®] 5.5 Modeler, each using a different classifier. These models were simple to construct as they all had the same basic elements (except for the classifier). Figure 3 shows an example of the Extra Trees model. The yellow nodes represent EN-VITask-based operations such as extracting training data, defining the classifier parameters, classifying the input image, and exporting the accuracy results to a JSON file on disk. Similar models were built for the Random Forest and AdaBoost classifiers.



Fig. 3: Extra Trees classification model.



Fig. 4: Minimum Distance classification model.

When building models, users must define the parameters for each classifier. The following values were used for the decision tree-based classifiers:

- Maximum Nodes: 75
- Number of Trees (Extra Trees and Random Forest): 125
- Number of Estimators (AdaBoost): 100

The Maximum Likelihood and Minimum Distance (Figure 4) classifiers did not require any userdefined parameters other than the input image and training ROIs.

3 Results and Discussions

Figures 5-9 show the classification images produced from each model. Tables 2-6 show the resulting accuracy statistics.

The Extra Trees and Random Forest classifiers yielded overall accuracies of 100% for 145,514 training samples. This high accuracy is not uncommon for ensemble-based machine learning classifiers, particularly with an increased number of nodes (75) and trees (125).

The AdaBoost machine learning classifier yielded an overall accuracy of 93% for 210,235 samples.

In general, these three classifiers effectively discriminated between the five mixed-vegetation land-use types.

The two traditional classifiers – Maximum Likelihood and Minimum Distance – yielded accuracies of 84% and 45% for 147,601 samples.



Fig. 5 & Tab. 2: AdaBoost classification image and classifier accuracy metrics.

AdaBoost	Wetland	Rangeland	Forest	Grass / Pasture	Cropland
Producer Accuracy	0.98	0.93	0.81	0.88	0.97
User Accuracy	1.00	0.92	0.92	0.87	0.95
Error of Commission	0.00	0.08	0.08	0.13	0.05
Error of Omission	0.02	0.07	0.19	0.11	0.02
F1	0.99	0.92	0.86	0.87	0.96
Recall	0.98	0.93	0.81	0.88	0.97
Overall accuracy: 0.93				Kappa coefficient: 0	.89



Fig. 6 & Tab. 3: Random Forest classification image and classifier accuracy metrics.

Random Forest	Wetland	Rangeland	Forest	Grass / Pasture	Cropland
Producer Accuracy	1.00	1.00	1.00	1.00	1.00
User Accuracy	1.00	1.00	1.00	1.00	1.00
Error of Commission	0.00	0.00	0.00	0.00	0.00
Error of Omission	0.00	0.00	0.00	0.00	0.00
F1	1.00	1.00	1.00	1.00	1.00
Recall	1.00	1.00	1.00	1.00	1.00
Overall accuracy: 1.00				Kappa coefficient:	1.00



Fig. 7 & Tab. 4: Extra Trees classification image and classifier accuracy metrics.

Extra Trees	Wetland	Rangeland	Forest	Grass / Pasture	Cropland
Producer Accuracy	1.00	1.00	1.00	1.00	1.00
User Accuracy	1.00	1.00	1.00	1.00	1.00
Error of Commission	0.00	0.00	0.00	0.00	0.00
Error of Omission	0.00	0.00	0.00	0.00	0.00
F1	1.00	1.00	1.00	1.00	1.00
Recall	1.00	1.00	1.00	1.00	1.00
Overall accuracy: 1.00				Kappa coefficient:	1.00



Fig. 8 & Tab. 5: Maximum Likelihood classification image and classifier accuracy metrics.

Maximum Likelihood	Wetland	Rangeland	Forest	Grass / Pasture	Cropland
Producer Accuracy	0.83	0.79	0.93	0.84	0.83
User Accuracy	0.51	0.67	0.94	0.68	0.98
Error of Commission	0.49	0.33	0.06	0.32	0.02
Error of Omission	0.17	0.21	0.06	0.16	0.17
Overall accuracy: 0.84				Kappa coefficient:	0.76



Fig. 9 & Tab. 6: Minimum Distance classification image and classifier accuracy metrics.

Minimum Distance	Wetland	Rangeland	Forest	Grass / Pasture	Cropland
Producer Accuracy	0.50	0.73	0.77	0.57	0.28
User Accuracy	0.26	0.25	0.88	0.40	0.69
Error of Commission	0.73	0.75	0.12	0.60	0.31
Error of Omission	0.50	0.27	0.23	0.43	0.72
Overall accuracy: 0.45				Kappa coefficient:	0.28

4 Conclusions

This research demonstrated how ensemble-based machine learning classifiers were more effective at discriminating among similar land-use classes than traditional classifiers. Python machine learning tools were imported into IDL and converted to ENVITasks. After this initial setup, the ENVITasks could be used in the ENVI® Modeler. The use of models made the process of testing classifiers and comparing results much more efficient, compared to writing API code or invoking tools through a user interface. Model nodes can be easily interchanged to accommodate different trainers and classifiers. Models can also be packaged and deployed to desktop and cloud-computing environments for reuse and further customization. As more machine learning classifiers are implemented in ENVI® software in the future (without the need for a Python-to-IDL interface), the opportunity for building and testing classification models will be simpler and more streamlined.

5 References

- EUROPEAN SPACE AGENCY (ESA), 2015a: Sentinel-2 MSI Technical Guide, Level-1C Algorithm. https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-1c/algorithm.
- EUROPEAN SPACE AGENCY (ESA), 2015b: Sentinel-2 User Handbook. https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook.
- FREUND, Y. & SCHAPIRE, R., 1995: A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting. Computational Learning Theory, EuroCOLT 1995, Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence), 904, VITÁNYI P. (ed), Springer, Berlin/Heidelberg, 23-37.
- GEURTS, P., ERNST, D. & WEHENKEL, L., 2006: Extremely Randomized Trees. Machine Learning, 63(1), 3-42.
- GREEN, W., 2012: Econometric Analysis. 7th Edition, Pearson Education, Boston.
- KULKARNI, A. & LOWE, B., 2016: Random Forest Algorithm for Land Cover Classification. International Journal on Recent and Innovation Trends in Computing and Communication, 4(3), 58-63.
- PEDREGOSA et al., 2011: Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12(Oct), 2825-2830.
- SAINI, R. & GHOSH, S.K., 2017: Ensemble Classifiers in Remote Sensing: A Review. 2017 International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, 1148-1152.
- WOLFE, J., BAHR, T. & HOLZER, N., 2018: Creating Models of Custom Image Classification Workflows Using Softmax Regression and Support Vector Machine. Journal for Geographic Information Science, 2-2018, JEKEL, T., CAR, A., STROBL, J. & GRIESEBNER, G. (Eds.): GI_Forum 2018, 6, 240-260.
- WOLFE, J., JIN, X., T. BAHR & HOLZER, N., 2017: Application of Softmax Regression and its Validation for Spectral-Based Land Cover Mapping. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., 42(1/W1), 455-459.
- ZHU, J., ZOU, H., ROSSET, S. & HASTIE, T., 2009: Multi-class AdaBoost. Statistics and its Interface, 2, 349-360.

Snapshot Hyperspectral Imaging for Field Data Acquisition in Agriculture (in Raspberry Plantation)

ANDRÁS JUNG^{1,2}, MICHAEL VOHLAND², MARIANNA MAGYAR³, LÁSZLÓ KOVÁCS³, TÍMEA JUNG⁴, NÓRA PÉTERFALVI³, BOGLÁRKA KELLER³, FANNI SILLINGER^{1,3}, RENÁTA RÁK³ & KORNÉL SZALAY³

Abstract: A snapshot spectral camera with more than 100 spectral channels was used and tested. A native snapshot imaging spectrometer captures all spectra and the entire image at the same time without any time delay. It enables this imaging system to capture motion pictures and producing hyperspectral videos.

In our horticulture study a snapshot camera was applied to spectrally document, map and characterize a raspberry plantation under differently coloured shade nets to analyse usability, flexibility and to record spectro-phenological parameters. Leading raspberry producers are located in Eastern Europe and contribute at least 80% of the world's raspberry production (FAOSTAT, 2016). Due to agricultural climate change scenarios raspberry plantations are at risk because evapotranspiration will be challenged by solar radiation and temperature changes. We concluded that spectral field data acquisition and length of data evaluation could be significantly reduced by snapshot spectral imaging.

1 Introduction

Hyperspectral technique is a powerful asset to rapidly assess qualitative and quantitative parameters of vegetation (FENYVESI 2008; DEÁKVÁRI et al. 2008). It provides with a unique solution in complex agro-ecosystem monitoring (MILICS et al. 2008; MILICS et al. 2010; JUNG et al. 2017). It has high importance in soil condition and nutrient-supply analysis (CSORBA et al. 2012, 2014; FEKETE 2016). Hyperspectral imaging techniques are used in many agricultural applications. Spectral scanners have successfully mastered several applications; however scanning is facing some limitations when the test object and/or the camera are randomly moving. To eliminate all these limitations snapshot hyperspectral imaging technique can be successfully applied. In principle, this system is capturing the entire hyperspectral image during a single integration time (one shot takes about 1 ms). For this reason we used and tested a snapshot spectral camera with more than 100 spectral channels for a novel agricultural research project.

We applied the snapshot camera to spectrally document, map and characterize a raspberry plantation under differently coloured shade nets to analyse usability, flexibility and to record spectro-phenological parameters. Snapshot hyperspectral imaging provided us a rapid and easy-to-use data acquisition tool for spectral mobile mapping. It worked as non-invasive and non-destructive sampling method with extreme short data acquisition time.

¹ Szent István University, Gödöllő, Hungary

² Leipzig University, Germany

³ NAIK Institute of Agricultural Engineering, Gödöllő, Hungary

⁴ Brightic Research GmbH, Nagymaros, Hungary

2 Materials and Methods

2.1 Experimental site

In order to find a reasonable solution to protect the plants and increase the stability of the production a raspberry plantation with different varieties was established. A sun protective shade tunnel system was constructed to create a test site at NAIK - Fruitculture Research Institute (FRI), Fruit Culture Research and Development Institute of Fertőd, Hungary. It provides opportunity to measure and evaluate relevant biological and physical parameters playing an important role in berry production (Keller et al., 2018). Experimental plantation (Figure 1) of three different raspberry varieties was set in two repetitions: covered and uncovered versions (Table 1).



Fig. 1: Experimental raspberry site at NAIK in Fertőd, Hungary

Each cover has characteristic interaction with light which is expected to generate different environmental conditions and also differences in plant growth and fruit quality. (Table 1).

	Covered								Uncoverred										
			1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8
1. plantation	lon	1. row	Julcsi								1. row	Julesi							
	ntat	2. row	Fertődi zamatos								2. row	Fertődi zamatos							
	I. plai	3. row	Eszterházai kétszertermő								3. row	Eszterházai kétszertermő							
2. plantation	Ion	1. row	Fertődi zamatos						1. row	Fertődi zamatos									
	lantat	2. row	Eszterházai kétszertermő							2. row	Eszterházai kétszertermő								
	7	3. row	Julesi						3. row	Julesi									

Tab. 1: Covered and uncovered plantations with the concerned species

To measure the spectral conditions (JUDD et al. 1964) under the shade nets and the spectral response and features of plants with proximal sensing a portable spectroradimeter and a snapshot hyperspectral camera were used (Figure 2). Portable spectroradiometers can widely be used both in field and laboratory. It is adequate to carry out independent, fast and precise evaluations in an

economic way. The full-range device extends the range of the detectable visible light to NIR and SWIR covering the spectral range of 350 to 2500 nm. These spectrometers have successfully mastered several applications; however scanning is facing some limitations when the test object and/or camera are randomly or rapidly moving in time and space. In agricultural field and close-field applications weight and speed have high importance (JUNG et al. 2017).

Portable spectral field measurements were carried out in the control area and under two different types of tunnels. Data acquisition was made by using ASD FieldSpec 3 MAX portable spectroradiometer (350-2500 nm) with Plant Probe head and Cubert snapshot spectral camera (400-1000 nm) on randomly selected leaves (>100). The same reference panel was used as a standard surface reflecting 95 % of all incident radiation for both acquisition methods.



Fig. 2: ASD FieldSpec 3 MAX (350-2500 nm) and Cuber UHD185 (400-1000 nm)

2.2 Indices

In order to compare the light utilization efficiency, the water and nitrogen management of plants under various shade nets contact measurements vegetation indices were calculated such as Photochemical Reflectance Index – PRI (GAMON et al. 1992), Water Index – WBI (PENUELAS et al. 1995), Normalized Nitrogen Index – NDNI (SERRANO et al. 2002) and Normalized Difference Vegetation Index – NDVI (ROUSE et al. 1973).

3 Results

3.1 Phenological and chemical properties

Studying the phenological features of strawberry plantation we can state that the occurring differences are species-specific therefore the effect of shaded nets weren't or just partly observable. Well distinguishable discrepancies between the plantations could be observed during the blooming and ripening period. Within the covered plantation the blooming was started and finished earlier, while the ripening process indicated the opposite in case of some species (*Eszterházai* and *Julcsi*). In case of *Fertődi zamatos* the blooming and ripening begun later compared to other species. The blooming and ripening period of covered plantation shifted later. Higher yields were reached in case of covered plantations under white Raschel net. The highest yield was measured in case of *Eszterházai*, which was 16-36% higher than in outdoor fields. In the case of *Julcsi*, the yield growing was also occurred under cover with 13-18% rate. The *Fertődi zamatos* showed weaker results than in outdoor, uncovered plantations. In case of average berries, the volume was different. The largest average berries were found at *Julcsi* and

Eszterházai varieties. The volume of production was doubled compared to 2016, as a result of the growth of the plantation. The species-specific features are visible (Table 2).

	20	16	2017				
Coverage	Whole amount (g)	Average berries (g)	Whole amount (g)	Average berries (g)			
White net	156 266	7,22	247 707	6,74			
Black net	89 874	6,83	199 904	6,52			
Control	85 499	7,17	206 167	6,31			

Tab. 2: The mass of whole amount and the average berries during the last two years under different cover system

3.2 Index responses

The processed spectral measurement dataset revealed a tendency of differences in case of photochemical reflectance index (PRI). Based on the dataset the light utilisation efficiency of all varieties under shade nets were higher than in the control plantation. Water index (WBI), normalized nitrogen index (NDNI) and normalized vegetation index (NDVI) did not show differences between treatments (Table 3).

Tab. 3: Different vegetation indices in case of different treatments

21.06.2017											
	White net	Black net	Control								
PRI	0,019	0,047	0,016								
WBI	0,984	0,985	0,989								
NDNI	0,142	0,127	0,142								
NDVI	0,822	0,830	0,822								
	02.08.2017										
	White net	Black net	Control								
PRI	0,039	0,031	0,021								
WBI	1,001	0,982	0,990								
NDNI	0,138	0,139	0,150								
NDVI	0,840	0,829	0,830								

It means the soil preparation; nutrient supply and irrigation assure the favourable homogeneity of the plantation and the only variable between treatments is the difference of illumination. The images acquired with snapshot camera were used to evaluate the shade nets and open sky from below to describe the spectral distribution of the incident light. On the other hand, vegetation survey was performed to visualize reflective features of the vegetation and also the heterogeneity arising within plants. Differences between ratios of significant ranges playing important role in vegetation monitoring to indicate alteration of plant condition (Figure 3).



Fig. 3: Hyperspectral images describing the incident and reflected radiation in the VIS and NIR region

Generally the main indicators of different efficiency of production technologies are yield and fruit quality. Within the frame of the research correlations between shade nets and fruit size was observed. *Julcsi* reacted positively to the white coloured shade net in two consecutive years with extra 72% (2016) and 24% growth of yield. *Eszterházai kétszertermő* reacted with extra 140% (2016) and 104% of yield moreover with extra fruit size 4,6% (2016) and 36,69% (2017). *Fertődi zamatos* did not show such a significant reaction to the treatment.

4 Conclusions

A portable spectroradiometer and a snapshot hyperspectral camera with more than 100 spectral channels were used to spectrally map a raspberry plantation under different type of shade nets. It was also a purpose to test and document the usability, flexibility of such techniques in collecting spectro-phenological parameters in a hand-held way. An in-situ ambient monitoring system was also set up to collect temperature, humidity and radiation data. Spectral techniques provided opportunity to reveal such differences in natural light conditions that are usually not detected by traditional weather stations and make possible to study the correlation between light condition and plant growth in a more complex way.

The results show significant differences between plantations with or without shade net cover. Measured ambient parameters and spectral analysis of the vegetation revealed differences in plant condition and indicated the effect of shade nets and also the differences between the two experimental shade net materials. Shade nets can increase the yield and also increase the average berry size but the reaction to shade nets seems to be variety-specific. Yield reacted very positively to white shade nets in case of *Eszterházai kétszertermő* and *Julcsi* varieties, in two consecutive years. Fruit size reacted significantly positively in 2017. In case of *Fertődi zamatos* variety significant positive effect of shade materials was not confirmed. Variety-specific shade

net-based production technology can offer a solution to improve yield, quality and production stability of raspberry. The tunnel system can be quickly deployed so it offers an immediate action to mitigate the effects of climate changes. It can also serve as an intermediate step towards agroforestry systems or remain a complementary technique with options to be used as spectral filters adjusted to plant needs or physical barriers for pests or to create a microclimate. Although the first synthesises already show useful correlations for the practice further crop years and additional measurements, analyses are necessary to identify the best production practice.

5 Acknowledgements

Authors would like to express their acknowledgements to the Hungarian Ministry of Agriculture to fund this research project and to colleagues at the NAIK - Fruitculture Research Institute (FRI), Fruit Culture Research and Development Institute of Fertőd, Hungary to measure and evaluate relevant biological parameters and for the everyday hard work to maintain and preserve the experimental site. Our special thanks go to Cubert GmbH, Germany, which made it possible to use and test the hyperspectral frame camera. Thank goes to the János Bolyai Research Scholarship of the Hungarian Academy of Sciences to support András Jung's participation in the research project.

6 References

- CSORBA, Á, LÁNG, V., FENYVESI, L. & MICHÉLI E., 2012: Reflektancia spektroszkópia alkalmazása talajok szervesszén-és CaCO3-tartalmának becslésében. Agrokémia és Talajtan, 6(2), 277-290.
- DEÁKVÁRI, J., KOVÁCS, L., PAPP, Z., FENYVESI, L., TAMÁS, J., BURAI, P. & LÉNÁRT, CS. 2008: Az AISA hiperspektrális távérzékelő rendszer használatának első eredményei. MTA AMB XXXII. Kutatási és Fejlesztési Tanácskozás. Gödöllő, 2008. I. kötet 46-50.
- FEKETE, GY., ISSA, I., TOLNER, L., CZINKOTA, I. & TOLNER, I. T., 2016: Investigation on the indirect correlation and synergistic effects of soil pH and moisture content detected by remote sensing. 5th Alps-Adria Scientific Workshop. Mali Lošinj, Croatia, Növénytermelés (Suppl), 203-206.
- FENYVESI, L., 2008: Characterization of the soil plant condition with hyperspectral analysis of the leaf and land surface. Cereal Research Communications, **36**, 659-662.
- GAMON, J. A., PENUELAS, J. & FIELD, C. B., 1992: A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. Remote Sensing of Environment, 41(1), 35-44.
- JUDD, D.B., MACADAM, D. L. & WYSZECLD G., 1964: Spectral distribution of typical daylight as a function of correlated color temperature. J. Opt. Soc. Am., **54**, 1031-1040.
- JUNG, A., DÉNES, F., KOVÁCS, L. & SZALAY, K., 2017: Snapshot Imaging Spectroscopy to Characterize Radiation Scenarios in an Experimental Raspberry Plantation.10th Earsel SIG Imaging Spectroscopy Workshop.

- KELLER, B., JUNG, A., NAGY, G. M., DÉNES, F., PÉTERFALVI, N. & SZALAY, K. D., 2018: Hiperspektrális távérzékelés alkalmazási lehetőségeinek bemutatása egy málna ültetvény példáján keresztül. NAIK – Gödöllő (ISBN 978-615-5748-09-7). 63-72.
- MILICS, G., BURAI, P. & LÉNÁRT, CS., 2008: Pre-Harvest Prediction of spring barley nitrogen content using hyperspectral imaging. Cereal Research Communications Volume 36, pp. 1863-1866. Proceedings of the VII. Alps-Adria Scientific Workshop. Szlovákia, Stara Lesna
- MILICS, G., VIRÁG, I., FAROUK, M. A., BURAI, P. & LÉNÁRT CS., 2010: Airborne hypersrectral imaging for data collection for resilient agro-ecosystems. 9th Alps-Adria Scientific Workshop. Növénytermelés. Špičák, Czech Republic, M. Harcsa (Eds.), Akadémiai Kiadó, 59, 593-596.
- PENUELAS, J., FILELLA, I. & GAMON, J.A., 1995: Assessment of Photosynthetic Radiation-Use Efficiency with Spectral Reflectance. New physiologist, **131**(3), 291-296.
- ROUSE, J. W., HAAS R. H., SCHELL, J. A. & DEERING, D. W., 1973: Monitoring vegetation systems in the Great Plains with ERTS. Third ERTS Symposium, NASA SP-351 I, 309-317.
- SERRANO, L., PENUELAS, J. & USTIN, S. L., 2002: Remote sensing of nitrogen and lignin inMediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. Remote sensing of Environment, 81(2-3), 355-364.

Fernerkundung für Globale Klima-Vegetationsmodelle

MATTHIAS FORKEL¹, WOUTER DORIGO¹, NUNO CARVALHAIS² & KIRSTEN THONICKE³

Zusammenfassung: Vegetationsmodelle als Teil von Klimamodellen sind essentiell um mögliche Auswirkungen zukünftiger Klimaveränderungen auf Ökosysteme abzuschätzen. Die Evaluierung und Verbesserung dieser Modelle anhand von Beobachtungsdaten ist dabei essentiell für eine realistische Modellanwendung. Satelliten erlauben es zahlreiche Aspekte von Landökosystemen wie Vegetationsbedeckung, Pflanzenproduktivität, Biomasse oder Waldbrände weltweit flächendeckend zu quantifizieren und stellen damit essentielle Informationen für die Evaluierung und Verbesserung von Vegetationsmodellen bereit. Dieser Beitrag gibt einen kurzen Überblick über Fernerkundungsdaten und deren Anwendung in globalen Vegetationsmodellen. Weiterhin wird gezeigt, wie multiple Fernerkundungsdaten genutzt werden, um Parameter eines globalen Vegetationsmodells zu kalibrieren und damit Modellsimulationen zu verbessern.

1 Einleitung

Die Fernerkundung der Landoberfläche mittels Satelliten erlaubt es langfristige Veränderungen in Ökosystemen zu beobachten, wie beispielsweise Phänologie, Vegetationsproduktivität, Landbedeckung, Biomasse oder Waldbrände (DE BEURS & HENEBRY 2004; BOYD & DANSON 2005). Satellitendaten zeigen beispielsweise einen früheren Beginn der Vegetationsperiode in vielen Regionen der gemäßigten Zone oder eine Zunahme der Vegetationsbedeckung und produktivität in den nördlichen Nadelwäldern (FORKEL et al. 2015; ZHU et al. 2016). Um jedoch den Einfluss von zukünftigen Klimaänderungen auf Änderungen in Ökosystemen abzuschätzen, müssen Klimamodelle mit Dynamischen Globalen Vegetationsmodellen (DGVM) (PRENTICE et al. 2007) angewendet werden. DGVMs simulieren Ökosystemprozesse wie Photosynthese, die Allokation von Biomasse. Kohlenstoffumsatz, Bodenkohlenstoff, Evapotranspiration, Bodenfeuchte bis hin Waldbränden oder landwirtschaftlicher Produktivität. zu Fernerkundungsdaten sind dabei eine wichtige Datengrundlage um die Plausibilität und Güte von Modellsimulationen zu überprüfen und um schließlich Modelle zu verbessern. Dabei können Fernerkundungsdaten für die Evaluierung von Modellergebnissen, für die Kalibrierung von Modellparametern oder für die Abschätzung von Modellzuständen genutzt werden (DORIGO et al. 2007).

Ziel dieses Beitrages ist es einen Überblick zu geben, wie die Vielfalt von Fernerkundungsdaten für die Evaluierung und Verbesserung von DGVMs genutzt wird. Es wird umrissen welche Fernerkundungsdaten für die Evaluierung und Kalibrierung von DGVMs zur Verfügung stehen

¹ Technische Universität Wien, Department für Geodäsie und Geoinformation, Gußhausstr. 27-29, A-1040 Wien, E-Mail: [matthias.forkel, wouter.dorigo]@geo.tuwien.ac.at

² Max-Planck-Institut für Biogeochemie, Hans-Knöll-Str. 10, D-07745 Jena,

E-Mail: ncarval@bgc-jena.mpg.de

³ Potsdam-Institut für Klimafolgenforschung, Telegraphenberg A31, D-14473 Potsdam, E-Mail: kirsten.thonicke@pik-potsdam.de

und ein Überblick über relevante Arbeiten gegeben. Desweiteren wird anhand des LPJmL-Modelles des Potsdam-Instituts für Klimafolgenforschung (PIK) (SCHAPHOFF et al. 2018a) gezeigt, wie Modellparameter für verschiedene Ökosystemprozesse anhand von multiplen Fernerkundungsdaten geschätzt werden können.

2 Fernerkundungsdaten für Globale Vegetationsmodelle

Zahlreiche satelliten-basierte Datensätze können für die Evaluierung und Kalibrierung von DGVMs genutzt werden (SCHOLZE et al. 2017). Vegetationsindizies wie der Normalized Difference Vegetation Index (NDVI) oder abgeleitete biophysikalische Parameter wie FAPAR (fraction of absorbed photosynthetic radiation) und der Blattflächenindex (LAI, leaf area index) werden genutzt um die simulierte saisonale Blattentwicklung zu evaluieren (BUERMANN et al. 2018; FORKEL et al. 2014). Dabei kommen oftmals Daten des MODIS-Sensors (Moderate Resolution Imaging Spectroradiometer) zum Einsatz, da dieser Sensor homogene Zeitreihen seit 2001 auf einer für globale Modelle idealen räumlichen Auflösung (250-500 m) bereitstellt. Desweiteren werden Vegetationsindizes auch genutzt um Modellparameter zu optimieren, welche die Phänologie (saisonale Blattentwicklung) steuern (FORKEL et al. 2014; KNORR et al. 2010; MACBEAN et al. 2015).

Landbedeckungskarten werden entweder als Eingangsdatensatz oder zur Evaluierung von DGVMs verwendet (POULTER et al. 2011). Dabei müssen die satelliten-basierten Landbedeckungsklassifikationen in die jeweiligen Vegetationstypen der Modelle (sogenannte PFTs, plant functional types) übersetzt werden. Zur Übersetzung wurden Ansätze entwicklt, die jeder Landbedeckungsklasse einen bestimmten Anteil von PFTs zuordnen (LI et al. 2018; POULTER et al. 2015; POULTER et al. 2011). Diese Übersetzungsansätze beinhalten jedoch eine relativ große Unsicherheit hinsichtlich der absoluten Bedeckung einer Gridzelle mit Wald-, Grasland- oder landwirtschaftlicher Bedeckung die sich auf die Unsicherheit in Modellsimulationen auswirkt (GEORGIEVSKI & HAGEMANN 2018).

Abschätzungen der Waldbiomasse aus Satellitenbeobachtungen wurden großräumig für die Tropen (AVITABILE et al. 2016; BACCINI et al. 2012; SAATCHI et al. 2011) und für die nördlichen temperierten und borealen Wälder (SANTORO et al. 2015; THURNER et al. 2014) entwickelt. Diese Biomassekarten können zur Modellevaluierung von Kohlenstoffspeichern oder des Kohlenstoffumsatzes verwendet werden (THURNER et al. 2017; THURNER et al. 2016). Bisher wurden diese Daten aber noch nicht für die Kalibrierung von Modellen genutzt.

Seit einigen Jahren kann die sonnen-induzierte Fluoreszenz (SIF) von Vegetation aus Fernerkundungssensoren abgeschätzt werden und zeigt einen proportionalen Zusammenhang zur Bruttoprimärproduktion (GUANTER et al. 2014). Daher können SIF-Daten zur Evaluierung und Kalibrierung von Photosynthese-Parametern in DGVMs verwendet werden (MACBEAN et al. 2018).

Aus mikrowellen-basierten Fernerkundungssensoren kann die oberflächennahe Bodenfeuchte abgeleitet werden (DORIGO et al. 2017). Satelliten-basierte Bodenfeuchtedaten wurden schon vielfach für die Evaluierung von hydrologischen Modellen verwendet, doch gelangen erst seit kurzem Bedeutung für global Vegetationsmodelle (RAOULT et al. 2018).

Satellitendatensätze der Brandfläche von Wald-, Grasland- oder Buschbränden sind seit vielen Jahren verfügbar und werden u .a. für die Abschätzung von Kohlenstoffemissionen aus Vegetationsfeuern verwendet (VAN DER WERF et al. 2006). Diese Datensätze werden auch für die Evaluierung von Simulationen von Vegetationsfeuern in DGVMs verwendet (FORKEL et al. 2018; LASSLOP et al. 2014; THONICKE et al. 2010).

Evaluierungen von DGVMs gegen einzelne Satellitendatensätze geben einen Einblick in die Leistung eines Modells einzelne Ökosystemprozesse zu simulieren. Jedoch wird daraus die Gesamtgüte des Modelles nicht ersichtlich. Daher wird versucht DGVMs gegen zahlreiche Fernerkundungs- und in situ-Daten zu evaluieren um die Modellgüte umfassend zu beurteilen und mögliche Schwachstellen aufzuzeigen (KELLEY et al. 2013; SCHAPHOFF et al. 2018b). Auch ist es nicht ratsam DGVMs gegen einzelne Datensätze oder Variablen zu kalibrieren, da damit mögliche Datenfehler die Modellsimulationen negativ beeinflussen können oder da einzelne Modellprozesse so zu Lasten von anderen Prozessen überparametrisiert werden können (KAMINSKI et al. 2013; MACBEAN et al. 2016). Daher wird zunehmend versucht Modelle gegen mehrere Beobachtungsdatensätze zu kalibrieren (Tab. 1) (KEENAN et al. 2011).

Studie	Modell	Skala der Optimierung	Fern- erkundungsdaten	In situ Daten	
(Knorr et al. 2010)	BETHY	Global / grob aufgelöst	FAPAR		
(Kaminski et al. 2012)	BETHY	Global / grob aufgelöst	FAPAR	CO ₂	
(Kuppel et al. 2012)	ORCHIDEE	Einzelne Standorte		Nettokohlenstoff- austausch, Latente Wärme	
(Kato et al. 2013)	BETHY	Einzelne Standorte	FAPAR	Latente Wärme	
(Forkel et al. 2014)	LPJmL	Mehrere Gridzellen	FAPAR, Albedo	Bruttoprimär- produktion	
(Kuppel et al. 2014)	ORCHIDEE	Mehrere Standorte		Nettokohlenstoff- austausch, Latente Wärme	
(MacBean et al. 2015)	ORCHIDEE	Mehrere Gridzellen	NDVI		
(Bloom et al. 2016)	DALEC	Global / pro Gridzelle	LAI, Biomasse		
(Scholze et al. 2016)	BETHY	Global / grob aufgelöst	Bodenfeuchte	CO ₂	
(Schürmann et al. 2016)	JSBACH	Global / grob aufgelöst	FAPAR	CO ₂	
(Norton et al. 2018)	BETHY	Global / grob aufgelöst	SIF		
(MacBean et al. 2018)	ORCHIDEE	Mehrere Gridzellen	SIF		

Tab. 1: Übersicht über Studien in denen Fernerkundungsdaten für die Optimierung von globalen Vegetationsmodellen verwendet wurden.
3 Integration multipler Fernerkundungsdaten in das LPJmL DGVM

Fernerkundungsdaten von multiplen Ökosystemparametern können helfen DGVMs umfassend zu verbessern. Bisher wurden jedoch noch nicht mehr als drei Datensätze in einer globalen Modelloptimierung zusammen verwendet (Tab. 1). Im Folgenden wird bewertet wie satellitenbasierte Daten wie FAPAR, SIF, Biomasse und die Landbedeckung verschiedener PFTs gemeinsam genutzt werden können um Simulationen von Vegetations- und Kohlenstoffkreislaufdynamiken im LPJmL-Modell zu verbessern.

3.1 Modell-Daten Integration

LPJmL ist ein DGVM, dass Kohlenstoff- und Wasserflüsse in Landökosystemen und globale Vegetationsdynamik simuliert (SCHAPHOFF et al. 2018a). LPJmL Version 4 berücksichtigt außerdem die Simulation von landwirtschaftlichen Flächen (BONDEAU et al. 2007), Waldbrände (THONICKE et al. 2010), Permafrost (SCHAPHOFF et al. 2013) und eine verbessertes Modul für Phänologie (FORKEL et al. 2014). LPJmL simuliert die globale Vegetationsverteilung als die anteilige Bedeckung von PFTs pro 0.5° x 0.5° Gridzellen. Veränderungen in PFTs sind abhängig von simulierten Änderungen in Pflanzenproduktivität und Biomasse, Sukzession und Pflanzenmortalität. Die simulierte Vegetationsverteilung beeinflusst wiederum FAPAR. Damit stellen Pflanzenbedeckung, FAPAR, Photosynthese und Biomasse im Modell eine positive Rückkopplung dar (Abb. 1).



Abb. 1: Überblick über den Ansatz wie multiple Fernerkundungsdaten mit dem LPJmL-Modell in einer gemeinsamen Kostenfunktion integriert werden um Modellparameter zu schätzen.

Multiple Fernerkundungsdaten wurden genutzt um Modellparameter zu schätzen. FAPAR wurde von MODIS für den Zeitraum 2000-2015 verwendet. SIF basiert auf Messungen von GOME-2 für den Zeitraum 2007-2014. Satellitenbasierte Abschätzungen der Waldbiomasse wurden von einer Karte für die Tropen (AVITABILE et al. 2016) und für die temperierten und borealen Wälder (THURNER et al. 2014) verwendet. Landbedeckung für verschiedene PFTs wurde von der Climate Change Initiative (CCI) der Europäischen Weltraumagentur (ESA) für den Zeitraum 1992-2015 verwendet (LI et al. 2018). Brandflächen stammen ebenfalls von einem ESA CCI-Produkt welches auf MODIS-Daten beruht (CHUVIECO et al. 2018).

Die Satellitenbeobachtungen wurden mit den Modellsimulationen in einer multi-variaten Kostenfunktion kombiniert, die auf dem Bias, der Varianz und der Korrelation zwischen Simulationen *s* und Beobachtungen *o* basiert:

$$Cost = \sqrt{\sum_{d=1}^{d=N} \left(\frac{\overline{s_d}}{\overline{o_d}} - 1\right)^2 + \left(\frac{\sigma_{\{s,d\}}}{\sigma_{\{o,d\}}} - 1\right)^2 + (r(s_d, o_d) - 1)^2}$$

 \bar{s} und \bar{o} sind die Mittelwerte über Raum und Zeit für Simulationen *s* und Beobachtungen *o*. σ_s und σ_o sind die Varianzen und r ist der Korrelationskoeffizient nach Bravais-Pearson. Es wurden insgesamt 27 Modellparameter für jeweils vier PFTs in der borealen Zone, vier PFTs in der temperierten Zone und drei PFTs in der tropischen Zone optimiert. Für die Optimierung von Modellparametern wurde ein genetischer Algorithmus ähnlich wie in einer früheren Studie verwendet (FORKEL et al. 2014).

3.2 Verbesserung von globalen Modellsimulationen mit Fernerkundungsdaten

Die Optimierung des LPJmL-Modells anhand von Fernerkundungsdaten verbessert die Modellsimulationen. Das optimierte Modell repräsentiert besser die räumliche Verteilung von Pflanzentypen (Abb. 2). Beispielsweise wurde die Verteilung von immergrünen Laubbäumen in den Tropen stark unterschätzt und die Verteilung von sommergrünen Laubbäumen überschätzt; diese Fehler wurden nach der Optimierung reduziert.



Abb. 2: Vergleich der prozentualen Bedeckung für zwei Baumtypen entlang der geographischen Breite für Satellitendaten (schwarz) und LPJmL Modellsimulation vor (blau) und nach (rot) der Optimierung.

Weiterhin kann das optimierte Modell die räumliche Verteilung von Biomasse und Gesamtbaumbedeckung besser weidergeben (Abb. 3). Insbesondere wurde die Überschätzung der Biomasse in den Tropen und in den nördlichen Nadelwäldern deutlich reduziert (Abb. 3 a). Die simulierte Baumbedeckung wurde insbesondere in einigen subtropischen (Australien, SW-USA) und borealen Regionen (Sibirien) verbessert (Abb. 3 b). Simulationen der Bruttoprimärproduktion verbesserten sich insbesondere in nördlichen Ökosystemen, wo eine Überschätzung reduziert wurde (Abb. 3 c).

Simulierte Zeitreihen von FAPAR, SIF und Bruttoprimärproduktion von LPJmL sind weltweit hoch korreliert (r > 0.6 in den meisten Regionen) mit den entsprechenden Satellitenbeobachtungen. Die Modelloptimierung mit Fernerkundungsdaten hat nur geringe Veränderungen in der Modellleistung bewirkt.



Abb. 3: Vergleich der Waldbiomasse, prozentualen Gesam-Baumbedeckung und Bruttoprimärproduktion für Satellitendaten (schwarz) und LPJmL Modellsimulation vor (blau) und nach (rot) der Optimierung.

4 Fazit

Fernerkundungsdaten sind zu einem unerlässlichen Datenstrom geworden um globale Klima-Vegetationsmodelle zu evaluieren, zu kalibrieren und damit zu verbesserten. Zahlreiche Fernerkundungsdaten erlauben es inzwischen Veränderungen der Vegetation umfassen zu beschrieben und zu modellieren. Jedoch muss die Photogrammetrie und Fernerkundung in Zukunft weitere Methoden entwickeln und Datensätze erzeugen um Klima-Vegetationsmodelle zu testen und zu verbessern. Insbesondere bedarf es Zeitreihen, welche Veränderungen in der Vegetationsstruktur (z. B. Baumhöhe, Unterwuchs, Biomasse), im Vegetationswassergehalt und

der Biodiversität (z. B. Artenzusammensetzung) beschreiben. Damit leistet die Fernerkundung nicht nur einen wesentlichen Beitrag zum Monitoring von Klimaänderungen sondern auch zu verbesserten Prognosen von Vegetationsveränderungen über die nächsten Tage bis Jahrhunderte.

5 Literaturverzeichnis

- AVITABILE, V., HEROLD, M., HEUVELINK, G.B.M., LEWIS, S.L., PHILLIPS, O.L., ASNER, G.P., ARMSTON, J., ASHTON, P.S., BANIN, L., BAYOL, N., BERRY, N.J., BOECKX, P., JONG, D., J, B.H., DEVRIES, B., GIRARDIN, C.A.J., KEARSLEY, E., LINDSELL, J.A., LOPEZ-GONZALEZ, G., LUCAS, R., MALHI, Y., MOREL, A., MITCHARD, E.T.A., NAGY, L., QIE, L., QUINONES, M.J., RYAN, C.M., FERRY, S.J.W., SUNDERLAND, T., LAURIN, G.V., GATTI, R.C., VALENTINI, R., VERBEECK, H., WIJAYA, A., & WILLCOCK, S., 2016: An integrated pantropical biomass map using multiple reference datasets. Global Change Biology, 22(4), 1406-1420.
- BACCINI, A., GOETZ, S.J., WALKER, W.S., LAPORTE, N.T., SUN, M., SULLA-MENASHE, D., HACKLER, J., BECK, P.S.A., DUBAYAH, R., FRIEDL, M.A., SAMANTA, S., & HOUGHTON, R.A., 2012: Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nature Climate Change, 2(3), 182-185.
- DE BEURS, K.M., & HENEBRY, G.M., 2004: Land surface phenology, climatic variation, and institutional change: Analyzing agricultural land cover change in Kazakhstan. Remote Sensing of Environment, **89**(4), 497-509.
- BLOOM, A.A., EXBRAYAT, J.-F., VELDE, I.R. van der, FENG, L., & WILLIAMS, M., 2016: The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times. Proceedings of the National Academy of Sciences, 113(5), 1285-1290.
- BONDEAU, A., SMITH, P.C., ZAEHLE, S., SCHAPHOFF, S., LUCHT, W., CRAMER, W., GERTEN, D., LOTZE-CAMPEN, H., MÜLLER, C., REICHSTEIN, M., & SMITH, B., 2007: Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Global Change Biology, (13), 679-706.
- BOYD, D.S., & DANSON, F.M., 2005: Satellite remote sensing of forest resources: three decades of research development. Progress in Physical Geography **29**(1), 1-26.
- BUERMANN, W., FORKEL, M., O'SULLIVAN, M., SITCH, S., FRIEDLINGSTEIN, P., HAVERD, V., JAIN, A.K., KATO, E., KAUTZ, M., LIENERT, S., LOMBARDOZZI, D., NABEL, J.E.M.S., TIAN, H., WILTSHIRE, A.J., ZHU, D., SMITH, W.K., & RICHARDSON, A.D., 2018: Widespread seasonal compensation effects of spring warming on northern plant productivity. Nature 562(7725), 110-114.
- CHUVIECO, E., LIZUNDIA-LOIOLA, J., PETTINARI, M.L., RAMO, R., PADILLA, M., TANSEY, K., MOUILLOT, F., LAURENT, P., STORM, T., HEIL, A., & PLUMMER, S., 2018: Generation and analysis of a new global burned area product based on MODIS 250 m reflectance bands and thermal anomalies. Earth System Science Data, **10**(4), 2015-2031.
- DORIGO, W., WAGNER, W., ALBERGEL, C., ALBRECHT, F., BALSAMO, G., BROCCA, L., CHUNG, D., ERTL, M., FORKEL, M., GRUBER, A., HAAS, E., HAMER, P.D., HIRSCHI, M., IKONEN, J., DE JEU, R., KIDD, R., LAHOZ, W., LIU, Y.Y., MIRALLES, D., MISTELBAUER, T., NICOLAI-SHAW, N., PARINUSSA, R., PRATOLA, C., REIMER, C., VAN DER SCHALIE, R., SENEVIRATNE,

S.I., SMOLANDER, T., & LECOMTE, P., 2017: ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. Remote Sensing of Environment, **203**, 185-215.

- DORIGO, W.A., ZURITA-MILLA, R., WIT, A.J.W. de, BRAZILE, J., SINGH, R., & SCHAEPMAN, M.E., 2007: A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. International Journal of Applied Earth Observation and Geoinformation, 9(2), 165-193.
- FORKEL, M., ANDELA, N., HARRISON, S.P., LASSLOP, G., MARLE, M. van, CHUVIECO, E., DORIGO, W., FORREST, M., HANTSON, S., HEIL, A., LI, F., MELTON, J., SITCH, S., YUE, C., & ARNETH, A., 2018: Emergent relationships on burned area in global satellite observations and fire-enabled vegetation models. Biogeosciences Discussions, 1-31.
- FORKEL, M., CARVALHAIS, N., SCHAPHOFF, S., V. BLOH, W., MIGLIAVACCA, M., THURNER, M., & THONICKE, K., 2014: Identifying environmental controls on vegetation greenness phenology through model-data integration. Biogeosciences, 11(23), 7025-7050.
- FORKEL, M., MIGLIAVACCA, M., THONICKE, K., REICHSTEIN, M., SCHAPHOFF, S., WEBER, U., & CARVALHAIS, N., 2015: Codominant water control on global interannual variability and trends in land surface phenology and greenness. Global Change Biology, **21**(9), 3414-3435.
- GEORGIEVSKI, G., & HAGEMANN, S., 2018: Characterizing uncertainties in the ESA-CCI land cover map of the epoch 2010 and their impacts on MPI-ESM climate simulations. Theoretical and Applied Climatology. Internet: https://doi.org/10.1007/s00704-018-2675-2 (19.12.2018).
- GUANTER, L., ZHANG, Y., JUNG, M., JOINER, J., VOIGT, M., BERRY, J.A., FRANKENBERG, C., HUETE, A.R., ZARCO-TEJADA, P., LEE, J.-E., MORAN, M.S., PONCE-CAMPOS, G., BEER, C., CAMPS-VALLS, G., BUCHMANN, N., GIANELLE, D., KLUMPP, K., CESCATTI, A., BAKER, J.M., & GRIFFIS, T.J., 2014: Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. Proceedings of the National Academy of Sciences, 111(14), E1327-E1333.
- KAMINSKI, T., KNORR, W., SCHOLZE, M., GOBRON, N., PINTY, B., GIERING, R., & MATHIEU, P.-P., 2012: Consistent assimilation of MERIS FAPAR and atmospheric CO2 into a terrestrial vegetation model and interactive mission benefit analysis. Biogeosciences, 9(8), 3173-3184.
- KAMINSKI, T., KNORR, W., SCHÜRMANN, G., SCHOLZE, M., RAYNER, P.J., ZAEHLE, S., BLESSING, S., DORIGO, W., GAYLER, V., GIERING, R., GOBRON, N., GRANT, J.P., HEIMANN, M., HOOKER-STROUD, A., HOUWELING, S., KATO, T., KATTGE, J., KELLEY, D., KEMP, S., KOFFI, E.N., KÖSTLER, C., MATHIEU, P.-P., PINTY, B., REICK, C.H., RÖDENBECK, C., SCHNUR, R., SCIPAL, K., SEBALD, C., STACKE, T., VAN SCHELTINGA, A.T., VOSSBECK, M., WIDMANN, H., & ZIEHN, T., 2013: The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges. Journal of Geophysical Research: Biogeosciences, 118(4), 1414-1426.
- KATO, T., KNORR, W., SCHOLZE, M., VEENENDAAL, E., KAMINSKI, T., KATTGE, J., & GOBRON, N., 2013: Simultaneous assimilation of satellite and eddy covariance data for improving terrestrial water and carbon simulations at a semi-arid woodland site in Botswana. Biogeosciences, 10(2), 789-802.

- KEENAN, T., CARBONE, M., REICHSTEIN, M., & RICHARDSON, A., 2011: The model-data fusion pitfall: assuming certainty in an uncertain world. Oecologia, **167**(3), 587-597.
- KELLEY, D.I., PRENTICE, I.C., HARRISON, S.P., WANG, H., SIMARD, M., FISHER, J.B., & WILLIS, K.O., 2013: A comprehensive benchmarking system for evaluating global vegetation models. Biogeosciences, 10(5), 3313-3340.
- KNORR, W., KAMINSKI, T., SCHOLZE, M., GOBRON, N., PINTY, B., GIERING, R., & MATHIEU, P.-P., 2010: Carbon cycle data assimilation with a generic phenology model. Journal of Geophysical Research: Biogeosciences, 115(G4), G04017.
- KUPPEL, S., PEYLIN, P., CHEVALLIER, F., BACOUR, C., MAIGNAN, F., & RICHARDSON, A.D., 2012: Constraining a global ecosystem model with multi-site eddy-covariance data. Biogeosciences, 9(10), 3757-3776.
- KUPPEL, S., PEYLIN, P., MAIGNAN, F., CHEVALLIER, F., KIELY, G., MONTAGNANI, L., & CESCATTI, A., 2014: Model–data fusion across ecosystems: from multisite optimizations to global simulations. Geoscientific Model Development, 7(6), 2581-2597.
- LASSLOP, G., THONICKE, K., & KLOSTER, S., 2014: SPITFIRE within the MPI Earth system model: Model development and evaluation. Journal of Advances in Modeling Earth Systems, **6**(3), 740-755.
- LI, W., MACBEAN, N., CIAIS, P., DEFOURNY, P., LAMARCHE, C., BONTEMPS, S., HOUGHTON, R.A., & PENG, S., 2018: Gross and net land cover changes in the main plant functional types derived from the annual ESA CCI land cover maps (1992–2015). Earth System Science Data, 10(1), 219-234.
- MACBEAN, N., MAIGNAN, F., BACOUR, C., LEWIS, P., PEYLIN, P., GUANTER, L., KÖHLER, P., GÓMEZ-DANS, J., & DISNEY, M., 2018: Strong constraint on modelled global carbon uptake using solar-induced chlorophyll fluorescence data. Scientific Reports, **8**(1), 1973.
- MACBEAN, N., MAIGNAN, F., PEYLIN, P., BACOUR, C., BRÉON, F.-M., & CIAIS, P., 2015: Using satellite data to improve the leaf phenology of a global terrestrial biosphere model. Biogeosciences, **12**(23), 7185-7208.
- MACBEAN, N., PEYLIN, P., CHEVALLIER, F., SCHOLZE, M., & SCHÜRMANN, G., 2016: Consistent assimilation of multiple data streams in a carbon cycle data assimilation system. Geoscientific Model Development, 9(10), 3569-3588.
- NORTON, A.J., RAYNER, P.J., KOFFI, E.N., & SCHOLZE, M., 2018: Assimilating solar-induced chlorophyll fluorescence into the terrestrial biosphere model BETHY-SCOPE v1.0: model description and information content. Geoscientific Model Development, **11**(4), 1517-1536.
- POULTER, B., CIAIS, P., HODSON, E., LISCHKE, H., MAIGNAN, F., PLUMMER, S., & ZIMMERMANN, N.E., 2011: Plant functional type mapping for earth system models. Geoscientific Model Development, 4(4), 993-1010.
- POULTER, B., MACBEAN, N., HARTLEY, A., KHLYSTOVA, I., ARINO, O., BETTS, R., BONTEMPS, S., BOETTCHER, M., BROCKMANN, C., DEFOURNY, P., HAGEMANN, S., HEROLD, M., KIRCHES, G., LAMARCHE, C., LEDERER, D., OTTLÉ, C., PETERS, M., & PEYLIN, P., 2015: Plant functional type classification for earth system models: results from the European Space Agency's Land Cover Climate Change Initiative. Geoscientific Model Development, 8(7), 2315-2328.
- PRENTICE, I.C., BONDEAU, A., CRAMER, W., HARRISON, S.P., HICKLER, T., LUCHT, W., SITCH, S., SMITH, B., & SYKES, M.T., 2007: Dynamic Global Vegetation Modeling: Quantifying

Terrestrial Ecosystem Responses to Large-Scale Environmental Change. In: CANADELL, J.G., PATAKI, D.E., & PITELKA, L.F. (Hrsg.), 2007: Terrestrial Ecosystems in a Changing World, Global Change — The IGBP Series. 175-192.

- RAOULT, N., DELORME, B., OTTLÉ, C., PEYLIN, P., BASTRIKOV, V., MAUGIS, P., & POLCHER, J., 2018: Confronting Soil Moisture Dynamics from the ORCHIDEE Land Surface Model With the ESA-CCI Product: Perspectives for Data Assimilation. Remote Sensing, 10(11), 1786.
- SAATCHI, S.S., HARRIS, N.L., BROWN, S., LEFSKY, M., MITCHARD, E.T.A., SALAS, W., ZUTTA, B.R., BUERMANN, W., LEWIS, S.L., HAGEN, S., PETROVA, S., WHITE, L., SILMAN, M., & MOREL, A., 2011: Benchmark map of forest carbon stocks in tropical regions across three continents. Proceedings of the National Academy of Sciences, 108(24), 9899–9904.
- SANTORO, M., BEAUDOIN, A., BEER, C., CARTUS, O., FRANSSON, J.E.S., HALL, R.J., PATHE, C., SCHMULLIUS, C., SCHEPASCHENKO, D., SHVIDENKO, A., THURNER, M., & WEGMÜLLER, U., 2015: Forest growing stock volume of the northern hemisphere: Spatially explicit estimates for 2010 derived from Envisat ASAR. Remote Sensing of Environment, 168, 316–334.
- SCHAPHOFF, S., VON BLOH, W., RAMMIG, A., THONICKE, K., BIEMANS, H., FORKEL, M., GERTEN, D., HEINKE, J., JÄGERMEYR, J., KNAUER, J., LANGERWISCH, F., LUCHT, W., MÜLLER, C., ROLINSKI, S., & WAHA, K., 2018a: LPJmL4 – a dynamic global vegetation model with managed land – Part 1: Model description. Geoscientific Model Development, 11(4), 1343–1375.
- SCHAPHOFF, S., FORKEL, M., MÜLLER, C., KNAUER, J., VON BLOH, W., GERTEN, D., JÄGERMEYR, J., LUCHT, W., RAMMIG, A., THONICKE, K., & WAHA, K., 2018b: LPJmL4 – a dynamic global vegetation model with managed land – Part 2: Model evaluation. Geoscientific Model Development, 11(4), 1377-1403.
- SCHAPHOFF, S., HEYDER, U., OSTBERG, S., GERTEN, D., HEINKE, J., & LUCHT, W., 2013: Contribution of permafrost soils to the global carbon budget. Environmental Research Letters, 8(1), 014026.
- SCHOLZE, M., BUCHWITZ, M., DORIGO, W., GUANTER, L., & QUEGAN, S., 2017: Reviews and syntheses: Systematic Earth observations for use in terrestrial carbon cycle data assimilation systems. Biogeosciences, **14**(14), 3401-3429.
- SCHOLZE, M., KAMINSKI, T., KNORR, W., BLESSING, S., VOSSBECK, M., GRANT, J.P., & SCIPAL, K., 2016: Simultaneous assimilation of SMOS soil moisture and atmospheric CO2 in-situ observations to constrain the global terrestrial carbon cycle. Remote Sensing of Environment, 180, 334-345.
- SCHÜRMANN, G.J., KAMINSKI, T., KÖSTLER, C., CARVALHAIS, N., VOBBECK, M., KATTGE, J., GIERING, R., RÖDENBECK, C., HEIMANN, M., & ZAEHLE, S., 2016: Constraining a landsurface model with multiple observations by application of the MPI-Carbon Cycle Data Assimilation System V1.0. Geoscientific Model Development, 9(9), 2999-3026.
- THONICKE, K., SPESSA, A., PRENTICE, I.C., HARRISON, S.P., DONG, L., & CARMONA-MORENO, C., 2010: The influence of vegetation, fire spread and fire behaviour on biomass burning and trace gas emissions: results from a process-based model. Biogeosciences, 7(6), 1991-2011.
- THURNER, M., BEER, C., CARVALHAIS, N., FORKEL, M., SANTORO, M., TUM, M., & SCHMULLIUS, C., 2016: Large-scale variation in boreal and temperate forest carbon turnover rate is

related to climate. Geophysical Research Letters. Internet: http://onlinelibrary.wiley.com/doi/10.1002/2016GL068794/abstract (02.05.2016).

- THURNER, M., BEER, C., CIAIS, P., FRIEND, A.D., ITO, A., KLEIDON, A., LOMAS, M.R., QUEGAN, S., RADEMACHER, T.T., SCHAPHOFF, S., TUM, M., WILTSHIRE, A., & CARVALHAIS, N., 2017: Evaluation of climate-related carbon turnover processes in global vegetation models for boreal and temperate forests. Global Change Biology, 23(8), 3076-3091.
- THURNER, M., BEER, C., SANTORO, M., CARVALHAIS, N., WUTZLER, T., SCHEPASCHENKO, D., SHVIDENKO, A., KOMPTER, E., AHRENS, B., LEVICK, S.R., & SCHMULLIUS, C., 2014: Carbon stock and density of northern boreal and temperate forests. Global Ecology and Biogeography, 23(3), 297-310.
- VAN DER WERF, G.R., RANDERSON, J.T., GIGLIO, L., COLLATZ, G.J., KASIBHATLA, P.S., & ARELLANO JR, A.F., 2006: Interannual variability in global biomass burning emissions from 1997 to 2004. Atmospheric Chemistry and Physics, 6(11), 3423-3441.
- ZHU, Z., PIAO, S., MYNENI, R.B., HUANG, M., ZENG, Z., CANADELL, J.G., CIAIS, P., SITCH, S., FRIEDLINGSTEIN, P., ARNETH, A., CAO, C., CHENG, L., KATO, E., KOVEN, C., LI, Y., LIAN, X., LIU, Y., LIU, R., MAO, J., PAN, Y., PENG, S., PEÑUELAS, J., POULTER, B., PUGH, T.A.M., STOCKER, B.D., VIOVY, N., WANG, X., WANG, Y., XIAO, Z., YANG, H., ZAEHLE, S., & ZENG, N., 2016: Greening of the Earth and its drivers. Nature Climate Change, 6(8), 791-795.

Using harmonic analysis of green LAI time series obtained from Sentinel-2 imagery for daily representation of crop growth in a hydro-agroecological model

LUKAS GRAF¹, ISABELLA KAUSCH¹, HEIKE BACH¹ & TOBIAS HANK²

Abstract: Harmonic Analysis of Time Series (HANTS) was used to obtain daily reconstructed values of green leaf area index (GLAI) from optical Sentinel-2 imagery for a maize and a winter wheat field plot in Southern Germany in order to represent crop growth in the hydroagroecological PROMET model. Daily GLAI time series were retrieved with high accuracy in case of winter wheat ($R^2=0.90$, RMSE= $0.78 \text{ m}^2/\text{m}^2$) and slightly lower accuracy in case of maize ($R^2=0.85$, RMSE= $1.05 \text{ m}^2/\text{m}^2$) from the imagery by radiative-transfer model inversion. Assimilating these reconstructed GLAI time series into PROMET using the model forcing technique allowed for retrieving winter wheat fruit biomass ($R^2=0.98$, RMSE= 0.08 kg/m^2) as well as maize above-ground biomass ($R^2=0.97$, RMSE= 0.12 kg/m^2) with considerable high accuracy. The proposed approach offers opportunities for contiguous crop growth monitoring as well as biomass and yield modelling.

1 Introduction

To tackle the rapidly growing demand for food and biomass, the contiguous monitoring of crop growing conditions was announced by the "Group on Earth Observations" (GEOS) as one of the keystones of applied Earth Observation (LAUTENBACHER 2006).

Green leaf area index (GLAI) is a dimensionless biophysical variable that is defined as the onesided area of green leaves per ground unit. It describes the surface area available for the exchange of mass and energy between plants and the atmosphere and therefore is often used for assessing crop development over the growing cycle (MADDONI & OTEGUI 1996). Time series of GLAI derived from optical satellite imagery provide essential information about crop growing conditions for wider areas. This information can be employed to augment physically-based land surface models that simulate the growth of plants according to fluxes of mass and energy.

HANK et al. (2015) have shown that the assimilation of GLAI values derived from optical high resolution satellite data using radiative transfer models improves the accuracy of the agroecological PROMET model by enhancing spatial detail and proving spatially continuous parametrization of crop growing conditions. The time-step of PROMET simulations is one hour, which allows for a highly dynamic representation of diurnal and inter-diurnal crop development based on physical constraints. While the strengths of agro-ecological models like PROMET lie in their generic spatial modelling capacity and high temporal resolution, crop parameters derived from optical satellite imagery, e.g. from Sentinel-2, can be used to re-initialize model runs based on concurrent observations of crop growing conditions to allow for more accurate yield forecasting

¹ VISTA Geowissenschaftliche Fernerkundung GmbH, Gabelsbergerstraße 51,

D-80333 München, E-Mail: [graf, kausch, bach]@vista-geo.de

² Ludwig-Maximilians-Universität München, Department für Geographie, Luisenstraße 37, D-80333 München, E-Mail: [tobias.hank]@lmu.de

and related aspects of agro-ecological modelling (DORIGO et al. 2007). In order to ensure high modelling accuracy, data availability of satellite imagery is a critical aspect, as data assimilation usually often occurs within relatively small time-steps.

Although Sentinel-2 provides a very high temporal resolution with revisit times of up to five days due to the twin constellation of Sentinel-2A and 2B, time series of GLAI derived by inverting radiative transfer models still suffer from data gaps and outliers due to undetected cloud shadows and adjacency effects. Thus, GLAI values cannot be provided on a daily base. In order to overcome these data gaps HANK et al. (2015) developed a methodology of stepwise crop growth simulation using different scenarios of model parameters. Whenever a satellite observation allows for an update of the GLAI values, the model is reinitialized from the last EO assimilation time step to the newest one using the best fitting scenario. This procedure is producing yield maps that correlate strongly with maps from combines (MIGDALL et al. 2009), however the workflow is complex and the computational time high. Therefore it would be considered optimal to use daily GLAI values for assimilation into the PROMET model in order to allow for continuous model forcing instead of stepwise updating.

In order to address this issue, the use of the Harmonic Analysis of Time Series (HANTS) algorithm is proposed (ROERINK et al. 2000). HANTS is capable to, firstly, detect and remove outliers from a time series and, secondly, to decompose a time series into its harmonic components, thereby allowing for the reconstruction of missing values. While HANTS was shown to return accurate results in a broad range of remote sensing domains, including the reconstruction of NDVI time series (JAKUBAUSKAS et al. 2001; ROERINK et al. 2000; ZHOU et al. 2015) and land surface temperature time series (XU & SHEN 2013) as well as the characterisation of land surface types (JAKUBAUSKAS & LEGATES 2000), only a relatively small number of studies has investigated the reconstruction of LAI time series (for instance TUM et al. 2016). Furthermore, to our knowledge, only little efforts have been undertaken to study the use of reconstructed GLAI time-series in agro-ecological models.

Therefore, the objectives of this study were twofold: Firstly, the retrieval accuracy of remote sensing derived GLAI time-series processed by the HANTS algorithm was investigated by comparing the daily outcomes of the HANTS algorithm with in-situ measurements for winter wheat (*Triticum aestivum*) and maize (*Zea mays*) for a test-site in Southern Germany. Secondly, crop biomass values obtained from PROMET simulations with assimilated HANTS-processed GLAI data were compared to field measurements.

The structure of this paper is based upon those objectives: In the first section a brief overview of the used data and the study area is provided. In the second section the workflow and methods are described in more detail while the results are presented and discussed in the third part of this paper.

2 Data and Study Area

In this study, multi-spectral Sentinel-2 imagery in L1B processing level was acquired from Copernicus Scientific Data Hub (https://scihub.copernicus.eu/). Sentinel-2, with its high spatial resolution (10m) as well as with its enhanced spectral band settings, has shown promising results in the retrieval of LAI and related biophysical plant parameters (ATZBERGER & RICHTER 2012; RICHTER et al. 2012). Moreover, its twin constellation allows for the monitoring of plant growing

conditions with high temporal resolution. This high temporal resolution is an essential prerequisite for providing operational monitoring services like GEOGLAM (http://geoglam.org/index.php/en/) that aim for contiguous global crop growth monitoring.

To allow for monitoring a whole growing period, 40 cloud-free Sentinel-2 scenes were downloaded for a study area located north of Munich (Germany). The study site covered two agricultural fields, one cultivated with winter wheat and one with maize. The fields were regularly monitored by insitu measurements of LAI and related leaf and canopy parameters between 1st March and 31st August 2018 (Fig. 1). Twelve field sampling dates at nine neighbouring points were available in case of maize, whereas ten field sampling dates at nine neighbouring points were used in case of winter wheat. Each of the sampling points accounts for exactly one adjacent Sentinel-2 pixel in 10m spatial resolution as can be derived from Fig. 1.



Fig. 1: Overview map showing the location of the study area north of Munich (Germany) and the position of the winter wheat (green) and maize (purple) sampling plots that were examined in this study. In the background a false-colour composite (bands 8, 4, 3) of a Sentinel-2A scene acquired on 31st July 2018 in 10m spatial resolution is displayed.

All in-situ LAI measurements were conducted using a LI-COR LAI 2200C Plant Canopy Analyser (LI-COR Inc.) that provides non-destructive estimates of total LAI. GLAI was computed from these measurements by multiplying the total LAI with the fraction of brown leaves that was estimated during the field visits. Besides GLAI measurements, crop biomass values were acquired destructively. Avoiding disturbance of the continuous non-destructive measurements, the destructive biomass sampling was conducted at adjacent points located in the same field plots. The

biomass samples thus represent locations which were subjected to identical management measures as the points where GLAI was non-destructively obtained.

The selected study area (Fig. 1) reflects a highly-intensified agricultural region, thereby accounting for characteristic farm management measures in terms of fertilizer utilization and nutrient supply. Considering the 1961-1990 records, the study area is dominated by humid conditions (annual precipitation about 810mm) and moderate air temperatures (average temperature 2m above ground about 9.0°C) which are typical for mid-latitude geographic regions characterized by a warm-temperate climate. Precipitation and air temperature in 2018 differed notably from average conditions as the months between April and August 2018 were exceptionally warmer and drier compared to the long-term average (years 1961 to 1990) as indicated by a nearby agarmeteorological weather station in Weihenstephan operated by the Bavarian State Research Centre for Agriculture (LfL Bayern).

3 Methods

3.1 Green LAI retrieval from Sentinel-2 imagery

The acquired Sentinel-2 imagery was converted to surface reflectance values using an approach proposed by VERHOEF & BACH (2003) making use of the MODTRAN (moderate resolution atmospheric transmission) model taking into account atmospheric conditions as well as adjacency effects. To ensure comparable geolocation all scenes were checked for systematic shifts in their x-and y-coordinates to avoid geometrically distorted time series.

First of all, the obtained surface reflectance values were used to run the Soil-Leaf-Canopy (SLC) radiative transfer model (VERHOEF & BACH 2007) in inverse-mode. SLC is a physically-based model that is based on a four-stream radiative transfer approach to account for direct and diffuse fluxes of incident and reflected radiation depending on biophysical as well as biochemical plant parameters and the properties of the underlying soil. Several studies demonstrated the potential of SLC to accurately derive plant parameters for agricultural purposes (MIGDALL et al. 2009; VERHOEF & BACH 2007).

As the inversion of radiative transfer models cannot be done analytically a numerical approach using a Lookup-table (LUT) was established. For each scene and each of the both crop types a LUT was compiled containing plant specific parameters and the current viewing and illumination conditions. Plant parameters were adapted parallel to the phenological development of the crops (expressed as BBCH growth stages) to account for changes of crop physiology and chemistry over the growing period.

During the inversion process, satellite-derived spectra were compared to the spectra stored in the LUT. The Root Mean Squared Error (RMSE) criterion thereby was used as cost function. The synthetic spectrum with the smallest RMSE provided then the GLAI value of a certain pixel. If the RMSE of the matching spectra exceeded a threshold of 2% (expressed as reflectance values) the result of the inversion was flagged as of poor quality and not included into the time-series stack of GLAI values.

3.2 Outlier removal and reconstruction of missing values using HANTS

The obtained GLAI time-series were processed by the HANTS algorithm (ROERINK et al. 2000) using a number of three frequencies. The number of frequencies was predetermined through a Fourier analysis by assessing the number of dominant frequencies in the frequency domain. These frequencies were used to decompose a periodic signal – such as GLAI time-series – expressed by an analytic function f(t) (Eq. 1) into its harmonic components γ and an error term ε in an iterative manner.

$$f(t) = \gamma + \varepsilon$$
 (Eq. 1)

Thereby iterative least square fitting was employed to minimize the error term according to the workflow proposed by JAKUBAUSKAS et al. (2001).

The harmonic components (also referred to as harmonic terms) are mathematically expressed as a superposition of sines and cosines characterized by different frequencies and amplitudes. Data points showing deviations too large from the harmonic terms are removed in each iteration until the number of maintained data points drops below a predefined minimum number of valid observations or a user-defined error-tolerance threshold is exceeded. After each iteration, the harmonic terms are recomputed based on the remaining data points. As a consequence, HANTS is capable to remove outliers from time-series and to restore GLAI values in user-defined equidistant temporal units (e.g. days).

In this study, for each time-series a minimum of nine valid data points was imposed to ensure that the results of HANTS were mathematically robust. Iteration was stopped when the number of valid observations dropped below the minimum number or an error convergence criterion was met. This criterion - expressed as the deviation between the remaining original data points and the fitted harmonic curve - was met when the overall error between the original and reconstructed GLAI values dropped below 1.0 m²/m². The resulting harmonic curve providing GLAI values on a per-day basis was then post-processed to remove possible over- or undershooting that would cause unrealistic representation of crop development. Using the daily reconstructed values, a comparison to field measurements was undertaken in order to investigate whether the retrieval accuracy of GLAI met the criterion of 10% relative accuracy expressed as Normalized RMSE as proposed by the Sentinel mission requirements (ESA 2007).

3.3 Assimilation into the PROMET model

Finally, the reconstructed time-series of GLAI were used to update the agro-ecological PROMET by assimilating the satellite-derived observations into the model (HANK et al. 2015). As PROMET is a raster-based model, assimilation is facilitated by using inter-changeable (artificial) spatial units that are called "proxel" in PROMET and pixel in image analysis. PROMET solves physical equations describing the exchange of mass and energy for each spatial unit hourly. Consequently, GLAI as a canopy variable is provided by PROMET for each spatial unit for every single hour. Assimilation of remotely sensed derived GLAI values into PROMET usually follows one of the three approaches proposed by DELÉCOLLE et al. (1992): Model calibration, parameter updating and model forcing. While the two former assimilation methods do not require continuous input of remote sensing data these approaches cause either large computational efforts in case of parameter updating or physically implausible crop growth simulation by introducing sharp steps in case of

model updating. Model forcing, in contrast, drastically reduces computational efforts by forcing PROMET to follow a given continuous GLAI time series as provided by HANTS.

Crop biomass development was then modelled by PROMET using model forcing over the whole growing period and compared to in-situ measurements in order to quantify crop growth modelling accuracy. Herein, a distinction between maize and winter wheat was made as for maize silage the whole above ground biomass including leaves and stems is usually harvested while in case of winter wheat the fruits (grains) are the most important compartment of the crop.

4 Results and Discussion

4.1 Analysis of GLAI time-series plots

The Sentinel-2 derived time series of GLAI covering the whole crop growing season are displayed for a single ground point in Fig. 2a) and 2b) for maize and winter wheat, respectively. For each time-series the HANTS interpolated values (solid black line), the removed outliers (red dots), the remaining remotely sensed derived values (green dots) as well as the in-situ measured GLAI values (blue dots) were shown.

Generally, all time-series followed an expected temporal pattern of GLAI development starting from values near zero LAI at the beginning of the crop growing period. After seeding, the plants started accumulating biomass and subsequently GLAI increased until a peak around August (maize) and June (winter wheat) was reached. After reaching these peaks, GLAI values declined within a relatively short time interval as the plants matured and green plant pigments were disintegrated. At the end of the observed growing period (harvest) GLAI values were near zero again. Therefore, the HANTS-reconstructed time-series can be considered as physically meaningful. Apart from that, a plateau-like pattern between begin of July and end of August could be identified from the in-situ data for maize (Fig 2a) indicating that GLAI development stagnated during this time period.

In more detail, the time-series plot of maize (Fig. 2a) contained a large amount of removed outliers that were identified from the remotely sensed derived GLAIs. The outliers were partly related to relatively low values during the summer period (July to September) that could be caused by cloud shadows. Additionally, very high GLAI values (> $5 \text{ m}^2/\text{m}^2$) derived during August were marked as outliers. The plateau-like pattern, however, was not completely captured by the HANTS-reconstructed GLAI curve as some of the rejected GLAI values were quite close to the corresponding in-situ measurements. Furthermore, within the plateau-like feature the in-situ measurements indicated the presence of relatively high frequencies as GLAI values showed deand increases within short time periods. Consequently, the HANTS derived curve reflected absolute values of GLAI fairly accurately but did not completely capture the high frequent dynamics present in the in-situ measurements.

In case of winter wheat (Fig. 2b) the derived time series revealed high agreement between the HANTS reconstructed curve and the in-situ measured values. Moreover, the amount of removed outliers was lower as only five outliers within this particular time-series could be identified. These outliers were characterized by clearly lower satellite-derived GLAI values compared to the in-situ measurements. Consequently, the HANTS derived curve quite accurately reflected the temporal



pattern expected from the in-situ measurements in terms of temporal dynamics and absolute values.

Fig. 2: Time series of green LAI values obtained from point number one of maize (a) and winter wheat (b) showing the HANTS reconstructed daily values (solid line), the removed outliers (red dots), the remaining remotely sensed derived values (green dots) and the in-situ measurements (blue dots).



4.2 Analysis of scatter plots of GLA

Fig. 3: Scatter plots of in-situ vs. HANTS-derived GLAI values based on the spatial average of all nine points per crop type showing maize (a) and winter wheat (b). Vertical and horizontal bars indicate the standard deviation of HANTS-derived and in-situ values, respectively. The one-to-one fit is displayed as red line.

Besides these assessments, scatter plots based on a spatial average of all nine points for both crop types (see Fig. 1) for all dates with in-situ measurements were plotted to compare HANTS-derived

GLAI values against in-situ measurements (Fig. 3). The solid red line indicates the desired one-toone relationship. In Fig. 3, for each date the spatial average together with the standard deviation (SD) of the in-situ measured (horizontal bars) and the HANTS-derived values (vertical bars) is shown to account for the observed spatial variation within the two field plots.

In case of maize (Fig. 3 a) HANTS-derived values were in considerable agreement to in-situ measured GLAIs resulting in $R^2 = 0.85$ and a RMSE of 1.05 m²/m². As shown by the vertical and horizontal bars in Fig. 3 a SD was up to +/- 0.5 m²/m² green LAI indicating a relatively high degree of spatial variation for each date. Normalized RMSE (NRMSE) was 37.5% that exceeded the desired accuracy of 10%.

The relatively high NRMSE for maize could be either related to an inaccurate parametrization of maize within the SLC model or to an inappropriate setting of the parameters used within HANTS (e.g. inappropriate number of frequencies). Furthermore, in multi-temporal model inversion processes, as presented in this study, phenological stages of the crop growing cycle are determined by a distinct set of leaf and canopy parameters that change over the time. If extreme weather conditions like during summer 2018 cause phenological patterns that differ significantly from average conditions it is feasible to assume that the model parametrization did not completely capture derivations from the usual crop growing cycle. Especially the stagnation of GLAI development during the summer months (see Fig. 2a) was not well captured by the HANTS reconstructed curve. This finding could explain the deviation between in-situ measured and HANTS-derived GLAI values.

Regarding winter wheat, the relationship between HANTS-derived and in-situ measured GLAI came close to the one-to-one line (red line) as also indicated by a high R² (0.90) and a low RMSE (0.78 m²/m²). The NRMSE was about 12.2%, thus almost reaching the targeted accuracy of 10%, and was in good agreement to results of recent studies on GLAI retrieval from optical satellite data (DUVEILLER et al. 2011, ALI et al. 2015).

4.3 Crop Growth Modelling Accuracy

In-situ measured and PROMET-simulated crop biomass values acquired during the growing period were plotted in Fig. 4 showing the development of leaf, stem, fruit (grain) and overall above ground biomass over time for maize (Fig. 4a) and of leaf and fruit biomass for winter wheat (Fig. 4b).

In general, all curves of biomass development followed an expected temporal pattern: Fruit biomass accumulated over the growing period and reached a maximum at the time of harvest (begin of September and mid of July for maize and winter wheat, respectively). Leaf and stem biomass, in contrast, reached their maximum earlier (begin of August and begin of June, respectively) and declined afterwards as plants matured.

In case of maize (Fig. 4a) overall above ground biomass was simulated with fairly high accuracy when compared to in-situ measurements ($R^2=0.97$, RMSE=0.12 kg/m²). The NRMSE was about 15.0% thus coming quite close to the targeted accuracy of 10%. As for maize silage the overall above ground biomass is usually harvested completely, the results reveal promising opportunities for maize yield modelling. Furthermore, the potential of the proposed approach for continuous monitoring of maize growing conditions was demonstrated as PROMET was capable to model biomass development over the whole growing period for all crop compartments.

Regarding winter wheat (Fig. 4b), fruit (=grain) biomass was simulated with considerable accuracy (R^2 =0.98) resulting in a RMSE value of 0.08 kg/m². NRMSE was about 24.2% that exceeded the targeted accuracy of 10%. At the time of harvest (13th July), however, the relative error between measured and simulated fruit biomass was about 10.6% thus fulfilling the accuracy requirements at the usually most important stage of crop growth modelling. In addition, the temporal patterns of crop development simulated by PROMET showed an overall good agreement to the in-situ measurements. Thus, the proposed approach seems to be capable for contiguous winter wheat growth monitoring as well as yield forecasting.



Fig. 4: Time series plot of above ground (red), fruit (blue), stem (brown) and leaf (green) biomass showing PROMET-simulated (solid lines) and in-situ measured (dots) values for maize (a) and winter wheat (b). For winter wheat only fruit and leaf biomass were assessed.

5 Conclusions and Outlook

The presented results indicate the potential of assimilating daily re-constructed GLAI values into the process-based agro-ecological PROMET model using the HANTS algorithm. HANTS was capable to provide physically meaningful and accurate curves of daily GLAI development that could be used for computational efficient model forcing assimilation techniques. PROMET simulations allowed to retrieve crop biomass with high accuracy over the whole growing cycle thereby offering opportunities for continuous crop growth monitoring and yield estimation. In the future, the usage of HANTS should be extended from the single-pixel level to entire agricultural fields or even wider areas to assess its capacity in an operational context. By doing so, the impact of daily reconstructed GLAI values on crop biomass modelling could be discussed with regard to spatial variability – an essential aspect in site-specific precision farming - that can be hardly completely captured by the presented single-point considerations.

6 Acknowledgement

This study was carried out within the VirtualWaterValues (ViWA) project. ViWA is a collaborative project of the funding program "Global Resource Water (GROW)" in the framework program FONA (Research for Sustainability) of the German Ministry for Education and Research (BMBF) under grant code 02WGR1423A. In-situ data were collected as part of the research project "EnMAP Scientific Advisory Group Phase III Developing the EnMAP Managed Vegetation Scientific Processor", supported by the Space Agency of the German Aerospace Center (DLR) through funding by the German Federal Ministry for Economic Affairs and Energy under the grant code 50EE1623.

7 Literature

- ALI, M., MONTZKA, C., STADLER, A., MENZ, G., THONFELD, F. & VEREECKEN, H., 2015: Estimation and Validation of RapidEye-Based Time-Series of Leaf Area Index for Winter Wheat in the Rur Catchment (Germany). Remote Sensing, 7, 2808-2831.
- ATZBERGER, C. & RICHTER, K., 2012: Spatially constrained inversion of radiative transfer models for improved LAI mapping from future Sentinel-2 imagery. Remote Sensing of Environment, 120, 208-218.
- DELÉCOLLE, R., MAAS, S., GUERIF, M. & BARET, F., 1992: Remote sensing and crop production models: present trends. ISPRS Journal of Photogrammetry and Remote Sensing, 47(2-3), 145-161.
- DORIGO, W.A., ZURITA-MILLA, R., WIT, A.J.W. DE & BRAZILE, J., 2007: A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. International Journal of Applied Earth Observation and Geoinformation, 9, 165-193.
- DUVEILLER, G., WEISS, M., BARET, F. & DEFOURNY, P., 2011: Retrieving wheat Green Area Index during the growing season from optical time series measurements based on neural network radiative transfer inversion. Remote Sensing of Environment, **115**, 887-896.

- EUROPEAN SPACE AGENCY (ESA), 2007: GMES: Sentinel-2 Mission Requirements Document, issue 2 revision 0-30/01/2007. 1-13.
- HANK, T.B., BACH, H. & MAUSER, W., 2015: Using a remote sensing-supported hydroagroecological model for field-scale simulation of heterogeneous crop growth and yield: Application for wheat in central Europe. Remote Sensing, 7, 3934-3965.
- JAKUBAUSKAS, M.E. & LEGATES, D.R., 2000: Harmonic analysis of time-series AVHRR NDVI data for characterizing us great plains land use/land cover. International Archives of Photogrammetry and Remote Sensing, **33**, 384-389.
- JAKUBAUSKAS, M.E., LEGATES, D.R. & KASTENS, J.H., 2001: Harmonic Analysis of Time-Series AVHRR NDVI Data. Photogrammetric Engineering and Remote Sensing, **67**(4), 461-470.
- LAUTENBACHER, C.C., 2006: The Global Earth Observation System of Systems : Science Serving Society. Space Policy, **22**, 8-11.
- MADDONI, G.A. & OTEGUI, M.E., 1996: Leaf area, light interception, and crop development in maize. Field Crops Research, 48, 81-87.
- MIGDALL, S., BACH, H., BOBERT, J., WEHRHAN, M. & MAUSER, W., 2009: Inversion of a canopy reflectance model using hyperspectral imagery for monitoring wheat growth and estimating yield. Precision Agriculture, **10**(6), 508-524.
- RICHTER, K., HANK, T.B., VUOLO, F., MAUSER, W. & D'URSO, G., 2012: Optimal Exploitation of the Sentinel-2 Spectral Capabilities for Crop Leaf Area Index Mapping. Remote Sensing, 4, 561-582.
- ROERINK, G.J., MENENTI, M. & VERHOEF, W., 2000: Reconstructing cloudfree NDVI composites using Fourier analysis of time series. International Journal of Remote Sensing, 21(9), 1911-1917.
- TUM, M., GÜNTHER, K.P., BÖTTCHER, M., BARET, F., BITTNER, M., BROCKMANN, C. & WEISS, M., 2016: Global Gap-Free MERIS LAI Time Series (2002 2012). Remote Sensing, 8, 1-19.
- VERHOEF, W. & BACH, H., 2007: Coupled soil-leaf-canopy and atmosphere radiative transfer modeling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data. Remote Sensing of Environment, 109, 166-182.
- VERHOEF, W. & BACH, H., 2003: Remote sensing data assimilation using coupled radiative transfer models. Physics and Chemistry of the Earth, **28**(Parts A/B/C), 3-13.
- XU, Y. & SHEN, Y., 2013: Reconstruction of the land surface temperature time series using harmonic analysis. Computers & Geosciences, **61**, 126-132.
- ZHOU, J., JIA, L. & MENENTI, M., 2015. Reconstruction of global MODIS NDVI time series : Performance of Harmonic Analysis of Time Series (HANTS). Remote Sensing of Environment, **163**, 217-228.

Seevermessungsprojekt Wörthersee als zukunftsweisendes Beispiel für Sensorfusion

THOMAS PIECHL¹, HERBERT DÖLLER²

Zusammenfassung: Für den Wörthersee sollte ein zukunftsweißendes 3D-Modell nach den modernsten Methoden der hydrographischen Vermessung geschaffen werden. Es kamen verschiedenste Sensortechniken zum Einsatz, der Ergebnisse zu den nunmehr vorliegenden 3D-Modell zusammengeführt wurden. Mit diesen hochauflösenden Daten ist sohin die Basis für eine Vielzahl von Aufgaben geschaffen worden. Eine neugewonnene Planungssicherheit ist das Ergebnis dieser Sensorfusion.

1 Einleitung - Motivation

Im Jahr 1976 wurde die letzte Tiefenvermessung des Wörthersees als "Neuauslotung des Wörthersees mit Hilfe eines Echographen" (SCHULZ & SCHULZ 1976) durchgeführt. Aus den 49 Tiefenprofilen wurde eine Karte mit Tiefenlinien (10m Isolinien) im Maßstab 1:25.000 abgeleitet.

40 Jahre später soll nunmehr das "integrale Seevermessungsprojekt Wörthersee" einerseits ein detailgenaues 3D-Modell des Seebeckens liefern, anderseits auch der Seeuntergrund sowie die Ufer- bzw. Flachwasserbereiche untersucht werden. Das neu erstellte digitale Unterwassergeländemodell des Wörthersees soll im Uferbereich nahtlos an das bestehende digitale Geländemodell aus Laserscandaten anschließen. Bei der Erfassung des Seeuntergrundes wird nun auch die Art und Mächtigkeit der Sedimentauflage dokumentiert. Die Einsatzbereiche der neuen Daten sind vielfältig und sicher noch nicht zur Gänze absehbar.

1.1 Wörthersee

Mit einer Fläche von 19,39 km² und 16,5 km Länge ist der Wörthersee der größte See Kärntens. Er liegt in einer Talfurche im Kärntner Zentralraum, die vom eiszeitlichen Gletscher überformt wurde. Der See erstreckt sich in Ost-West Richtung und ist topographisch in drei Becken unterschiedlicher Tiefe gegliedert. Das westliche und tiefste reicht von Velden bis Pörtschach (Fläche 7,9 km², max. Tiefe ca. 85 m), das mittlere von Pörtschach bis Maria Wörth (Fläche 3,5 km², max. Tiefe ca. 40 m) und das östliche von Maria Wörth bis Klagenfurt (Fläche 8,0 km², max. Tiefe ca. 73 m).

Den Zufluss des Wörthersees bilden eine Vielzahl kleinerer Bäche, die von allen Seiten in den See einmünden. Der größte von ihnen ist der Reifnitzbach mit einer mittleren Wasserführung von 0,63 m³/s. Der Abfluss, die Glanfurt/Sattnitz, verlässt den See im Osten und entwässert über die Glan und die Gurk zur Drau. (SCHULZ et al. 2004)

¹ Land Kärnten, Abteilung Umwelt, Energie und Naturschutz, Flatschacherstraße 70,

A-9020 Klagenfurt, E-Mail: thomas.piechl@ktn.gv.at

² Dr. Döller Vermessung ZT GMBH, Johannesgasse 15/1, A-1010 Wien, E-Mail: office.wien@doeller.biz

2 Ziele & Nutzen

Die exakte Erfassung von See-Einbauten oberhalb und unterhalb des Wasserspiegels (Einleitungen, Wasserentnahmen, Seedruckleitungen, Stege, ...) stellt die Basis für die Erarbeitung von Notfallplänen für die Betreiber von Leitungen im Seen dar. Für die Detailplanung von neuen Seedruckleitungen werden genaue Tiefenkarten und die exakte Lage bestehender Leitungen benötigt. Mit zeitgemäßer Echolot-Technologie (Single-Beam-Aufnahmen im Uferbereich und Multi-Beam-Aufnahmen im Tiefenbereich) wird der exakte 3D Verlauf der bestehenden Seedruckleitungen erfasst. Mit Hilfe der Ergebnisse können z.B. Tauchgänge für Reparaturarbeiten oder der Verlauf von zusätzlichen Leitungen exakt geplant werden

Neben der schon genannten Planungssicherheit werden die gewonnenen Daten auch für die Bereiche Sicheres Tauchen, Beobachtung von Hangrutschungen im See mit neuen Methoden (Faseroptische Systeme), Kartengrundlagen für Kampfmittelbeseitigung, Wracks und historische Einbauten, Grundwasserzutritte, Trinkwasserentnahmen und Dokumentation des Sedimenttransports benötigt.

Die EU-Wasserrahmenrichtlinie sieht bei der Erhebung des ökologischen Zustandes stehender Gewässer nicht nur chemisch-physikalische bzw. biologische Parameter vor, sondern auch die Bewertung morphologischer Parameter im Bereich der Seeufer. Ein wesentlicher Parameter dabei ist der Verbauungsgrad der Gewässer. Dazu sollen die überbauten Seeflächen und Uferbereiche detailliert erfasst werden. Für diese Kartierung der Morphologie werden hochauflösende Orthofotos aus einer Drohnenbefliegung eingesetzt und mit den Daten aus der Echolotaufnahme verschnitten (Sensorfusion).

3 Methodik & Sensoren

Die 3D-Tiefenerfassung erfolgte im Zeitraum 3./4. Ouartal 2017 sowie mit Ergänzungsmessungen im 2. Quartal 2018. Die Erfassung der Gewässersohle wurde mit einer hochpräzisen Sensorgruppe bestehend aus einem Fächerlot der neuersten Generation SeaBat, einem GNSS Heading System Aste RX-U Marine und einem Wasserschallprofilmessgerät VALEPORT Swift SVP durchgeführt. Die Positionierung erfolgte über RTK-DGNSS. Die Sedimentecholotmessung (Sub Bottom Profiling) erfolgte mit System SES-2000. Die Durchführung der Singlebeam-Echolotung wurde mit der neuesten Generation des Admodus Sonar durchgeführt. Für die Bestimmung der 3D-Koordinaten der Messschiffe wurde ein einheitliches homogenes Bezugssystem realisiert.

3.1 Subbottom Echolot

Für die Erfassung des Seeuntergrundes wurde das parametrische Sedimentecholotsystem SES-2000 light der Innomar Technologie GmbH benutzt, welches den parametrischen akustischen Effekt zur Schallerzeugung nutzt. Zwei hochfrequente Schallwellen werden abgestrahlt, interagieren in der Wassersäule und bilden eine tieffrequente Komponente. Das SES-2000 System benutzt Primärfrequenzen um die 100 kHz und kann Differenzfrequenzen zwischen 5 kHz und 15 kHz erzeugen. Die hohen Frequenzen können zur Wassertiefenbestimmung

verwendet werden und die tiefen Frequenzen sind in der Lage in den Boden einzudringen. Die reflektierten tiefen Frequenzen geben Informationen über Bodenstrukturen, Sedimentschichten und eingebettete Objekte.

Für die konkrete Messung wurde das S-2000 light System an Bord eines elektrischen Motorbootes installiert. Der Schallwandler wurde mit einer Tauchtiefe von 27 cm installiert, damit keine Blasen die Aufzeichnung stören. Zur Positionierung wurde eine Leica GS 16 GPS Antenne 2 m direkt über dem Schallwandler montiert. Alle Höhenangaben basieren auf dem offiziellen Wasserpegel, publiziert vom Hydrographischen Dienst des Landes. Für den Wörthersee liegt der Pegelnullpunkt bei 439.01 m (müA). Der Pegelstand während der Messperiode lag durchgehend bei 1.26 m. Die Wassertemperatur war mit 14.3° C angegeben, was in einer gemittelten Wasserschallgeschwindigkeit von 1469 m s-1 resultiert und welche für die gesamte Laufzeit-Tiefenumrechnung benutzt wurde.



Abb. 1: Wörthersee mit Profilspuren – Subbottom -Echolot

3.1.1 Ergebnisse

Nur in flachen Bereichen des Wörthersees ist die erste Multiple sichtbar (Eigenschaft 9 – ein scheinbares Echo, welches von Schallwellen herrührt, die mehrfach zwischen Boden und Seeoberfläche hin und her laufen). Sehr weiche Sedimente können überall im See gefunden werden und resultieren in guter akustischer Eindringung bis zum festen Seegrund (Basement oder auch Grundgebirge) oder älteren kompakten Sedimenten (Eigenschaft 5) mit unebener Morphologie. Das Basement ist mit parallel geschichteten, weichen Sedimenten verfüllt (Eigenschaften 2 und 4), welche feine parallel Reflektoren ausbilden. Helle Schichten wechseln sich im Echogramm mit dunkleren Schichten höherer Reflektivität ab. Einige große Rutschungen, wie z.B. Eigenschaft 3, können im ganzen See beobachtet werden und schneiden und überdecken die fein geschichteten Sedimente. Die Sedimente einiger Rutschungen sind akustisch transparent, andere zeigen chaotische und kissenförmige Strukturen. Einige jüngere Rutschungen (Eigenschaft 1) sind noch nicht von parallel geschichteten Sedimenten überdeckt.



Abb. 2: 2D Sektion von Profil Q18_20171026_130243LF mit einigen akustischen und geologischen Eigenschaften (1-9), der gelbe Pfeil gibt die Lage und Richtung des Profils an

Der Auslöser für die Rutschungen ist wahrscheinlich verbunden mit tektonischer Aktivität.

3.2 Fächerecholot und ergänzende SingleBeam Messungen

Alle Messungen erfolgten mit einem Fächerlot der neuesten Generation SeaBat T50-P .Dieses System arbeitet mit einer Frequenz von 400kHz. Der Öffnungswinkel variiert zwischen 150° und 165°. Die entsprechende Beamöffnung beträgt quer 0,5° und längs 1°. Die Tiefenauflösung kann mit 6mm angenommen werden. Die maximale Ping Rate beträgt 50Hz. Als Inertialsystem wird die Messeinheit Octans V mit einer Richtungsgenauigkeit von 0,1° und von 0,01° bei Roll, Pitch und Yaw sowie einer Auflösung von 0,001° verwendet.

Für die exakte Positionsbestimmung wurde ein GNSS Heading System der AsteRx-U Marine von Septentrio zum Einsatz. Die Temperatur und das Wasserschallgeschwindigkeitsprofil wird mit der VALEPORT Swift SVP Sonde gemessen. Die Messfahrten wurden bei idealem Herbstwetter bei extrem ruhiger Seeoberfläche in den ersten beiden Novemberwochen 2017 durchgeführt.



Abb. 3: Spuren der Fächerecholotmessungen

Die Daten aus der Fächerecholotmessung, die flächendeckend bis zu einer Wassertiefe von minimal zwei Metern durchgeführt wurden stellen die Basis für ein hochauslösendes Unterwassergeländemodell des Wörthersees dar. Für die Erfassung der Fachwasserzonen wurden ergänzenden SingleBeam Messungen durchgeführt.



Abb. 4: Spuren der ergänzenden Singlebeam Messungen

Das berechnete Unterwassergeländemodell ist in den Flachwasserzonen war aufgrund der wesentlich geringeren Punktdichte im Vergleich zu den Fächerecholotdaten natürlich ungenauer, eine wesentlicher Verbesserung wird durch die zusätzliche Verwendung des, aus der Drohnenbefliegung berechnete Oberflächenmodell erwartet.



Abb. 5: Unterwassergeländemodell des Wörthersees (Auflösung 2×2m)

3.2.1 Erste Auswertungen

Der Wasserverband Wörthersee Ost betreibt ein weiterverzweigtes Netz an Abwasserleitung im Kärntner Zentralraum. Die exakte Verortung von See-Einbauten unterhalb des Wasserspiegels (Einleitungen, Wasserentnahmen, Seedruckleitungen) wird für die Erarbeitung von Notfallplänen dringend benötigt. Bisher konnte die Lage der bestehenden Seeleitung nur punktuell im Rahmen von Tauchgängen ermittelt werden. Mit Hilfe des hochauflösenden Tiefenmodells konnte der genaue Verlauf der Leitungen ermittelt werden.



Abb. 6: Bestehende Druckleitung in der Wörthersee Ostbucht

3.3 Orthofoto und Oberflächenmodell aus Drohnenbefliegung

Im Rahmen einer Forschungskooperation zwischen dem Land Kärnten und der FH Kärnten -Studiengang Geoinformation und Umwelttechnologien wurde eine hochauflösende photogrammetrische Auswertung des gesamten Wörtherseeufers und ausgewählter Flachwasserbereiche auf Grundlage einer Drohnenbefliegung durchgeführt. Erfasst wurde dabei ein ca. 100m breiter Uferstreifen sowie die anschließende Flachwasserzone. Im Zuge dieses Projektes wurden einerseits Orthofotos (Auflösung 5×5cm) und ein Oberflächenmodell (Auflösung 10×10cm) berechnet. (Eigener Beitrag im Rahmen dieser Tagung).



Abb. 7: Orthofoto Ostufer des Wörthersees



Abb. 8: Oberflächenmodell Ostufer des Wörthersees

3.4 Produkte

Als solche liegen derzeit vor bzw. sind abrufbar:

- Hochauflösendes Tiefenmodenmodell mit nahtlosem Übergang zum bestehenden Höhenmodell
- Auswertungen des Tiefenmodells (Isohypsenkarte, Mächtigkeit der Weichauflage, Sedimenttypen, Grundwasserzutritte, Wracks, historische Einbauten ...)
- Erfassung von See-Einbauten oberhalb und unterhalb des Wasserspiegels (Einleitungen, Wasserentnahmen, Seedruckleitungen, Stege, ...)
- Erfassung von ökologisch wertvollen Flächen im Uferbereich, Change Detection zu bestehenden Daten
- Räumliche Darstellung der Sedimentmächtigkeit

Die Präzision der 3D-Tiefendaten hat mit den verwendeten Messgeräten einen beeindruckenden Level erreicht. Die Aufgabenstellung verschiedene Sensoren zu fusionieren, darf als gelungen bezeichnet werden. Für Nachfolgeprojekte ist sohin ein zukunftsweisender Weg aufgezeigt worden.

4 Literaturverzeichnis

- DÖLLER, H., JACHS, W. & NEUNTEUFEL, G., 2016: New Concepts and Applications for Geodetical Monitoring with Fibre optic Sensoring (FOS). Joint International Symposium on Deformation Monitoring (JISDM), TU Wien. https://www.fig.net/resources/proceedings/ 2016/2016 03 jisdm pdf/nonreviewed/JISDM 2016 submission 106.pdf
- SCHULZ, L., FRESNER, R., AMBROS, M., REICHMANN, M., MAIRITSCH, M. & HOLAREK, C., 2004: Der Wörthersee Limnologische Langzeitentwicklung des Wörthersees und limnologische Untersuchungen des Jahres 1999 unter besonderer Berücksichtigung der Planktonbiocönosen. Publikationen des Kärntner Instituts für Seenforschung, 24, 1-58. https://www.zobodat.at/pdf/Pub-Kaerntner-Institut-fuer-Seenforschung 24 0001-0058.pdf
- SCHULZ, N. & SCHULZ, L. 1976: Neuauslotung des Wörthersees mit Hilfe eines Echographen. Carinthia II, **86**, 463-466. https://www.zobodat.at/pdf/CAR_166_86_0463-0466.pdf

Hochauflösende unterwasserarchäologische Prospektion oberösterreichischer Pfahlbauten und Seen mit Fächerecholot und Sediment-Sonar

IMMO TRINKS¹, WOLFGANG NEUBAUER^{1,2}, TIMOTHY TAYLOR², MARIO WALLNER¹, KLAUS LÖCKER^{1,3} & JUTTA LESKOVAR⁴

Zusammenfassung: Das Projekt "Archäologische Prospektion oberösterreichischer Pfahlbauten" hat zum Ziel mit moderner Sonartechnik den Grund des Attersees und des Mondsees sowie archäologische Ablagerungen und Sedimente im Bereich prähistorischer Pfahlbausiedlungen detailliert zu kartieren. Bisher wurden die Seen und ihr Untergrund nur relativ grob vermessen. Mit einem modernen Fächerecholot ist es möglich die genaue Wassertiefe effizient digital zu vermessen, während ein Mehrkanal-Sedimentsonar Einblicke in den Untergrund liefert. Daraus resultieren detaillierte 3D-Modelle des Seebodens, die seine Struktur sowie darauf befindliche künstliche und natürliche Objekte zeigen. Neben archäologischen Strukturen sind detektierte geologische Formationen für das Verständnis der Entwicklung der Seen von großer Bedeutung.

1 Motivation

In den circum-alpinen Seen liegen bedeutende, zum Weltkulturerbe zählende neolithische und bronzezeitliche Pfahlbauten. In Oberösterreich befinden sich vier Stätten des UNESCO-Welterbes Prähistorische Pfahlbauten um die Alpen. Es handelt sich dabei um die Station See im Mondsee und die Stationen Abtsdorf I, Abtsdorf III und Litzlberg Süd im Attersee (POHL 2016). Im Rahmen der Vorbereitung der geplanten oberösterreichischen Landesausstellung zur Kultur der prähistorischen Pfahlbauten soll das Projekt "Archäologische Prospektion oberösterreichischer Pfahlbauten" durch den Einsatz moderner zerstörungsfreier archäologischer Prospektionsverfahren (MAN-DERS & GREGORY 2015) neue Erkenntnisse über diese speziellen Siedlungsplätze und ihre Umgebung liefern. Während an einigen Fundstellen auch noch über den Seegrund aufragende Reste von 4.000-5.000 Jahre alten Holzpfählen bis heute im flachen Wasser erhalten sind, haben sich an anderen Stellen lediglich die Ablagerungen der prähistorischen Siedlungen erhalten (OFFENBER-GER & RUTTKAY 1997). Das Ludwig Boltzmann Institut für Archäologische Prospektion und Virtuelle Archäologie (LBI ArchPro) hat in Zusammenarbeit mit dem Vienna Institute for Archaeological Science der Universität Wien (VIAS) und der Direktion Kultur des Landes Oberösterreich ein neues Prospektionssystem für die effiziente hochauflösende Unterwasserprospektion aufgebaut, mit dem Ziel den Seegrund und den Seeuntergrund des Mondsees wie auch des Attersees

¹ Ludwig Boltzmann Institute für Archäologische Prospektion und Virtuelle Archäologie, Hohe Warte 38, A-1190 Wien, E-Mail: [immo.trinks, wolfgang.neubauer, mario.wallner]@archpro.lbg.ac.at

² Universität Wien, Vienna Institute for Archaeological Science, Interdisziplinäre Forschungsplattform Archäologie, Franz Klein-Gasse 1/III, A-1190 Wien, E-Mail: timftaylor@gmail.com

³ Zentralanstalt f
ür Meteorologie und Geodynamik, Hohe Warte 38, A-1190 Wien, E-Mail: klaus.loecker@zamg.ac.at

⁴ Oberösterreichisches Landesmuseum, Land Oberösterreich, Museumstraße 14, A-4010 Linz, E-Mail: j.leskovar@landesmuseum.at

detailliert zu vermessen und zu erkunden. Bisher war die Wasserstiefe dieser Seen nur in Form von Tiefenkonturlinien mit 10m bis 5m Tiefenabstand bekannt (Abb. 1) (BEHBEHANI 1987). Die im Projektrahmen geplante hochauflösende Vermessung soll wesentliche neue Erkenntnisse über die archäologischen Siedlungsreste im Wasser wie auch geologische Strukturen am Seegrund und im Seeuntergrund liefern. Das hier präsentierte wissenschaftliche Vorhaben ist am ehesten mit dem Projekt "Tiefenschärfe Bodensee" (WESSELS et al. 2016) zu vergleichen, welches den Bodensee erfolgreich mit einem vergleichbaren Forschungsansatz durch Einsatz moderner Fächerecholot und Sedimentsonar Systeme komplett vermessen hat.



Abb. 1: Bathymetrische Karte des Mondsees nach bisherigem Kenntnisstand basierend auf 10m bis 5m Tiefenlinien

2 Methoden

Moderne Sonartechnik ermöglicht die detaillierte Vermessung und Kartierung von Seeböden sowie die dreidimensionale Untersuchung von Sedimentschichten und Strukturen im Untergrund. Für dieses Projekt wurden modernste Fächerecholot- und Sedimentsonarsysteme mit geeigneter Positionierungstechnologie auf einem Messboot installiert und am Mondsee und Attersee zum Einsatz gebracht.

2.1 Fächerecholot

Im Rahmen des hier vorgestellten Projekts wurde nach Testmessungen mit verschiedenen hochauflösender Fächerecholotsysteme am Mondsee im Mai 2016 von der Universität Wien ein hochauflösendes Teledyne Reson SeaBat T50-P Fächerecholotsystem erworben, welches eine Vermessung des Seebodens mit 512 Strahlen (1°×0,5° Strahlbreite @ 400kHz) mit einem interaktiv justierbaren Fächer ermöglicht und pro Sekunde bis zu 25.600 Punkte und die zugehörige Reflexionsintensität aufzeichnet. Dieses kompakte Fächerecholot-System ist speziell für den Einsatz auf kleineren Booten geeignet und bietet im Frequenzbereich zwischen 190 Hz und 420 Hz ausgesprochen saubere Daten mit sehr hoher Auflösung. Die Strahldichte kann flexibel angepasst werden und Reflexionen im Bereich 0,5m bis 200m (maximal 575m) können detektiert werden. Es ist möglich für jeden Strahl Mehrfachdetektionen sowie die gesamte Wassersäule aufzuzeichnen um komplexe Strukturen detaillierter abbilden zu können. Die angegebene Tiefenauflösung beträgt 6mm. Zur Kalibrierung des Fächerecholotsystems werden regelmäßig sogenannte Patch Tests gefahren und Wasserschallgeschwindigkeitsprofile gemessen um automatisch die Roll-, Nick-, und Richtungswinkel der aktuellen Sensorlage zu berechnen und für die Streifenanpassung anzuwenden.

2.2 Sedimentsonar

Zusätzlich zum Fächerecholot wurde mit Mitteln des Landes Oberösterreich ein innovatives vierkanaliges parametrisches Sedimentsonar SES-2000 quattro erworben, welches von Innomar in Rostock für höchstauflösende dreidimensionale Messungen des Seeuntergrundes in extrem flachem Wasser entwickelt wurde (LOWAG et al. 2010). Der parametrische Sub-Bottom Profiler besteht aus vier Schallgebern und Empfängern die mit nur 25cm Abstand als Linienanordnung montiert werden können um sehr dicht im Flachwasser den Untergrund dreidimensional untersuchen zu können. Es ist prinzipiell möglich die Sender/Empfänger im Single-, Dual-, Triple- oder Quad-Strahlmodus einzusetzen, je nach Aufgabenstellung und benötigter Eindringtiefe, bzw. gewünschtem horizontalem Auflösungsvermögen. So bietet es sich an, bei größeren Wassertiefen eine quadratische Sender/Empfängeranordnung zu wählen und gemeinsam zu pulsen, während im Flachwasser mit einem 1m breiten Sensorbalken quer zur Fahrtrichtung der Untergrund mit nur 25cm Profilabstand vermessen werden kann. Dieser Quad-Strahlmodus lässt sich im Tiefenbereich von 0,5m bis 30m einsetzen.

2.3 Positionierung und Navigation

Die Positionierung der Messdaten erfolgt mit einem Doppelantennen Echtzeit Global Navigation Satelliten System (Septentrio AsteRx-U) mit 6.5m Distanz zwischen Bug- und Heckantenne sowie einem hochgenauen SBG Apogee-E inertialen Navigationssystem (0.008° Rollen und Nicken, 0.02° Richtungswinkel, 9mm Position, 20mm delayed heave). Mittels des Echtzeit Positionierung Dienstes (EPOSA) und des Empfangs von Korrekturdaten über NTRIP ist es möglich die Sonardaten effizient hochgenau zu positionieren. Die Navigation und Aufzeichnung der Fächerecholotdaten erfolgt mit QPS QINSy Software auf einer Workstation mit SSD Festplatte und einer Ge-Force GTX 1080 Grafikkarte. Am Fächerecholotkopf wird die Wasserschallgeschwindigkeit mit einer Valeport UltraSV Sonde gemessen, während vertikale Wasserschallgeschwindigkeitsprofile mittels einer Valeport SWiFT Sonde gezogen werden.



Abb. 2: Lage und Verteilung der Messprofile der Fächerecholotmessungen auf dem Mondsee mit einer Gesamtlänge von 304km

2.4 Messboot

Als Vermessungsboot wurde ein Buster XXL AWC Kabinenmotorboot aus Finnland angeschafft mit 6,35m Länge und nur 35cm Tiefgang, 3,5mm Aluminiumrumpf und 150PS Außenbordmotor, sowie separater Stromversorgung für die Messelektronik, ein Bootstyp wie er auch von mehreren professionellen hydrographischen Vermessungsbüros in Deutschland, Finnland und Estland eingesetzt wird. Eine schwenkbare Halterung wurde für die Sonarsysteme konstruiert und am Bug des Bootes montiert und eingemessen.

3 Ergebnisse

Seit Mai 2018 wurde der Seegrund des Mondsees nahezu vollständig mit Fächerecholot-Messungen kartiert, mit Ausnahme des extrem flachen Küstenstreifens, der aufgrund des im Sommer 2018 herrschenden Niedrigwassers bisher nicht vermessen werden konnte (Abb. 2). Seit Anfang Oktober wurde das Boot an den Attersee verlegt und es wurden erste Messungen durchgeführt. Die erzeugten bathymetrischen Daten sind nicht nur aus archäologischer und historischer Sicht von großem Interesse, sondern liefern auch eine detaillierte Karte des Seebodens in bisher unerreichter Auflösung, die die Morphologie des Sees aufzeigt, zahlreiche geologisch interessante Strukturen wie Felsvorsprünge, Seebodenfazies, Erosionskanäle, Ablagerungen von Massenbewegungen, Rohrleitungen, Wracks sowie Spuren am Seegrund von Ankern, Bojen oder anderen Störungen der lakustrinen Sedimente deutlich abbildet (Abb. 3).



Abb: 3: A: Schleifspuren eines Bojenankers. B: Fischschwarm und dessen Sonarschatten. C: Zwei gesunkene Boote. D: ein Citroen Traction Avant Autowrack. E: Rohrleitung und Topographie. F: Schwoiradienkrater einer Boje mit Bojenkette. G: Pfahlbaustation Mooswinkel mit Pfosten, Steinen und liegenden Holzstämmen. H: Bootswrack. I: Rutschung als 3D Punktwolke (links) und als simulierter Sidescan Plot (rechts oben) und normalisierte Backscatter Energie Darstellung (rechts unten)

Die insgesamt 333 auf dem Mondsee aufgezeichneten Fächerecholot-Datenstreifen, welche eine Gesamtfläche von mehr als 13km² abdecken, werden bearbeitet um ein optimiertes dreidimensionales Gesamtmodell des Seegrundes erzeugen zu können. Hierzu wird in Zusammenarbeit mit der TU Wien die für Airborne Laserscanning Datenbearbeitung und Streifenanpassung entwickelte Software OPALS (Orientation and Processing of Airborne Laser Scanning data) für die Anwendung auf Fächerecholotdaten zu adaptiert (PFEIFER et al. 2014). Die Datenbearbeitung ist aufgrund der Datenmenge vom Mondsee noch nicht abgeschlossen. Es wird derzeit an einem optimierten Bearbeitungsverfahren gearbeitet, damit die Daten des wesentlich größeren Attersees entsprechend effizient bearbeitet werden können. Die dreidimensionalen Sedimentsonardaten werden nach Zeit-Tiefen Konvertierung in horizontale Tiefenscheiben geschnitten, um eine flächenhafte Analyse archäologisch relevanter Strukturen zu ermöglichen.



Abb: 3: Mit dem Innomar SES2000 quattro Sedimentsonar aufgezeichnetes Einzelprofil aus dem Bereich der Pfahlbaufundstätte Station See am Mondsee. Die horizontalen Linien haben einen vertikalen Abstand von 1m. Die über den Seeboden herausragenden jahrtausendealten Holzpfähle erzeugen deutlich erkennbare Reflektionen

4 Zusammenfassung und Ausblick

Die Datenaufnahme an Mondsee und Attersee ist noch nicht abgeschlossen und wird sobald möglich fortgesetzt werde. Der extreme Föhnsturm in der Nacht vom 29. auf den 30. Oktober 2018 hat bedauerlicherweise zum Sinken des Messbootes wie auch der Havarie zahlreicher anderer Boote am Attersee geführt. Es wird derzeit daran gearbeitet ein Ersatzboot zu beschaffen und nach abgeschlossener Befundung der Messgeräte mit der Vermessung der Seen fortzusetzten. Die Daten werden nach erfolgter Bearbeitung publiziert und im Rahmen der geplanten Landesausstellung der Öffentlichkeit anschaulich präsentiert werden. Dieses Projekt wird wesentliche Grundlagendaten für die weitere Erforschung dieser bedeutenden Seen generieren, von großem Nutzen für zukünftige Umweltüberwachungsvorhaben sein und gezielte unterwasserarchäologische Untersuchungen wie auch geologische und limnologische Forschungen in der Zukunft erleichtern.

5 Danksagung

Wir danken Frau Mag. Nicoletta Waechter für die freundliche Genehmigung den Grund des Mondsees untersuchen zu dürfen. Das Institut für Limnologie in Mondsee hat das Projekt ganz wesentlich durch das Bereitstellen einer Boje und Zugang zu seinem Bootshaus unterstützt, wofür dem Institutsleiter Herrn Prof. Martin Hahn wie auch dem Werkstattleiter Herrn Hannes Höllerer herzlichst gedankt sei. Herr Dipl.-Ing. Christian Klug hat sehr großzügig Zugang zur Echtzeit Positionierung Austria (EPOSA) gewährt und dadurch eine effiziente hochgenaue differentielle Satellitenpositionierung und Navigation ermöglicht. DI Johannes Pfeffer war eine große logistische Hilfe bei der Planung der Arbeiten am Mondsee. Herr Rudolf Wilflingseder sei für die Zurverfügungstellung von Bootshaus und Steg gedankt. Nikolaus Studnicka von Riegl Laser Measurement Systems hat freundlicherweise mit dem neuen Riegl VZ-4000i Laser Scanner dreidimensionale Testmessungen des Mondseeufers ermöglicht. Johannes Häupl aus Attersee hat freundlicherweise sein Bootshaus für Messarbeiten am Attersee zur Verfügung gestellt. DI Jörg Brunken, Florian Hämmerling und Alexander Schmidt von MacArtney Germany haben großartige Unterstützung bei den Messarbeiten geleistet. Mag. Heinz Gruber vom Bundesdenkmalamt hat das Projekt wohlwollend unterstützt. Nach der Bootshavarie im Oktober 2018 haben die Wasserrettung Attersee und die Freiwillige Feuerwehr Attersee und Nußdorf das gesunken Messboot unter großem Einsatz gehoben und geborgen. Heimatforscher Gerald Egger war bei der Bootsbergung und Reinigung eine enorme Hilfe. Marco Pajer hat beim Aufbau des Messsystems geholfen. Hannes Schiel, Geert Verhoeven und Lisa Aldrian waren an den Messarbeiten beteiligt.

Das Ludwig Boltzmann Institut für Archäologische Prospektion und Virtuelle Archäologie (archpro.lbg.ac.at) beruht auf einer internationalen Kooperation der Ludwig Boltzmann Gesellschaft (A), des Amtes der Niederösterreichischen Landesregierung (A), der Universität Wien (A), der Technischen Universität Wien (A), der Zentralanstalt für Meteorologie und Geodynamik (A), Airborne Technologies (A), 7reasons (A), des Römisch-Germanisches Zentralmuseum Mainz (D), des Landschaftsverbandes Westfalen-Lippe – Archäologie für Westfalen (D), des Norwegischen Instituts für Kulturgüterforschung – NIKU (N) und von Vestfold fylkeskommune – Kulturarv (N).

6 Literaturverzeichnis

- BEHBEHANI, A.-R. 1987: Sedimentations- und Klimageschichte des Spät- und Postglazials im Bereich der Nördlichen Kalkalpen (Salzkammergutseen, Österreich). Göttinger Arbeiten zur Geologie & Paläontologie, 34, Geologische Institute der Georg-August-Universität Göttingen.
- JANSA, V. 2013: Probleme und Lösungsansätze beim Monitoring unterwasserarchäologischer Fundstellen am Beispiel des UNESCO-Weltkulturerbes Pfahlbauten. Diplomarbeit. Universität Wien.
- LOWAG, J., WUNDERLICH, J. & HUEMBS, P. 2010: Three-dimensional investigation of buried structures with multi-transducer parametric sub-bottom profiler as part of hydrographical applications. Proc. Hydro 2010, Rostock-Warnemünde, Germany, 02-05 November 2010.
- MANDERS, M. & GREGORY, D. 2015: SASMAP Guideline Manual 2. Best Practices for locating, surveying, assessing, monitoring and preserving underwater archaeological sites, 29-49.
- OFFENBERGER, J. & RUTTKAY, E. 1997: Pfahlbauforschung in den österreichischen Salzkammergutseen. Pfahlbauten rund um die Alpen. Archäologie in Deutschland, Sonderheft, 76-86.
- PFEIFER, N., MANDLBURGER, G., OTEPKA, J. & KAREL, W. 2014: OPALS A framework for Airborne Laser Scanning data analysis. Computers, Environment & Urban Systems, 45, 125-136.
- POHL, H. 2016: Drei Jahre unterwasserarchäologisches Monitoring an den österreichischen UNE-SCO-Welterbestätten. Archäologie Österreichs, **27**(1), 29-35.
- WESSELS, M., ANSELMETTI, F.S., BARAN, R., HILBE, M., GESSLER, S. & WINTERSTELLER, P. 2016: Tiefenschärfe – Hochauflösende Vermessung Bodensee. Internationale Gewässerschutzkommission für den Bodensee, Blaue Reihe, 61.

Das Klassenzimmer im Weltraum – Anwendungen zur Erdbeobachtung zwischen Realität und Virtualität

CLAUDIA LINDNER¹, CHRISTINA MÜLLER², ANNETTE ORTWEIN¹, HENRYK HODAM¹, CARSTEN JÜRGENS¹, JOHANNES SCHULTZ¹, FABIAN SELG¹ & ANDREAS RIENOW¹

Zusammenfassung: Das Projekt Columbus Eye / KEPLER ISS hat sich zum Ziel gesetzt, die MINT-Bildung an Schulen mit Bildern aus dem All zu verbessern, und stellt dafür digitale Lehreinheiten bereit. Diese fördern das wissenschafts-propädeutische Lernen und motivieren Schülerinnen und Schüler, sich stärker mit MINT-Themen auseinanderzusetzen und diese als Berufsfelder für sich zu entdecken. Neben den Lernmodulen werden auch Augmented-Reality-App-unterstützte Arbeitsblätter entwickelt und die Lehre mit Virtual Reality erprobt. In Kooperation mit einem Gymnasium in Nordrhein-Westfalen werden die Lehrmodule in einem eigenständigen Fernerkundungs-Wahlpflichtfach evaluiert. Nach Kursabschluss des ersten Jahrgangs erfolgte kürzlich die erste vollständige Evaluation.

1 Einleitung

"Man erkennt, dass die Erde wirklich nur eine Ansammlung aus kosmischem Staub ist, der sich zu einem Felsen verklumpt hat und über dem eine hauchdünne, zerbrechlich wirkende Atmosphäre liegt. Um das zu begreifen, habe ich den Blick aus dem Fenster gebraucht." – so der erste deutsche Kommandant auf der Internationalen Raumstation (ISS), Alexander Gerst. Die Erde im Weltraum schweben zu sehen, ist für Astronauten ein einschneidendes Erlebnis. Es verschiebt den Blick hin zu mehr Nachhaltigkeit im ökologischen wie auch im sozialen Sinne. Ohne tiefgreifende Erkenntnisse in Naturwissenschaft und Technik wäre dieser außergewöhnliche Blick auf unseren Planeten nicht möglich. Und so wird die Notwendigkeit für MINT-Berufe deutlich, die nicht nur die Raumfahrt selbst erst ermöglichen, sondern auch die Grundlage für Erdbeobachtung im Sinne des Umweltschutzes, der Ernährungssicherheit oder der Katastrophenhilfe bilden.

Ziel des Projektes Columbus Eye und seines Nachfolgers "KEPLER ISS – Kompetenzorientiertes, erfahrungsbasiertes, praktisches Lernen mit Erdbeobachtung von der ISS" ist es einerseits, Schülerinnen und Schülern (SuS) den Blick der Astronauten im Rahmen des regulären Unterrichtes zu vermitteln, andererseits, ihnen mit wissenschafts-propädeutischem Lernen die Naturwissenschaften näher zu bringen und ihre Methodenkompetenzen zu stärken. Das Konzept aus Intermedialität, Interdisziplinarität und Interaktivität hat sich im mehrjährigen Versuch bei SuS bewährt. Dabei verschiebt sich der Fokus mehr und mehr von der Nutzung von stationären Computern hin zu mobilen Smartphones, Tablets und neuerdings VR-Brillen. Mit diesen Geräten lässt sich nicht nur die Realität "erweitern", indem statische Bilder beispielsweise mit Videos, 3D-Animationen oder

¹ Ruhr-Universität Bochum, Geographisches Institut, Arbeitsgruppe Geomatik, Universitätsstr. 150, D-44801 Bochum, E-Mail: [claudia.lindner, annette.ortwein, henryk.hodam, carsten.juergens, johannes.schultz, fabian.selg, andreas.rienow]@rub.de

² Gymnasium Siegburg Alleestraße, Alleestraße 2, D-53721 Siegburg, E-Mail: christina.mueller@gymnasium-alleestrasse.de
sonstigen Zusatzinformationen angereichert werden, sondern lassen sich auch virtuelle Erlebnisse wie ein Flug über die Erde als wissenschafts-propädeutische Lerneinheit nutzen.

2 MINT-Nachwuchs: Mangelware

Laut dem MINT-Frühjahrsreport 2018 (IW 2018) betrug die Differenz zwischen offenen Stellen und Arbeitslosen mit entsprechender Qualifikation im April 2018 in Deutschland 314.800 Personen, davon 33% im Experten- bzw. Akademikerbereich. Die reale Differenz, die u.a. durch Mobilitätsfaktoren und sonstige Mismatch-Probleme entsteht, liegt mit 485.600 Personen noch deutlich darüber. Gleichzeitig wählen immer weniger junge Menschen eine MINT-orientierte (Aus-)Bildung, obwohl die persönlichen und gesellschaftlichen Perspektiven gut aussehen. Befristung ist sowohl bei MINT-Akademikern als auch bei MINT-Fachkräften geringer als in sonstigen Fachrichtungen, beide sind häufiger in Vollzeitstellen und leitenden Positionen tätig und erzielen höhere Bruttolöhne als Personen mit vielen vergleichbaren Ausbildungs- und Studienabschlüssen anderer Fachrichtungen (IW 2018).

Gründe für die immer seltenere Wahl von MINT-Berufen Deutscher Jugendlicher werden vom Report (IW 2018) mit Blick auf die PISA-Studie von 2015 (OECD 2018) genannt: MINT-Lehrkräfte fehlen in so großer Anzahl, dass 41% der Schulen eine Beeinträchtigung des Unterrichts spüren, 18% eine starke Beeinträchtigung (IW 2018) und für die Zukunft werden durch eine zusätzliche Nachwuchsproblematik in MINT-Lehrfächern noch größere Lücken erwartet (KLEMM 2015); SuS haben zu wenig Freude an MINT-Fächern und daher auch nicht den Wunsch, später einen MINT-Beruf zu ergreifen; Computernutzung findet in zu geringem und teilweise falschen Ausmaß statt, da entsprechende Unterrichtskonzepte und Weiterbildungen fehlen (IW 2018).

Dementsprechend gibt es mehrere Anknüpfungspunkte, um die Zukunft der MINT-Berufe zu sichern. MINT-Bildung muss jedoch schon lange vorher in der frühkindlichen Ausbildung beginnen, und die Kinder anschließend durch ihre Schullaufbahn begleiten. Hierzu muss Infrastruktur auf allen Schulebenen geschaffen oder ausgebaut werden. Vor allen Dingen sollte ein Naturwissenschaftlicher Unterricht die SuS durch Lebensnähe und Anwendungsorientiertheit motivieren. Naturwissenschaftliche Wettbewerbe und Arbeitsgruppen wirken sich ebenfalls positiv auf die Kompetenzen der SuS aus. In diesem Zusammenhang muss auch die Arbeit mit Computern in den Schulen verbessert werden. Konzepte zur Nutzung, die sich vorteilhaft auf die technischen und Fachkompetenzen auswirken, müssen erstellt und den Lehrerinnen und Lehrern (LuL) vermittelt werden. V.a. Recherche- und Gruppenarbeit am Rechner führen zu Leistungsverbesserungen. Schließlich muss die Berufsorientierung verbessert werden, die SuS über die vielen Vorteile einer MINT-Ausbildung oder eines MINT-Studiums informiert (IW 2018).

3 Erbeobachtung von der ISS als Anknüpfungspunkt

MINT-Schulmaterialien haben die Aufgabe, komplexe und abstrakte Themen anschaulich und spannend zu vermitteln. Das mangelnde Interesse, das im MINT-Report (IW 2018) identifiziert wurde, weist jedoch darauf hin, dass dies nicht immer gelingt. Eine Möglichkeit, das Interesse zu wecken, ist der bis heute beliebte Kinder-Berufswunsch Astronaut, der bei einer Befragung 2018

(APPINIO RESEARCH 2017) unter den zehn beliebtesten Berufen als Kind rangierten. 2% der Mädchen und 4% der Jungen gaben diesen Berufswunsch an. MINT-Berufe sind in den Befragungsergebnissen rar: Abgesehen vom Wunsch "Astronaut" haben unter den beliebten Kinder-Berufswünschen nur die medizinischen Berufe einen MINT-Bezug (APPINIO RESEARCH 2017).

Der Beruf des Astronauten mag vordergründig als nicht realistischer Berufswunsch abgetan werden. Doch er bildet den Einstiegspunkt für die Faszination Raumfahrt, die vor allen Dingen eine Ingenieurstechnische Teamleistung ist und ein hohes Maß an Interdisziplinarität bietet. Ihre zahlreichen Anwendungsfelder reichen von der Mathematik über die Physik, und Informatik, aber auch Biologie, Biotechnologie Geowissenschaften und Astronomie zählen dazu. Für nahezu jedes Thema aus den MINT-Fächern der Schullehrpläne findet sich eine Anwendung in der bemannten Raumfahrt. Hierzu gehört auch die Erdbeobachtung, die sowohl von Astronauten mit Fotos und Videos aus der Cupola, als auch mit mehreren Erdbeobachtungsensoren an der Außenhülle der Internationalen Raumstation ISS erfolgt. Insgesamt wurden und werden an der ISS 52 Erdbeobachtungsexperimente durchgeführt und es kommen in Kürze noch weitere hinzu (NASA 2018).

Wie die Raumfahrt ist auch die Erdbeobachtung, bzw. Fernerkundung, vielseitig, was die benötigten MINT-Kenntnisse und Anwendungsgebiete angeht. Themen aus den Curricula aller MINT-Fächer sind notwendig, um Satellitenbilder aufzubereiten und quantitativen Analysen zu unterziehen. Die Ergebnisse fließen in Untersuchungen aller Bereiche der Geowissenschaften ein und lassen sich daher in vielen verschiedenen Lehrthemen einsetzen, vom Kartenlesen in der 5. Klasse bis zur Landwirtschaft nach Klima- und Vegetationszonen am Ende der Qualifikationsphase (MSB NRW 2014). Gleichzeitig sind Satellitenbilder hervorragend geeignet, um Themen gezielt und aktuell zu visualisieren und stellen somit ein sinnvolles Lehrmaterial dar. 13 von 16 der Deutschen Bundesländer haben daher die Benutzung von Satellitenbildern in ihren Lehrplänen verankert (SIEGMUND 2011). Die qualitative und quantitative Analyse der Bilder lehrt die SuS räumliches Denken und Entscheiden sowie Verantwortungsbewusstsein für unseren Planeten. Zudem fördert das empirische Arbeiten mit Satellitendaten die Methodenkompetenz der SuS. In einer zunehmend digitalisierten Welt stellen Fähigkeiten in der Verarbeitung von digitalen Fernerkundungsdaten zudem eine wichtige Qualifikation dar (VOß et al. 2007).

Das Projekt Columbus Eye / KEPLER ISS nutzt Erdbeobachtungsdaten von der ISS um die Faszination für bemannte Raumfahrt mit dem Kompetenzerwerb für MINT-Fächer zu verbinden und so die SuS für eine spätere MINT-(Aus-)Bildung zu begeistern und zu stärken (RIENOW et al. 2015a). Zusätzlich nutzt das Projekt "FIS – Fernerkundung in Schulen" Fernerkundung von vielen verschiedenen Satelliten für die gleichen Ziele. Alle drei Projekte stellen kostenlos Lehrmaterialen zur Verfügung, die als eigenständige Schulstunden oder als Ergänzung zum regulären Unterricht eingesetzt werden können (RIENOW et al. 2015b). FIS, Columbus Eye und KEPLER ISS werden bzw. wurden vom DLR aufgrund eines Beschlusses des Deutschen Bundestages gefördert (Förderkennzeichen 50EE1703, 50JR1307 und 50JR1701).

3.1 Sensoren

Von den 52 vergangenen und laufenden Erdbeobachtungsexperimenten auf der ISS wurden 5 für die Erarbeitung von Lehrmaterialien ausgewählt. Diese stellen ihre Daten frei zur Verfügung, die Daten sind in weit verbreiteten GIS-Programmen zu verarbeiten und nutzbar für Themen aus den Lehrplänen der Deutschen Bundesländer.

Das Columbus-Eye-Projekt startete in 2014 mit den Daten des NASA "High Definition Earth Viewing"-Experiments (HDEV), welches im März 2014 an der ISS angebracht wurde und noch immer andauert. Vier HD-Kameras, die nicht speziell für die Bedingungen im All angepasst, sondern im Handel erhältlich sind, wurden am Columbus External Payload Adapter des europäischen Columbus-Moduls angebracht, um zu testen, wie lange diese unter den Extrembedingungen durchhalten würden (MURI et al. 2017; RUNCO 2011). Die Ergebnisse des Experiments stehen noch aus, da die Kameras weiterhin klare und verwertbare Bilder liefern und die Anzahl der ausgefallenen Pixel im einstelligen Bereich liegt (Stand: Dezember 2018). Die Kameras sind fest auf ihrem Modul angebracht, sodass in der Theorie immer eine Kamera in Flugrichtung nach vorn, eine nach unten (Nadir) und zwei nach hinten gerichtet sind. In der Realität wird die ISS jedoch regelmäßig gedreht, bspw. für Andockmanöver der Sojus-Kapseln, sodass die Kameras gelegentlich von ihren gedachten Sichtrichtungen abweichen. Auch die Einstellungen der Kameras sind fixiert, sodass das Bild weder vergrößert noch die Helligkeit verändert werden kann. Nacht- oder Sternenaufnahmen sind daher nicht möglich. Die Kameras werden immer im Wechsel betrieben und senden in einem kontinuierlichen Zyklus einen Live-Stream. Das Team von Columbus Eye / KEPLER ISS hat die Erlaubnis, in den Kamera-Zyklus einzugreifen, um beispielsweise bestimmte Phänomene wie Vulkanausbrüche, Überflutungen, Waldbrände oder Hurrikane zu verfolgen. Die Videos werden exklusiv auf den Servern des Projektes gespeichert, aktuell liegen 52 TB vor (Stand: Dezember 2018). Die Aufbereitung einer Video-Auswahl findet in Matlab© statt (SCHULTZ et al. 2018). Eine Auswahl von Videos befindet sich auf der Columbus-Eye-Website unter columbuseye.unibonn.de/highlights/.

Nachtvideos entstehen im japanischen Projekt METEOR, das ebenfalls eine handelsübliche HD-Kamera an der ISS befestigt hat, welche seit März 2016 läuft und noch aktiv ist. Anders als bei HDEV wird hier kein kontinuierlicher Stream an die Erde gesendet, sondern auf der Station mit einem Algorithmus für das Projekt relevante Video-Ausschnitte identifiziert und nur diese übertragen. Dies hängt mit der niedrigen Datenübertragungsrate von der ISS zusammen. Ziel des Projektes ist die Suche und chemische Analyse von Meteoriten (PERC O.J.). Eines der Videos wird eingesetzt, um die Problematik der Lichtverschmutzung sowie des Energieverbrauchs in Mitteleuropa zu vermitteln, und wurde zu diesem Zweck in zwei Apps zu diesen Themen eingebaut (Co-LUMBUS EYE; FIS).

Der Hyperspektralsensor HICO (Hyperspectral Imager for the Coastal Ocean, gestartet Semptember 2009, inaktiv seit September 2014) nahm insgesamt mehr als 10.000 hyperspektrale Bilder von Küstengebieten, Seen, Flüssen und Ästuaren auf, die der Untersuchung der Klarheit von Küstengewässern, Bodentypen, Bathymetrie und Küstenvegetation dienten (CORSON et al. 2008). HICO-Aufnahmen vom Eerie-See werden in einer App-gestützten Lehreinheit eingesetzt, die den SuS Wasserverschmutzung und Algenblüte näherbringen soll. Auch Daten des neuen ISS-gebundenen Hyperspektralsensors DESIS (DLR Earth Sensing Imaging Spectrometer) sollen eingesetzt werden, sobald eine Auswahl frei zur Verfügung steht.

ISS RapidScat (gestartet September 2014, inaktiv seit August 2016) nutzte ein Scatterometer, um oberflächennahe Windgeschwindigkeiten und -Richtungen über dem Meer zu messen. Die Daten werden in einer App-gestützten Lehreinheit über die Westwindzone, Winterstürme und ihr Potential zur Windenergieerzeugung eingesetzt, das sich auch mit verschiedenen Atmosphärenmessungen auseinandersetzen soll.

Das LiDAR-Instrument von CATS (Cloud-Aerosol Transport System, gestartet Januar 2015, inaktiv seit Oktober 2017) erstellte Profile der Atmosphäre unterhalb der ISS, um Aerosole und Wolken zu analysieren. Wöchentliche Highlights mit Erklärungen wurden vom CATS-Team bereits während des Projektes erstellt (MICGILL et al. 2015) und werden von KEPLER ISS als Basis für Lehreinheiten genutzt, beispielsweise um Natürliche und Anthropogene Ursachen von Luftverschmutzung und ihre Auswirkungen zu diskutieren.

Am 9. Juli 2018 wurde mit dem ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) ein vielversprechender multispektraler Erdbeobachtungssensor auf der ISS in Betrieb genommen. Zusammen mit drei weiteren Instrumenten – Global Ecosystem Dynamics Investigation (GEDI), dem Orbiting Carbon Observatory 3 (OCO-3) und der Hysperspectral Imager Suite (HISUI) bildet er ein ISS-basierter Erdbeobachtungskomplex, der für mindestens 12 Monate synchronisierte Beobachtungen in verschiedenen spektralen, zeitlichen und räumlichen Auflösungen zur Analyse von Kompositionen, Strukturen und Funktionen von Ökosystemen liefern wird (STAVROS et al. 2017). In diesem Zusammenhang wurde bereits am 8.2.2018 bekanntgegeben, dass die Firma ,Airbus Defence & Space' im kommenden Jahr die Plattform ,Bartolomeo' am Columbus-Labor der ISS befestigen wird, auf der ebenfalls Erdbeobachtungsexperimente durchgeführt werden sollen.

3.2 Einbindung in den Schulunterricht

Neben den Sachkompetenzen der zahlreichen, mit Satellitenbildern vermittelbaren Schulinhalte und den Methodenkompetenzen der Verarbeitung und Analyse dieser Bilder werden die Lehrmodule auch unter Berücksichtigung der Urteils- und Handlungskompetenzen der MINT-Fächer gestaltet. Interaktivität, sowohl mit technischen Hilfsmitteln als auch zwischen den SuS, ist ein Kernelement der Lehreinheiten. Durch die Verbindung von mathematischen, physikalischen und chemischen Grundlagen mit geographischen und biologischen Anwendungsgebieten in digitalen Umgebungen sind die Lehreinheiten grundsätzlich interdisziplinär. Die verschiedenen Medien, die von klassischen Arbeitsblättern über Videos hin zu kleinen GIS-Anwendungen und Augmented-Reality-Apps reichen, sorgen zudem für eine intermediale Ausbildung der SuS. Interaktivität, Interdisziplinarität und Intermedialität werden in einem moderatkonstruktivistischen, wissenschafts-propädeutischen Ansatz miteinander verbunden und fördern so die verschiedenen Kompetenzbereiche (VOB et al. 2011; ORTWEIN et al. 2016). Hierzu gehören auch Soft Skills, wie die Fähigkeit zum selbstständigen Lernen und Problemlösen, die Anwendung wissenschaftlichen Arbeitens, sowie räumliche Orientierung, Entscheidungsfindung und Verantwortungsbewusstsein (SCHULTZ et al. 2017).

Viele der in den Projekten erstellten Lehrmodule stützen sich auf kleine GIS-Anwendungen, in denen SuS zwischen ein und drei typischen GIS-Methoden auf eingebettete Satellitenbilder anwenden können (LINDNER et al. 2018). Dies erfordert die Arbeit mit Schul-Computern. Die Ausstattung vieler Schulen ist in dieser Hinsicht jedoch mangelhaft. Viele Klassen müssen sich die Nutzung weniger Rechner(-räume) teilen, die Rechner selbst sind oft veraltet und das Internet, wenn es funktioniert, ist langsam (INITIATIVE D21 2016). Dementsprechend sind die digitalen Lehrmodule der Projekte sowohl in ihrer Datengröße als auch in ihrem Ressourcenverbrauch so klein wie möglich gehalten und wurden programmiert, um auch auf veralteten Rechnern laufen zu können. Die Anwendung auf aktueller IT-Infrastruktur gestaltet sich dadurch jedoch schwierig.

Der Fokus der Projekte verschiebt sich daher langsam von Elektronischem Lernen (E-Learning) hin zu Mobilem Lernen (M-Learning), das HTML5 und Javascript auf Websiten für die Nutzung auf Desktop-PCs genauso wie auf mobilen Geräten einsetzt, sowie Smartphone-Apps, die klassische Arbeitsblätter mit Augmented Reality bereichern und auf den Smartphones der SuS laufen, wodurch der Stand der Schul-IT-Infrastruktur für den Einsatz der Lehreinheiten irrelevant wird.

Insgesamt stehen auf der Columbus-Eye-Website 9 Lehreinheiten zur Verfügung, davon 5 für Geographie, 3 für Physik und 1 für Mathematik. Weiterhin gibt es 4 Observatorien, in denen die SuS annotierte HDEV-Bilder für eine Klassifizierung und somit Kartenerstellung nutzen können (COLUMBUS EYE o.J.). Auf der FIS-Website stehen 23 Lehreinheiten zur Verfügung, davon 12 für Geographie, 4 für Biologie, 3 für Physik, 2 für Mathematik und 1 für Informatik (FIS o.J.). LuLfortbildungen finden regelmäßig statt und werden kostenfrei auf Anfrage angeboten.

3.3 Augmented & Virtual Reality

E-Learning ist für die Arbeit mit Satellitenbildern unabdingbar. Die digitalen Lernmodule, die bereits seit einigen Jahren in FIS und Columbus Eye eingesetzt werden, benutzen daher eigens programmierte kleine GIS-Anwendungen mit nur wenigen Funktionen, die das Erlernen des Umgangs mit GIS vereinfachen, im Gegensatz zu auf Spezialisten ausgerichteter Software, die auch die LuL Zeit zum Erlernen kostet und/oder kostenintensiv ist bzw. erheblich bessere und mehr Rechner erfordert, als Schulen i.d.R. zur Verfügung stehen (LINDNER et al. 2018).

Mit M-Learning wird das Konzept der hohen Verfügbarkeit und kurzer Anlernphase noch vertieft. Gerade im MINT-Bereich, in dem Experimente oft mit hohen Kosten, Platz- und Zeitverbrauch verbunden sind, bietet sich M-Learning für Experimente an, die dadurch jederzeit, überall und kostengünstig durchgeführt werden können. So kann auch das Lernen selbst mobil und noch interaktiver, als beim E-Learning, gestaltet werden. Augmented Reality stellt einen wichtigen Teil des M-Learnings dar, in dem sie die reale Umgebung mit zusätzlichen Informationen durch Abbildungen, Videos, Karten und vielem mehr erweitert (DUNLEAVY et al. 2009). Derartige Apps werden seit einigen Jahren zu Bildungs- und Forschungszwecken eingesetzt, besonders, seit die notwendigen Geräte wie Smartphones und Tablets für die breite Masse der Bevölkerung zur Verfügung stehen (WOJCIECHOWSKI & CELLARY 2013). MINT-AR-Apps sind besonders beliebt, da sie eine positive Auswirkung auf das Lernverhalten der SuS haben. Sie bleiben länger auf den Unterricht fokussiert und interessieren sich stärker für die vorgestellten Themen. Darüber hinaus lernen sie effizienter, trainieren ihre Wahrnehmung, Aufmerksamkeit, Selbstvertrauen und soziale Interaktion (LI et al. 2017).

Die AR-Apps werden in Unity 5.6 mit der Vuforia-Erweiterung 6.2 erstellt, wodurch sie für Android- und iOS-Smartphones anwendbar werden (VUFORIA 2016). Aktuell stehen die Apps nur für Android zur Verfügung, die Implementierung für iOS wird in Kürze erfolgen.

3.3.1 Im Auge des Sturms

Sturmsysteme, insbesondere tropische Zyklone, sind in vielen deutschen Lehrplänen einsetzbar. Das Arbeitsblatt erläutert den SuS den Aufbau eines tropischen Zyklons. Hierzu wird der Aufbau von Wetterkarten, insbesondere Isobaren, erläutert und verschiedene Beispiele gegeben. Mit Hilfe der Augmented Reality wird aus einer Wetterkarte eine Animation und aus einem Schaubild ein Video des Taifuns von oben, aufgenommen von der ISS mit dem HDEV-Experiment.



Abb. 1: Ansicht der AR-App "Im Auge des Sturms". Die Schaubilder werden auf dem Bildschirm zu Animationen und Videos.



Abb. 2: Ansicht der AR-App "Aralkum". Die Umrisse des Aralsees zwischen 2000 und 2016 werden animiert über ein Landsat-Bild von 2001 projiziert.

3.3.2 Aralkum – Vom See zur Wüste

Der Aralsee ist schon seit Jahrzehnten das Paradebeispiel von negativem menschlichem Einfluss auf die Umwelt und von Desertifikationsprozessen. Es findet Anwendung in den Lehrplänen vieler deutscher Bundesländer. Im App-gestützten Arbeitsblatt "Aralkum – Vom See zur Wüste" lernen die SuS nicht nur etwas über die Hintergründe der Entstehung der Aralkum (-kum bedeutet Wüste), sondern auch über die aktuelle Entwicklung. Unterstützt werden sie dabei von einem HDEV-Video, das einen Überflug von September 2016 zeigt, von einer Animation von Landsat- und Sentinel-Bildern des Sees von 2000 bis 2016 sowie einer Animation der Küstenlinien im gleichen Zeitraum, die über eine Aufnahme von Sentinel-2 gelegt wird und anhand derer sie die Seefläche abschätzen sollen.

3.3.3 Erde bei Nacht – Lichtverschmutzung in Mitteleuropa

Die Metropolregion Rhein-Ruhr und das Land Belgien haben viele Gemeinsamkeiten, was Industriegeschichte und heutige Demographie angeht – und ihre Lichtverschmutzung. Der Großraum zählt mit zu den hellsten der Welt (STARE o.J.). Mit Hilfe eines Videos des japanischen METEOR-Projektes, das von London bis nach Frankfurt a.M. reicht und das mit dem Druck eines Virtual Buttons mit Orientierungspunkten versehen werden kann, wird die Lichtverschmutzung im Raum anschaulich vermittelt. Eine Lichtverschmutzungskarte basierend auf Daten des Sensors VIIRS von 2017 zeigt das Problem auf quantitative Weise. C. Lindner, C. Müller, A. Ortwein, H. Hodam, C. Jürgens, J. Schultz, F. Selg & A. Rienow



Abb. 3: Ansicht der AR-App "Erde bei Nacht". Über eine Sentinel-2-Szene des Ruhrgebiets wird ein METEOR-Video gelegt, an das über das Verdecken eines bestimmten Bildteils Städtenamen zur Orientierung gelegt werden.

3.3.4 Das Erde-Mond-System

Als einzige Physik-App beschäftigt sich diese mit Gravitation und Kepler'schen Gesetzen. SuS können anhand des Beispielraums Norddeutsche Bucht experimentieren, wie sich der Tidenhub bei verschiedenen Abständen des Mondes von der Erde auswirken würde. Anhand eines SRTM-Höhenmodells werden Hoch- und Niedrigwasser bei Halbmond simuliert und wahlweise über einem Sentinel-2-Bild oder dem Höhenmodell dargestellt. Durch einen virtuellen Mond, dessen Position relativ zur Kamera berechnet wird, lässt sich nicht nur die Mondrückseite beobachten, sondern auch der Schatten, den dieser auf die Erde wirft, und so kann die Frage nach der Seltenheit von Sonnen- und Mondfinsternissen selbst beantwortet werden.



Abb. 4: Ansicht der AR-App "Erde-Mond-System". Das Markerbild wird zur Erde, das Smartphone zum Mond. Informationen über Entfernungen zwischen den beiden Körpern und die Auswirkung dieser Entfernung können eingeblendet werden. Bei Überschreiten der Roche-Grenze, an der die Differenz der Gravitationskräfte zwischen der erdzu- und erdabgewandten Seite des Mondes wird dieser in der App von den Gravitationskräften zerrissen. Zwei Videoausschnitte aus HDEV, die den Mond über der Erde beim auf- und untergehen zeigen, verdeutlichen die große Distanz zwischen den beiden Körpern. Eine 3D-Animation zur Darstellung des Baryzentrums, anhand derer nicht nur Gravitation im Erde-Mond-System, sondern auch im Sonnensystem und Methoden zur Exoplanetensuche diskutiert werden, kann ebenfalls eingeblendet werden.

3.3.5 Virtual Reality

Virtual-Reality-Equipment ist dank des Wettbewerbs führender Hersteller auch für Privatpersonen erschwinglich geworden. Die Auswirkungen dieser Potentiale auf den Lernerfolg sowie die Einbindung der Technik für den Schulunterricht sind jedoch bislang noch vergleichsweise wenig erforscht. Erste Übersichtsarbeiten der "neuen" VR mit einem Field of View von 100°+, konstatieren einen Gewinn für den Unterricht im Hinblick auf das Erleben schwer erreichbarer Orte oder Phänomene sowie die Durchführung von Experimenten, die in der realen Welt als gefährlich oder unethisch eingestuft werden (JENSEN & KONRADSEN 2018; OTT & FREINA 2015; MIKROPOULOS & NATSIS 2011). Entlegene Gebiete wie z.B. der Mount Fuji können mit Hilfe von naturgetreuen 3D-Modellen in die virtuelle Realität überführt und mit Animationen (etwa Lawinen, Bergstürze etc.) angereichert werden, die erst bei Kontakt ausgelöst werden. Für die Lernenden bedeutet das realistische Erleben eine größere Identifikation mit der Situation als z.B. in computerbasierten Settings, unter anderem bleibt der Sinn für gefährliche Situationen und die natürliche Reaktion auf diese erhalten (REINERS et al. 2014). Vielen bestehenden VR-Anwendungen ist gemein, dass sie entweder nur einen einzigen Nutzer im Fokus haben, oder einen ganzen Klassenraum im "Second Life" als Kommunikationsplattform in die virtuelle Realität überführen. Für die aktuelle Situation in Schulen sind beide Konzepte wenig geeignet, da sie entweder eine sehr gute technische Ausstattung benötigen oder das Lernen im Klassenverband vernachlässigen. Somit ist die Herausforderung, eine immersive Virtual-Reality-Umgebung zu entwickeln, in der SuS dennoch mit der Außenwelt in Kontakt bleiben. Die SuS kommunizieren mit dem Expeditionsleiter außerhalb der VR, welche die Expeditionsteilnehmer mit Hilfe von analogem Kartenmaterial zu den zu kartierenden Arealen lotsen. Hier können die Teilnehmer helfen, die weißen Flecken auf der Landkarte zu erforschen und somit eine vollständige Karte des gesamten Mount Fuji zu erstellen. In der VR stehen ihnen dazu ein Lasermessgerät sowie ein Fernauslöser für Lawinen zur Verfügung.

4 Fernerkundung als Schulfach

Unter den zahlreichen Partnerschulen, mit denen Columbus Eye/KEPLER ISS und FIS zusammenarbeiten, findet eine besonders enge Kooperation mit dem Gymnasium Siegburg Alleestraße (GSA) statt. Hier werden die Lehrmodule von LuL und SuS getestet, evaluiert und mit Verbesserungsvorschlägen versehen. An der Schule lernen mehr als 1000 SuS der Jahrgangsstufen 5 bis 12. Seit Oktober 2009 werden die Lehrmodule von FIS regulär im Geographieunterricht eingesetzt. Zusätzlich finden Expertenbesuche und von Experten begleitete Exkursionen statt, die den SuS erste Kontakte zu ESA, DLR und NASA einbringen, beispielsweise bei einem Live-Call mit Astronaut Alexander Gerst an Bord der ISS. Hauptziel der Kooperation ist die Evaluation und Überarbeitung der Lehreinheiten der Projekte. Diese sollen in ganz Deutschland und darüber hinaus in die Lehrpläne der Bundesländer passen und gut im Unterricht anwendbar sein. Da alle Projektmitarbeiter eine rein wissenschaftliche akademische Ausbildung ohne Pädagogikanteil abgeschlossen haben und die Themen dementsprechend konzipieren, ist der praktische, pädagogische Ansatz der LuL und die Bewertung der SuSunabdingbar für die Anwendbarkeit der Lehreinheiten. Sowohl die Projekte, als auch die LuL und SuS profitieren von der Zusammenarbeit.

Seit dem Schuljahr 2016/17 wird am GSA das MINT-Differenzierungsfach (Wahlpflichtfach) "Geographie-Physik" angeboten, das sich an die 8. und 9. Jahrgangsstufe richtet und sich in je zwei Halbjahren mit den physikalischen Grundlagen und den geographischen Anwendungsgebieten der Fernerkundung auseinandersetzt. Um die zwei Teilbereiche abzudecken, arbeitet je eine Lehrkraft der Geographie und der Physik eng zusammen, was die Interdisziplinarität des Unterrichts, aber auch der LuL-Fachschaften stärkt.

Zusätzlich zu den Geographie-Physik Lehrkräften werden am GSA regelmäßig LuL-fortbildungen für alle interessierten Lehrkräfte der Fächer Geographie und Physik angeboten. Trotz anfänglicher Skepsis bei einigen Fach-LuL hinsichtlich der Komplexität des Unterrichts, sowie seiner Vorbereitung mit den digitalen Lernmaterialien, konnte durch die Fortbildungen der Kooperationspartner entgegengewirkt werden. Die Bereitschaft der Kolleginnen und Kollegen für den Einsatz als unterrichtende Lehrkraft des neuen MINT-Differenzierungsfaches konnte durch die regelmäßigen Fortbildungen am GSA insgesamt gesteigert werden. Besonders, wenn nur wenig Zeit zur Vorbereitung einer Stunde zur Verfügung steht, sind die digitalen Lehrmodule aufgrund ihrer intuitiven Anwendbarkeit beliebt.

Bei den SuS ist das MINT-Differenzierungsfach beliebt: Im Jahr 2016 starteten 8 Mädchen und 17 Jungen das Fach, 2017 10 Mädchen und 6 Jungen, im Jahr 2018 13 Mädchen und 10 Jungen.

4.1 Vorwissen und Erwartungen der SuS

Die nationalen Bildungsstandards fordern die Nutzung von Satellitenbildern im Unterricht bereits seit zehn Jahren (DGFG 2008; DGFG 2014), doch durch die Befragung der SuS der ersten drei Wahljahrgänge des Kurses "Geographie-Physik" zeigt, dass dies noch nicht überall umgesetzt wurde: Die SuS beantworten zu Beginn des Kurses in Klasse 8 einige Fragen zu ihrem Vorwissen und ihren Erwartungen. Von 62 SuS aus drei Jahrgängen hatten nur 23 (37%) vorher schon einmal mit Fernerkundungsmethoden gearbeitet, davon jedoch 12 aus dem Jahrgang 2018. 36 der 62 (58%) hatten dagegen schon privat Satellitendaten benutzt, in der Form von Google Earth bzw. Maps oder auch Navigationsgeräten. Diese Zahlen bestätigen den Eindruck, dass zwischen der Alltagswelt der SuS hinsichtlich des Gebrauchs von Satellitenbildern und der Realität im Schulunterricht ein Defizit vorliegt. Die konkrete Realisierung im Schulalltag stellt sich dementsprechend – trotz theoretischer Ambitionen – bis heute nur punktuell ein (STÜMPER 2011; FUCHSGRUBER et al. 2017).

In einer Befragung bezüglich der Erwartungen antworteten fast die Hälfte aller SuS, dass sie Unterricht mit vielen Experimenten und Exkursionen mit dem Koorperationspartner Universität Bonn, aber auch dem DLR erwarteten. Ebenfalls mehrfach genannt wurden die Erwartungen an Expertenbesuche.

4.2 Satelliten- und ISS-Bilder zur Motivation und Verständnisförderung

Die SuS des ersten Jahrgangs "Geographie-Physik" von 2016 haben den Kurs im Jahr 2018 abgeschlossen und nicht nur die einzelnen Module, sondern auch die Halbjahre in anonymen Fragebögen evaluiert. Die Fragebögen sind an das in Deutschland übliche Benotungssystem von 1 (sehr gut) bis 6 (schlecht) angelehnt und fragen i.d.R. nach Zustimmung zu bestimmten Aussagen, die dann von 1 (trifft völlig zu) bis 6 (trifft überhaupt nicht zu) benotet werden.

Die erste Testgruppe bestätigt, dass Satellitenbilder und Live-Bilder und Videos von der ISS nicht nur interessant (Durchschnittsnoten 2,25 und 2,08, respektive), sondern auch für das gesamte Verständnis hilfreich sind (Durchschnittsnoten 1,88 und 1,84, respektive). Unter der Annahme, dass die SuS sich bei der Betrachtung der Live-Videos in die Perspektive eines auf der ISS lebenden Astronauten hinein versetzen können, lassen die Evaluationsergebnisse auf eine gewisse Faszination sowie dadurch gesteigerte Lernmotivation schließen.

4.3 Auf SuS-Bedürfnisse angepasste GIS-Tools

"Um die Möglichkeiten der Fernerkundung zur Förderung einer naturwissenschaftlichen und geographischen Handlungskompetenz gewinnbringend in den Unterricht einbringen und damit auch den Forderungen der Bildungspläne Rechnung tragen zu können, bedarf es dem entsprechend geeignete Lernumgebungen und SuSgerechte Softwarelösungen" (DITTER 2013). Auch wenn bereits seit einiger Zeit ein vielfältiges Angebot entsprechender Software zur Verfügung steht, sind diese Applikationen für die Verwendung im Unterricht in der Regel zu komplex (DITTER 2013). Der Einsatz der Bildbearbeitungssoftware LeoWorks 4 im Wahlpflichtunterricht der Jahrgangsstufe 9 zum Themenkomplex Klimawandel im Teilbereich "Geographie" spiegelt dieses Problem wider. Fast die Hälfte der SuS der Testgruppe (11 SuS) hatte deutliche Probleme bei der Bearbeitung von Satellitenbildern mit der Bildbearbeitungssoftware, die sich eher an die Ausbildung von GIS-Fachkräften richtet (Zustimmungsnoten 5 und 6, Frage: Die Arbeit mit der Bildbearbeitungssoftware LeoWorks4 war einfach..., Durchschnitt: 3,4, n = 24).

Dementsprechend wichtig ist es, die digitalen Fernerkundungsdaten für den Einsatz im Schulunterricht in ein didaktisches Gesamtkonzept zu bringen. Durch den Einsatz der multimedialen Lernmodule und der interaktive Lernplattform durch Recherche- und Analysetools der Projekte FIS und ColumbusEye/KEPLER ISS in den MINT-Unterricht am GSA ist es möglich, die Fülle an Informationen, die ein multispektrales Satellitenbild bietet, der entsprechenden Zielgruppe anschaulich verständlich zu machen. Im Vergleich zur Arbeit mit der Software LeoWorks 4, die vielmehr für professionelle Nutzer gedacht ist, wird aus der Evaluation deutlich, dass die digitalen FIS-Lernmodule SuS-gerecht aufbereitet sind. Zustimmungsnoten wurden in der Testgruppe (n = 24) für Interesse (Durchschnitt: 2,52), Logischen Aufbau (Durchschnitt: 1,83) und verständliche Aufgabenstellung und Anweisungen (Durchschnitt: 2,08) vergeben.

5 Fazit & Ausblick

Die Erde im bzw. aus dem Weltall zu betrachten, kann unsere Sichtweise auf die Geschenisse auf der Erde nachhaltig ändern. Diese Änderung auch bei Schulkindern im Rahmen des normalen Unterrichts hervorzurufen ist einer der Ansätze der Projekte Columbus Eye / KEPLER ISS und Fernerkundung in Schulen. Mit Satellitenbildern und ISS-Videos werden SuS motiviert, sich

stärker mit den MINT-Fächern auseinanderzusetzen, dabei fachliche und methodische wissenschaftliche Kompetenzen zu erlangen und weitere Soft Skills zu trainieren. Dies soll ihnen den Weg zu einer MINT-Bildung oder –Ausbildung eröffnen, wo sie hervorragende Chancen auf Arbeitsplätze mit guten Konditionen vorfinden, aber auch langfristig den Arbeitsmarkt stärken, auf dem ein Mangel an MINT-Arbeitskräften aller Ausbildungsstufen herrscht.

Um dies umzusetzen, wird stark auf E-Learning gesetzt, das in der Fernerkundung zur quantitativen Analyse von Satellitendaten unabdingbar ist, und auch die methodischen Kompetenzen der SuS im Umgang mit digitalen Medien stärkt. Die Weiterentwicklung von E-Learning stell das M-Learning dar, das die oft mangelhafte IT-Infrastruktur an Schulen umgeht, in dem es die hohe Verfügbarkeit von Smartphones bei SuS nutzt. Hierzu werden Apps mit Augmented Reality erstellt, die klassische Arbeitsblätter auf Papier mit Videos, Animationen und kleinen Experimenten erweitern. Zusätzlich wird aus ISS- und weiteren Fernerkundungsdaten eine erste Virtual-Reality-Anwendung entwickelt.

Das neue Fach "Geographie-Physik" in Zusammenarbeit mit den universitären Arbeitsgruppen erweitert durch die Einbindung von Fernerkundungsmethoden und Live-Bildern von der ISS den MINT-Fächerkanon im Differenzierungsbereich, sowie die MINT-Schulenwicklung des GSA und ermöglicht so eine facherverbindende Zusammenarbeit zwischen Kollegen der Fachschaften Geographie und Physik. Darüber hinaus kann durch die ausgearbeiteten digitalen Unterrichtsmaterialien, sowie einem entsprechenden Lernportal die didaktische Reduzierung hochkomplexer und aktueller Datensätze für den Schulunterricht im Bereich der Fernerkundung erfolgen. Die Einbindung der entwickelten Unterrichtsmaterialien der universitären Projekte, die über eine visuelle Deutung von Echtfarbenbildern hinausgehen (GOETZKE et al. 2014; RIENOW et al. 2015a), erleichtert es den Lehrkräften, motivierenden und gleichzeitig fächerverbindenen Unterricht zu gestalten, ohne dass die persönlichen Ressourcen bei der Unterrichtsvorbereitung voll ausgeschöpft werden müssen. Die Kooperation mit außerschulischen Partner in Verbindung mit dem neuen MINT-Fach ermöglicht dabei nicht nur die Chance, das Interesse und die Motivation der SuS für die Fächer Geographie und Physik zu wecken, sondern auch eine praxisnahe Qualifikation für die MINT-Berufswelt anzubahnen.

In Zukunft wird die Evaluation der Lehreinheiten inklusive der neuen AR- und VR-Anwendungen fortgeführt und die mittelfristigen Auswirkungen beobachtet, um zu untersuchen, ob die Projektmaterialien und das Fach Geographie-Physik den Wunsch der heutigen SuS nach einer MINT-Karriere erhöhen.

6 Literaturverzeichnis

APPINIO RESEARCH, 2017: Studie zu Traumberufen: Das sind die Berufswünsche der Kinder von heute. www.appinio.com/de/blog/studie-zu-traumberufenberufsw%C3%BCnsche-der-kinder-von-heute, letzter Zugriff 20.08.2018.

COLUMBUS EYE o.J.: columbuseye.uni-bonn.de/, letzter Zugriff 21.08.2018.

CORSON, M.R., KORWAN, D.R., LUCKE, R.L. & SNYDER, W.A., 2008: The Hyperspectral Imager for the Coastal Ocean (HICO) on the International Space Station. IEEE International Geoscience and Remote Sensing Symposium, 4, IV-101.

- DGFG DEUTSCHE GESELLSCHAFT FÜR GEOGRAPHIE, 2008: Bildungsstandards im Fach Geographie für den Mittleren Schulabschluss, Bonn.
- DGFG DEUTSCHE GESELLSCHAFT FÜR GEOGRAPHIE, 2014: Bildungsstandards im Fach Geographie für den Mittleren Schulabschluss, Bonn.
- DITTER, R., 2013: Die Wirksamkeit digitaler Lernwege in der Fernerkundung. Eine empirische Untersuchung zu Lernmotivation und Selbstkonzept bei Schülerinnen und Schülern der Sekundarstufe. Dissertation an der Universität Heidelberg.
- DUNLEAVY, M., DEDE, C. & MITCHELL R., 2009: Affordances and Limitations of Immersive Participatory Augmented Reality Simulations for Teaching and Learning. Journal of Science Education and Technology, 18, 7-22.
- FIS: FERNERKUNDUNG IN SCHULEN o.J.: fis.uni-bonn.de/, letzter Zugriff 21.08.2018.
- FUCHSGRUBER, V., DITTER, R. & SIEGMUND, A., 2017: Geographieunterricht mit Satellitenbildern innovativ gestalten. Praxis Geographie, **47**(3), 8-9.
- GOETZKE, R., HODAM, H., RIENOW, A. & VOB, K., 2014: Floods. Dealing with a Constant Threat. T. Jekel et al. (Hrsg.): Learning and Teaching with Geomedia, Cambridge. Newcastle, 90-102.
- KLEMM, K., 2015: Lehrerinnen und Lehrer der MINT-Fächer: Zur Bedarfs- und Angebotsentwicklung in den allgemeinbildenden Schulen der Sekundarstufen I und IIam Beispiel Nordrhein-Westfalens. Telekom-Stiftung.
- INITIATIVE D21, 2016: Sonderstudie "Schule Digital" Lehrwelt, Lernwelt, lebenswelt. Digitale Bildung im dreieck SchülerInnen-Eltern-Lehrkräfte. initiatived21.de/app/uploads/2017/01/ d21 schule digital2016.pdf, letzter Zugriff 24.08.2018.
- IW: INSTITUT DER DEUTSCHEN WIRTSCHAFT, 2018: MINT-Frühjahrsreport. www.arbeitgeber.de/www/arbeitgeber.nsf/res/MINTFruehjahrsreport_2018.pdf/\$file/MINTFruehjahrsreport_2018.pdf (Abgerufen am 20.08.2018).
- JENSEN, L. & KONRADSEN, F., 2018: A review of the use of virtual reality head-mounted displays in education and training. Education and Information Technologies, **23**(4), 1515-1529. https://doi.org/10.1007/s10639-017-9676-0 null.
- LI, J., VAN DER SPEK, E.D., FEIJS, L., WANG, F. & HU, J., 2017: Augmented Reality Games for Learning. A Literature Review. In International Conference on Distributed, Ambient, and Pervasive Interactions Springer, Cham, 612-626.
- LINDNER, C., HODAM, H., ORTWEIN, A., SELG, F., SCHULTZ, J. & RIENOW, A., 2018: Sentinel-Daten für digitale und interaktive Anwendungen im Schulunterricht. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 27, 38. Wissenschaftlich-Technische Jahrestagung der DGPF und PFGK18 Tagung in München, 117-128.
- MICGILL, M.J., YORKS, J.E., SCOTT, V.S., KUPCHOCK, A.W. & SELMER, P.A., 2015: The Cloud-Aerosol Transport System (CATS). A Technology Demonstration on the International Space Station. Proc. SPIE. 9612, Lidar Remote Sensing for Environmental Monitoring XV.

- MIKROPOULOS, T. A. & NATSIS, A., 2011: Educational virtual environments. A ten-year review of empirical research (1999–2009). Computers & Education, **56**(3), 769-780.
- MSB NRW: MINISTERIUM FÜR SCHULE UND WEITERBILDUNG DES LANDES NORDRHEIN-WESTFA-LEN, 2014: Kernlehrplan für die Sekundarstufe II Gymnasium/Gesamtschule in Nordrhein-Westfalen. Geographie. <u>https://www.schulentwicklung.nrw.de/lehrplaene/ lehrplan/180/KLP_GOSt_Geographie.pdf</u>, letzter Zugriff 06.12.2018.
- MURI, P., RUNCO, S., FONTANOT, C. & GETTEAU, C., 2017: The High Definition Earth Viewing (HDEV) payload. IEEE Aerospace Conference. Big Sky, MT, 1-7.
- NASA: Space Station Research Explorer o.J. <u>www.nasa.gov/mission_pages/station/research/ ex-periments/explorer/</u>, letzter Zugriff 20.08.2018).
- OECD, 2108: PISA 2015 Results in Focus. www.oecd.org/pisa/pisa-2015-results-in-focus.pdf, letzter Zugriff 23.08.2018.
- ORTWEIN, A., GRAW, V., HEINEMANN, S., MENZ, G., SCHULTZ, J., SELG, F. & RIENOW, A., 2016: Pushed Beyond the Pixel – Interdisciplinary Earth Observation Education from the ISS in Schools. Proceedings of the 67th International Astronautical Congress 2016 26-30 September in Guadalajara, Mexico, 1-6.
- OTT, M. & FREINA, L., 2015: A literature review on immersive virtual reality in education: state of the art and perspectives. Conference proceedings of »eLearning and Software for Education« (eLSE).
- PERC: PLANETARY EXPLORATION RESEARCH CENTER o.J.: ISS Meteor Observation Project "ME-TEOR". www.perc.itchiba.ac.jp/project/meteor/, letzter Zugriff 20.08.2018.
- REINERS, T., WOOD, L. & GREGORY, S., 2014: Experimental study on consumer-technology supported authentic immersion in virtual environments for education and vocational training, Hegarty, B., McDonald, J. & Loke, S. (Hrsg.), Proceedings of the 31st Annual Ascilitite Conference (ascilite 2014): Rhetoric and Reality: Critical perspectives on educational technology, Nov 23-26 2014, Dunedin, 171-181.
- RIENOW, A., HODAM, H., SELG, F. & MENZ, G., 2015a: Columbus Eye: Interactive Earth Observation from the ISS in Class Rooms, GI-Forum, Journal for Geographic Information Science, Wichmann, Berlin, 349-353.
- RIENOW, A., GRAW, V., HEINEMANN, S., SCHULTZ, J., SELG, F. & MENZ, G., 2015b: Inspecting the Blue Dot: Goals, Methods, and Developments of the Project Columbus Eye. 64. Deutscher Luft- und Raumfahrtkongress (DLRK) 2015, September in Rostock.
- RUNCO, S., 2015: International Space Station High Definition Earth Viewing (HDEV). www.nasa.gov/mission_pages/station/research/experime nts/917.html, letzter Zugriff 20.08.2018.
- SCHULTZ, J., LINDNER, C., HODAM, H., ORTWEIN, A., SELG, F., WEPPLER, J. & RIENOW, A., 2017: Augmenting Pupil's Reality from Space – Digital Learning Media based on Earth Observation Data from the ISS. Proceedings of the 68th International Astronautical Congress 2017, 25-29 September in Adelaide, Australia.

- SCHULTZ, J., ORTWEIN, A. & RIENOW, A., 2018: Technical note: using ISS videos in Earth observation implementations for science and education. European Journal of Remote Sensing, 51(1), 28-32.
- SIEGMUND, A., 2011: Satellitenbilder im Unterricht. Eine Ländervergleichsstudie zur Ableitung fernerkundungsdidaktischer Grundsätze. Dissertation an der Pädagogischen Hochschule Heidelberg.
- STARE, J. o.J.: www.lightpollutionmap.info (Abgerufen am 06.12.2018).
- STAVROS, E. N., SCHIMMEL, D., PAVLICK, R., SERBIN, S., SWANN, A., DUNCANSON, L., FISHER, J.B., FASSNACHT, F., USTIN, S., DUBAYAH, R., SCHWEIGER, A. & WENNBERG, P., 2018: ISS observations offer insights into plant function. Nature Ecology & Evolution, 1, 1-5.
- STÜMPER, C., 2011: Möglichkeiten und Grenzen der Umsetzung der Fernerkundung mit Hilfe der Methode "Lernen durch Lehren (LdL)" im bilingualen deutschfranzösischen Erdkundeunterricht der Jahrgangsstufe 11 für einen Projekttag zum Thema "chances et limites de l'application des méthodes de la télédéction pour la prévention des catastrophes naturelles" für die Jahrgangsstufe 9. Schriftliche Hausarbeit im Rahmen der Zweiten Staatsprüfung für das Lehramt an Gymnasien und Gesamtschulen.
- VOB, K., GOETZKE, R., HODAM, H. & RIENOW, A., 2011: Remote Sensing, New Media and Scientific Literacy - A New Integrated Learning Portal for Schools Using Satellite Images. Learning with GI 2011 - Implementing Digital Earth in Education, Berlin, 172-180.
- VOß, K., GOETZKE, R. & THIERFELDT, F., 2007: Integration von Fernerkundung im Schulunterricht. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., 16, 41-50. https://www.dgpf.de/src/pub/DGPF2007.pdf
- VUFORIA, 2016: Developer's Guide. library.vuforia.com/.
- WOJCIECHOWSKI, R. & CELLARY, W., 2013: Evaluation of learners' attitude toward learning in ARIES augmented reality environments. Comput. Educ., **68**, 570-585.

Bildungsprojekte im Fachbereich Photogrammetrie, Fernerkundung und GIS: Herausforderungen, Chancen und Grenzen

REINFRIED MANSBERGER¹, THOMAS BAUER¹, MARKUS IMMITZER¹ & FRANCESCO VUOLO¹

Zusammenfassung: Derzeit ist das Institut für Vermessung, Fernerkundung und Landinformation (IVFL) der Universität für Bodenkultur Wien (BOKU) an sieben Bildungsprojekten mit unterschiedlichen Geldgebern und Partnern beteiligt. Die Aktivitäten in den einzelnen Projekten sind vielfältig und reichen von reiner Lehrtätigkeit über die Entwicklung von Studienprogrammen, der Aufbereitung von Lehrmaterial, der Abhaltung von Summer Schools bis hin zu gemeinsamen Forschungsaktivitäten.

Im Artikel werden die Projekte nach einer kurzen Vorstellung im Hinblick auf Inhalte, gewählte Methodik, Finanzierung und Organisation charakterisiert. Herausforderungen und Chancen, Grenzen und Schwachpunkte bei der Beantragung und Durchführung von Bildungsprojekten werden aufgezeigt.

Abschließend dokumentieren die Autoren ihre gewonnenen Erfahrungen in Form von Empfehlungen für eine erfolgreiche Beantragung und Abwicklung von Bildungsprojekten.

1 Einleitung

Wissensvermittlung und Wissensaustausch sind zentrale Elemente in der akademischen Lehre. Daher werden Bildungsprojekte zur Förderung von akademischen Partnerschaften, zur Steigerung der Lehrenden- und Studierendenmobilität sowie zum Aufbau fachlicher Kompetenzen auf universitärer, nationaler, europäischer und internationaler Ebene angeboten und finanziert. Die Vergabe erfolgt normalerweise in einem kompetitiven Prozess.

Das Institut für Vermessung, Fernerkundung und Landinformation (IVFL) der Universität für Bodenkultur Wien (BOKU) hat sich aufgrund der an der BOKU umfangreich durchzuführenden Lehraktivitäten Expertise im Bereich der Aus- und Weiterbildung angeeignet. Dies manifestiert sich auch in zahlreichen vom IVFL durchgeführten Bildungsprojekten, die inhaltlich, in der Altersstruktur und im Vorwissen der Auszubildenden, in der geographischen Ausrichtung sowie von den Fördergebern sehr weit gestreut sind. Die ersten Mitwirkungen an Bildungsprojekten begannen vor mehr als zwanzig Jahren und die Anzahl an gleichzeitigen Projekten hat sich stetig erhöht. Derzeit ist das IVFL an sieben Bildungsprojekten beteiligt, wovon zwei im Abschlussstadium sind. MitarbeiterInnen des IVFL haben mit jedem Projekt Erfahrungen in der Beantragung und Abwicklung von Bildungsprojekten gewonnen. Mit diesem Artikel wollen die Autoren diese Erfahrungen mit FachkollegInnen teilen und speziell die Chancen und die Grenzen von Bildungsprojekten dokumentieren. Zudem sollen die aktuellen Bildungsprojekte vorgestellt werden und anhand dieser

¹ Universität für Bodenkultur Wien, Institut für Vermessung, Fernerkundung und Landinformation, Peter Jordanstraße 82, A-1190 Wien,

E-Mail: [t.bauer, markus.immitzer, mansberger, franceso.vuolo]@boku.ac.at

die unterschiedlichen Charakteristika von Bildungsprojekten aufgezeigt werden. Zudem werden auch die vom IVFL im Rahmen der Regellehre vermittelten Lehr- und Lerninhalte vorgestellt. Die abschließenden Empfehlungen sind aus der Sicht der Autoren Erfolgsfaktoren bei der Beantragung und Abwicklung von Bildungsprojekten. Sie sollen FachkollegInnen bei der Mitwirkung bei internationalen Lehr- und Lernprojekten unterstützen.

2 Lehre am IVFL

Die Lehraufgaben des IVFL sind sehr umfangreich. Dies resultiert einerseits aus der thematischen Breite der vom Institut abzuwickelnden Lehrveranstaltungen. Andererseits bedient das IVFL mit seinem Lehrangebot vier Bachelor- und ca. zehn Masterstudien an der BOKU.

Jährlich werden vom IVFL 28 Lehrveranstaltungen mit in Summe 90 ECTS in den Fachbereichen Vermessung, Landadministration, Photogrammetrie, Fernerkundung, Bildverarbeitung und Geoinformationssysteme (GIS) abgehalten. Die Lehrveranstaltungen werden pro Jahr von etwa 2.200 Studierenden besucht. Lehrveranstaltungen mit praktischen Übungen werden aufgrund der hohen Studierendenzahlen in Parallelgruppen mit fünf bis maximal 30 Studierenden angeboten.

Die Lehre am IVFL wird größtenteils von institutseigenem Personal durchgeführt. Im Studienjahr 2018/19 wurden die zehn in der Lehre tätigen InstitutsmitarbeiterInnen mit insgesamt 147 Semesterwochenstunden (eine Semesterwochenstunde entspricht einer Stunde Lehre pro Woche im Semester, d.s. ca. 14 Stunden pro Semester) betraut. 26 Semesterwochenstunden werden durch externe ExpertInnen abgewickelt. Zudem wirkt das IVFL auch aktiv an Weiterbildungsprogrammen der BOKU mit (MANSBERGER & PAULUS 2015).

Aufgrund der umfangreichen Lehrtätigkeiten hat das IVFL große Erfahrungen bei der Wissensvermittlung an Studierende gewonnen. Die Lehrabwicklung wurde durch die Einführung moderner Lehr- und Lernmethoden kontinuierlich verbessert. Das IVFL war ein Pionier in der Verwendung von E-Learning an der BOKU (BAUER & MANSBERGER 2006, MANSBERGER et al. 2008) und auch heute wird die Kommunikation mit Studierenden, die Bereitstellung von Lehr- und Lernmaterialien sowie die Durchführung interaktiver Lehr- und Lernaktivitäten von BOKUlearn, dem auf moodle (MOODLE 2018) basierende E-Learning-System der BOKU, unterstützt.

3 Bildungsprojekte am IVFL

Die durch die Lehre gewonnen Erfahrungen veranlassten die IVFL-MitarbeiterInnen verstärkt zu einem Wissensaustausch mit FachkollegInnen sowie mit pädagogisch und didaktisch geschultem Personal auf nationaler und internationaler Ebene. Damit wurde am IVFL ein Know-How im Bereich der akademischen Aus- und Weiterbildung im Fachbereich aufgebaut, welches nunmehr im Rahmen von Bildungsprojekten auch an unterschiedlich Zielgruppen weitergegeben wird. Die vom IVFL durchgeführten Bildungsprojekte werden von der EU sowie von österreichischen Institutionen gefördert.

3.1 Derzeitige und kürzlich abgeschlossene Projekte

Die Europäische Union bietet zahlreiche Programme zur Unterstützung von akademischen Partnerschaften, zur MitarbeiterInnen- und Studierendenmobilität als auch zum Aufbau von Ausbildungskapazitäten (EU-EACEA 2018). ERASMUS+ sei hier als das wohl bekannteste und von den Budgetmitteln umfangreichste Programm genannt. Von den derzeit sieben am IVFL abgewickelten Bildungsprojekten werden vier über das ERASMUS+ Programm gefördert.

Drei der IVFL-Bildungsprojekte werden national gefördert. Zwei davon sind durch das Österreichische Partnerschaftsprogramm für Höhere Ausbildung und Entwicklungsforschung (APPEAR 2018) finanziert. Die Geldmittel dafür werden von der Österreichischen Entwicklungszusammenarbeit beigestellt. Ein Projekt wird vom österreichischen Bundesministerium für Forschung, Innovation und Technologie gefördert und von der Forschungsförderungsgesellschaft (FFG) abgewickelt.

3.1.1 SULANET

Das kürzlich abgeschlossene Projekt "Sustainable Land Management Network (SULANET)" war ein ERASMUS + -Projekt (Aktionsprogramm: Jean Monnet) und hatte eine Laufzeit von vier Jahren (September 2015 bis August 2018). Unter der Leitung der Slowakischen Landwirtschaftsuniversität in Nitra (SUA) kooperierten die folgenden vier Universitäten: Universität Foggia (Italien), Universität Bialystok (Polen), Universität Cordoba (Spanien), Universität für Agribusiness und Ländliche Entwicklung in Plovdiv (Bulgarien) und BOKU/IVFL (Österreich). Das Netzwerkprogramm ermöglichte multidisziplinäre Lehr- und Forschungsaktivitäten im Fachbereich. Übergeordnetes Ziel von SULANET war die Schaffung eines virtuellen Forschungs- und Bildungsnetzes für ein "nachhaltiges Landmanagement" mit den spezifischen Zielen, Forschung und Lehre im Bereich nachhaltiger europäischer Landbewirtschaftung zu fördern und die Wirksamkeit von Land Governance in der EU zu stärken (SULANET 2018).

3.1.2 ECAP

ECAP - "Enhancing Competencies of Central Asian Universities in Agricultural Policy focused on Environmental Protection and Land Management" ist ebenfalls ein ERASMUS + -Projekt (Leitaktionsprogramm 2 - Capacity Building). ECAP startete November 2015 und wurde im Oktober 2018 abgeschlossen. Auch im ECAP-Projekt agierte SUA (siehe oben) als Koordinator. Drei europäische Universitäten, neben SUA die Tschechische Universität für Lebenswissenschaften und die BOKU und vier zentralasiatische Universitäten, die Kasachische Nationale Agraruniversität und die Kostanay Staatsuniversität (beide aus Kasachstan) sowie die Karakalpak Staatsuniversität und die Samarkand Staatsuniversität (beide aus Usbekistan) bildeten das Konsortium. Im Projekt wurden Curricula für Umweltschutz und Landmanagement entwickelt. Die BOKU hat vorrangig ihr Wissen im Bereich der Landbeobachtung mit Fernerkundungsmethoden eingebracht (ECAP 2018).

3.1.3 EduLAND2

Das Projekt "Implementation of Academic Land Administration Education in Ethiopia for Supporting Sustainable Development (EduLAND2)" wurde im März 2016 gestartet und hat eine Lauf-

zeit von vier Jahren. Es ist ein bilaterales Projekt zwischen österreichischen und äthiopischen Universitäten, das von der Austrian Development Agency im Rahmen des APPEAR-Programms finanziert wird. Die BOKU ist Projektkoordinator und die Universität in Debre Markos (DMU), die Bahir Dar Universität (beide Äthiopien) sowie die Technische Universität Wien (Österreich) sind Projektpartner. Im Rahmen des Projekts wird ein Kompetenzzentrum für Landverwaltung an der Debre Markos Universität aufgebaut, welches den Mangel an akademischen Arbeitskräfte auf dem Gebiet der Landadministration beheben soll. EduLAND2 leistet einen wichtigen Beitrag zur nachhaltigen Entwicklung in Äthiopien, durch den Aufbau von akademischem Personal, durch die Einrichtung eines Bachelor-Studienprogramms und durch gemeinsame Forschungsaktivitäten. Das Projekt soll dazu beitragen, die Lebensbedingungen der Bevölkerung vor Ort zu verbessern, die Sicherheit des Landbesitzes zu gewährleisten und eine objektive Entscheidungsgrundlage für eine gute Regierungsführung zu schaffen. Ein Fokus des Projekts liegt auch in der Förderung von benachteiligten Gruppen mit dem Erfolg, dass der Frauenanteil bei Studierenden in den bisherigen drei Jahrgängen mehr als 35 Prozent (im ersten Jahrgang sogar 70 Prozent) beträgt (APPEAR 2018).

3.1.4 EO4Africa

"Strengthening remote sensing data processing and interpretation capacities for operational use in agricultural system monitoring (EO4Africa)" ist ein von APPEAR gefördertes Projekt, welches vom IVFL bilateral mit der Universität Eduardo Mondale aus Mozambique bearbeitet wird. Die BOKU ist Koordinator des im Juni 2017 begonnenen Projekts mit einer Laufzeit von drei Jahren. EO4Africa stärkt die Ausbildung, Forschung, das Management sowie Netzwerke, um landwirtschaftliche Kulturen zu kartieren, Anomalien im Wachstum von Pflanzen zu erkennen, sowie die Politik und die Administration laufend mit statistischen Daten über Erträge zu unterstützen. Die Untersuchungsgebiete umfassen die Provinzen Inhambane, Sofala bzw. Manica. Die dazu notwendigen Werkzeuge und zu verwendenden Daten werden im Rahmen des Projektes erarbeitet und erfasst. EO4Africa trägt damit zu einer erhöhten Transparenz und verstärkten Effizienz in der Entscheidungsfindung für das Management und die Produktion von Nahrungsmitteln bei (APPEAR 2018).

3.1.5 Netzwerk Wald unendlich

Im Projekt "Netzwerk Wald unendlich - Analog und digital - Perspektiven für Mensch und Natur" beschäftigen sich SchülerInnen aus sieben Schulen mit einer großen Bandbreite an Bilddaten- und Geoinformation, um mit diesen und anderen Untersuchungsmethoden die Natur in ihrer unmittelbaren Umgebung zu erforschen. Das Projekt wird durch den Unternehmenspartner Österreichische Bundesforste AG geleitet und hat neben dem IVFL einen weiteren universitären Partnern (Universität für Veterinärmedizin Wien) sowie einen Firmenpartner (LEHI-Copters KG). Fördergeber ist das Bundesministerium für Verkehr, Innovation und Technologie (BMVIT) im Zuge der FFG Ausschreibung "Talente regional". Das IVFL bietet im Zuge des Projekts halbtägige Workshops für SchülerInnen an, welche sich mit der Beobachtung des Waldes mit Fernerkundungssensoren und den daraus ableitbaren Informationen beschäftigen (NETZWERK WALD 2018).

3.1.6 SUSDEV

"Lifelong Learning for Sustainable Development (SUSDEV)" ist ein ERASMUS+-Projekt der (Leitaktionsprogramm 2 – Capacity Building). An dem von der Universität für Lebenswissenschaften in Warschau koordinierten Projekt arbeiten neben sechs europäischen Universitäten (zwei aus Polen, je einer aus Portugal, Schweden, Irland und Österreich), der Vereinigung der europäischen Universitäten für Lebenswissenschaften (ICA), acht Ausbildungsinstitutionen aus Russland, sechs Universitäten aus Kasachstan sowie die Bildungsministerien in Russland und Kasachstan mit. Das Projekt begann im Dezember 2016 und hat eine Laufzeit von drei Jahren. Allgemeines Ziel des Projekts ist die Stärkung der Hochschuleinrichtungen bei der Sicherstellung einer nachhaltigen Entwicklung von Industrie und Gesellschaft, die Unterstützung nationaler grüner Politiken in den Partnerländern und die Förderung der grünen Kultur durch lebenslanges Lernen. Im Projekt werden Lehr- und Lernmodulen zur Förderung von "Green Skills" erarbeitet. Der thematische Fokus von SUSDEV liegt in den Fachbereichen Ökologie, Lebensmittelwissenschaften und Landmanagement (SUSDEV 2018).

3.1.7 IRSEL

Das Projekt IRSEL – "Innovation on Remote Sensing Education and Learning" zählt auch zu den ERASMUS+ Projekten, an denen das IVFL beteiligt ist. Das Projekt wurde Ende 2017 gestartet und läuft bis 2020. Projektkoordinator ist die Óbuda Universität (OU) aus Ungarn. Neben vier europäischen Universitäten besteht das Konsortium aus vier asiatischen Universitäten. Neben OU und BOKU sind aus Europa die Jagiellonian Universität aus Krakau und das ITC der Universität Twente involviert. Die asiatischen Partner kommen von der Fujian Normal University und der Yunnan Normal University aus China sowie der Universität in Khon Kaen und dem Asian Institute of Technology aus Thailand. Im Einklang mit den Zielen der "2016 Communication on a Space Strategy for Europe" wurde das Projekt IRSEL ins Leben gerufen, um die Anwendung von Weltraumtechniken zu fördern und deren Nutzen für die Gesellschaft und die Wirtschaft zu optimieren. Das Ziel ist die Entwicklung einer innovativen Lernplattform für asiatische Länder im Bereich Fernerkundung. Dazu werden von den teilnehmenden Partneruniversitäten Lerninhalte für 20 Module aus dem Fachbereich entwickelt, die auf einer E-Learning-Plattform frei verfügbar sein werden (IRSEL 2018).

3.2 Charakterisierung der Projekte

3.2.1 Projektpartner

Die Projektpartner bei den EU-Projekten sind aus europäischen und vorrangig asiatischen Ländern. Die Anzahl der Partner im Konsortium variiert dabei zwischen sieben und 25 akademischen Institutionen. Die beiden APPEAR-Projekte sind bilateral zwischen der BOKU und Universitäten aus einem afrikanischen Land (Äthiopien bzw. Mosambik). Beim FFG Projekt arbeiten ausschließlich österreichische Institutionen zusammen.

3.2.2 Thematische Aspekte

Alle Projekte haben einen Bezug zu den Themenbereichen Vermessung, Landadministration, Photogrammetrie, Fernerkundung und Geographische Informationssysteme. Die Schwerpunktsetzungen in den einzelnen Projekten sind in Tabelle 1 ersichtlich.



Tab. 1: Fachliche Bereiche und Zielgruppen

Die Zielgruppen für die Ausbildung sind in den Projekten sehr unterschiedlich. Das Spektrum reicht hier von SchülerInnen (Sekundärer Bildungsbereich) bis hin zu FachexpertInnen. Dementsprechend sind auch die Methoden und Inhalte der Wissensvermittlung anzupassen.

3.2.3 Administrative Aspekte

Bei zwei Projekten agiert das IVFL als Projektkoordinator. Dies ist mit einer zusätzlichen administrativen Arbeitsbelastung verbunden. In zwei der vier ERASMUS+ Projekte ist das IVFL nationaler Koordinator. Bei drei Projekten hat das IVFL keine koordinierenden Aufgaben (siehe Tabelle 2).

Die Einbindung der IVFL bei der Akquisition und Ausarbeitung des Antrags war in den einzelnen Projekten sehr unterschiedlich. Bei SULANET, SUSDEV und IRSEL kam die Anfrage um Mitarbeit zu einem sehr späten Zeitpunkt. Damit war für das IVFL die Gestaltungsmöglichkeit bezüglich Ausrichtung und Budget nur sehr eingeschränkt.





Die Finanzierung von Bildungsprojekten ist sehr unterschiedlich. Während bei APPEAR-Projekten Personal, Investitionen und Reisen abgedeckt werden, müssen bei allen anderen Projekten eigene Beitragsleistungen erbracht werden. Die höchsten fallen dabei mit 20 Prozent bei Jean Monnet – Projekten (SULANET) an.

3.2.4 Methodische Aspekte

Die Charakteristik des Wissensmanagements ist in den vorgestellten Projekten sehr vielfältig. Abhängig davon, ob Wissen vermittelt, ausgetauscht oder erarbeitet werden soll, reichen die Aktivitäten von reiner Lehrtätigkeit, über die Entwicklung von Studienprogrammen, der Aufbereitung von Lehr- und Lernmaterial, der Abhaltung von Summer Schools bis hin zu gemeinsamen Forschungsaktivitäten (Tabelle 3). Die APPEAR-Projekte inkludieren auch Stipendien für akademische Studien (Master oder PhD) an einer österreichischen Universität.

Tabelle 4 gibt einen Überblick über die Charakteristik und die in den IVFL-Bildungsprojekten angewandten Aktivitäten.



Tab. 3: Charakterisierung und Aktivitäten des Wissensmanagements

4 Chancen und Stärken von Bildungsprojekten

Die in diesem Kapitel angeführten Punkte, welche die Chancen und Stärken von Bildungsprojekten aufzeigen sollen, basieren auf den von den Autoren in der Abwicklung zahlreicher Projekte gewonnenen Erfahrungen. Die einzelnen Aspekte, welche projekt- und personenbedingt als subjektiv zu sehen sind, werden im Folgenden auch kurz begründet.

Wissensgewinn für alle Projektbeteiligten

Einer der wichtigsten Gründe für die Teilnahme an Bildungsprojekten ist der Gewinn an Wissen und Erfahrung ganz allgemein. Auch als europäischer Partner lernt man mit jedem Projekt, da die außereuropäischen Länder sehr viele neue Problemstellungen und lokales Wissen in die Projekte einbringen.

Lehrerfahrung in einer Fremdsprache und anderen Kultur (Englisch)

Das Kennenlernen anderer Kulturen vor allem bezogen auf die Lehr- und Lernkultur in anderen Ländern ist ein wichtiger Motivationsfaktor. Dazu zählt die Abhaltung von Lehrveranstaltung in englischer Sprache.

Capacity Building in Partnerländern (Studierende, Lehrende, ExpertInnen)

Gerade in weniger entwickelten Ländern hat das Schlagwort Capacity Building, also die Stärkung der Problemlösungskapazität von Universitäten vor Ort, eine große Bedeutung.

Verbesserung der Lehrinfrastruktur in Partnerländern

In einigen der angeführten Projekte besteht auch die Möglichkeit, die Lehrinfrastruktur in den Partnerländern zu verbessern. So konnten beispielsweise für außereuropäischen Projektpartner Vermessungsinstrumente und anderes technisches Equipment angeschafft werden.

Möglichkeit eigenentwickelte Verfahren und Methoden in anderen Ländern zu implementieren

Der Wissenstransfer steht im Mittelpunkt der meisten Projekte. Verfahren und Methoden, die in einem Land unter diversen Bedingungen entwickelt wurden, können im Rahmen der Projekte in anderen Ländern unter anderen Voraussetzungen getestet, an die lokalen Gegebenheiten angepasst und verbessert werden.

Wissenschaft für SchülerInnen (be-)greifbar machen und Bewusstseinsbildung für den Fachbereich stärken (in Schulen, Partnerländern, ...)

Dieser Punkt betrifft vor allem Projekte, die mit Schulen gemeinsam durchgeführt werden. Dadurch kann der Fachbereich beworben werden und somit vielleicht das Interesse für ein späteres Studium geweckt werden. Aber auch der Kontakt mit lokalen FachexpertInnen, UniversitätsmanagerInnen und Interessensvertretungen ermöglicht eine Stärkung des Fachbereichs.

Kennenlernen anderer Rahmenbedingungen

Lehre in anderen Ländern stellt manchmal eine Herausforderung dar. Die in den Partnerländern nicht ausreichend vorhandene Lehrinfrastruktur erfordert für die europäischen Lehrenden eine hohe Flexibilität – speziell bei der Durchführung von praktischen Übungen mit IT-Infrastruktur.

Kultureller Austausch zwischen den Partnerländern und ProjektmitarbeiterInnen

Bildungsprojekte ermöglichen ein vertieftes Kennenlernen der Geschichte und Kultur von Partnerländern. Dieses ist für ein gegenseitiges Verstehen ein nicht zu unterschätzender Faktor. Aber auch das persönliche Kennenlernen und die in den Projekten oftmals entstehenden Freundschaften zwischen ProjektmitarbeiterInnen sind ein Anreiz für die Teilnahme an Bildungsprojekten und Erfolgsfaktoren für die Kontinuität der Zusammenarbeit.

Ausweitung der kritischen Masse für fachliche Diskussionen

Jeder Projektpartner bringt sich fachlich in der Projektarbeit ein. Dadurch wird sowohl die kritische Masse an FachexpertInnen als auch der gemeinsame Wissenspool vergrößert.

Gewinn an internationaler Erfahrung – besonders für junge ForscherInnen

Gerade bei jungen ForscherInnen spielt oft die internationale Erfahrung im Lebenslauf eine große Rolle, die im Rahmen von internationalen Kooperationen gewonnen werden können.

Beitrag zur Leistungssteigerung von Lehrenden – Instituten – Universitäten

Bildungsprojekte bieten die Chance eines Motivationsschubes für die Lehrenden, um neue Methoden, Ideen bei Lehrveranstaltungen umzusetzen, die man bei Partneruniversitäten kennengelernt hat.

Netzwerkbildung als Basis für die Anbahnung von Follow-Up Forschungsaktivitäten

Der Aufbau von Netzwerken spielt in der Wissenschaft eine große Rolle. Bildungsprojekte bilden eine ideale Basis für die Anbahnung von zukünftigen Kooperationen, da man im Zuge der Zusammenarbeit sehr gut die Fähigkeiten und Kapazitäten der Partner kennen lernt.

5 Grenzen und Schwachpunkte von Bildungsprojekten

Neben den vielen positiven Aspekten von Bildungsprojekten sollen an dieser Stelle auch die Probleme angesprochen, die auftreten können.

Keine Cash-Cow Projekte

In der Realität sind bei allen Bildungsprojekten Eigenleistungen zu erbringen. Die refundierten Kostensätze für Personal decken die tatsächlichen Personalkosten in den meisten Fällen nicht ab, und auch die in den Projekten ausbezahlten Overhead-Kosten entsprechen keineswegs den von der Universität zu tragenden Kosten für Raum und Verwaltung.

Zusätzlicher Workload besonders für Senior Lecturers & Senior Scientists

Für Senior Lecturers (MitarbeiterInnen, die fast ausschließlich in der Lehre tätig sind) bedeuten Bildungsprojekte in der Regel eine Ergänzung zur normalen Lehrverpflichtung und daher zusätzlichen Arbeitsaufwand. Senior Scientists, die vorwiegend in der Forschung tätig sind, werden durch derartige Projekte ebenfalls stärker belastet.

Eingeschränkte Möglichkeiten der Einbindung junger MitarbeiterInnen mit geringer Lehrerfahrung in die Projekte (EU-Länder)

Bildungsprojekte setzen eine gewisse Lehrerfahrung voraus. Für junge WissenschafterInnen stellt dies eine Herausforderung dar. Daher ist es oft schwierig, junge MitarbeiterInnen für diese Projekte motivieren zu können.

Des-Interesse vereinzelter ProjektmitarbeiterInnen bei Veranstaltungen / Projekttreffen

Die fehlende Motivation von ProjektmitarbeiterInnen kann ein hemmender Faktor bei Bildungsprojekten sein. Nicht immer nehmen die eigentlichen FachexpertInnen an den Projektentreffen /-workshops teil, sondern andere Überlegungen und Intentionen stehen im Vordergrund für die Teilnahme (Hierarchie, schöne Reisen, Einkaufsmöglichkeit, ...). Bei Schulprojekten zeigt sich, dass manche SchülerInnen "zwangsbeglückt" werden und/oder zu wenig von den LehrerInnen auf das Projekt vorbereitet wurden.

Homogenität der Projektpartner im Hinblick auf Fachwissen, Organisation, Output

Gerade bei internationalen Projekten ist das unterschiedlich Niveau in Bezug auf das Fachwissen und die Organisation bei den Partnern eine Herausforderung. Es besteht die Gefahr, dass es dadurch zu Reibungsverlusten und Zeitverzögerungen kommen kann bzw. dass der Arbeitsaufwand zwischen den Projektpartnern ungleich verteilt wird.

Personalwechsel

Bei Projekten mit Entwicklungsländern wurde die Erfahrung gemacht, dass es häufig zu Personalwechseln an den Universitäten kommt, da gut ausgebildete MitarbeiterInnen in der Hierarchie rasch aufsteigen und neue Jobchancen wahrnehmen. Dies bedingt wiederum einen Wissens- und damit einen Zeitverlust, da sich neues Personal erst einarbeiten muss.

Sprachbarrieren

Sprachliche und dazu kulturelle Barrieren erschweren oft die verbale Kommunikation zwischen den Projektpartnern. Englisch als Weltsprache kann nicht immer vorausgesetzt werden. Besonders deutlich wird dies bei Telefonkonferenzen. Zwar werden manchmal DolmetscherInnen als Unterstützung eingesetzt, welche tlw. jedoch Fachbegriffe falsch oder missverständlich übersetzen.

Forschungsaktivitäten nur beschränkt gefördert, daher wenig Forschungs-Output

Bei den meisten Bildungsprojekten ist nur ein geringer Anteil der Förderung für Forschungsaktivitäten reserviert. Bildungsprojekte bieten daher wenig Spielraum für wissenschaftliche Publikationen. Als Folge davon spiegelt sich die in diesen Projekten erbrachte Leistung oftmals nicht in den Kennzahlen von Universitäten wider (SCI-Publikationen, ...).

Keine Garantie auf Kontinuität der Kooperation

Im Zuge von Kooperationen werden sehr häufig neue Netzwerke aufgebaut, deren Weiterführung sinnvoll wäre. Mit dem Projektende bricht der Kontakt aber häufig ab. Nur wenige Projektschienen bieten die Möglichkeit für unmittelbar anschließende Folgeprojekte.

Überprüfung der Projektergebnisse mehr auf administrative (finanzielle) als auf wissenschaftliche Kriterien beschränkt

Bildungsprojekte werden manchmal nicht nach wissenschaftlichen Kriterien überprüft, sondern häufiger steht die reine administrative Überprüfung im Vordergrund. Dadurch besteht die tendenziell die Gefahr, dass nur die minimalen Projektziele erfüllt werden.

Aufwändige Reiseaktivitäten

Internationale Bildungsprojekte sind meist mit zeit- und kostenintensiven Reisen verbunden. In Zeiten der Klimaerwärmung sind aufwändige Flugreisen kritisch zu hinterfragen.

Kommunikationsmöglichkeiten (Telekonferenzen, Internetzugang)

Neben den oben angeführten sprachlichen Barrieren treten häufig auch technische Probleme bei der Kommunikation auf. Telekonferenzen erfordern stabile Internetverbindungen, die nicht immer gegeben sind.

Unregelmäßige bzw. längere Abstände zwischen Treffen

Eine regelmäßige Kommunikation und häufige Treffen zwischen den Partnern haben einen positiven Einfluss auf die Fortschritte bei einem Projekt. Gerade bei längeren Projektlaufzeiten kommt es aber oft zu größeren Abständen zwischen Treffen, was durch Telekonferenzen nicht immer ausgeglichen werden kann.

6 Empfehlungen für die Beantragung und Abwicklung von Bildungsprojekten

Anhand der oben angeführten Stärken und Schwächen von Bildungsprojekten lassen sich Empfehlungen für die Beantragung und Abwicklung von Bildungsprojekten ableiten. Die im Folgenden angeführten Empfehlungen sind durch die in den von den Autoren bearbeiteten Projekten gewonnenen Erfahrungen geprägt. Damit sind diese nicht frei von Subjektivität.

Rechtzeitiges Bilden von Konsortien

Sobald es eine Projektidee gibt, sollte mit der Zusammenstellung eines Konsortiums begonnen werden. Geeignete Projektpartner sollen so bald wie möglich eingebunden werden, da diese ansonsten für andere Projekte abgeworben werden. Eine sehr späte Einladung zur Mitarbeit versperrt dem potentiellen Partner auch das Einbringen eigener Ideen und damit in weiterer Folge auch oft das Engagement in der Projektabwicklung.

Zeit nehmen für die Ausarbeitung des Projektvorschlags - Gemeinsame Entscheidungen aller Projektpartner

Mit der Ausarbeitung des Projektvorschlags sollte rechtzeitig begonnen werden. Inhalte, Budget und anzuschaffende Infrastruktur müssen zwischen allen Partnern abgesprochen und gemeinsam beschlossen werden. Dies benötigt Zeit. Ideal ist dabei eine Finanzierung der Projektanbahnung, wie es bei den APPEAR-Projekten üblich ist. Dabei kann mit einer Kurzbeschreibung und einer Kostenabschätzung um ein "Preparatory Funding" angesucht werden. Wird dieses genehmigt, stehen dem Konsortium Geldmittel (im Fall von APPEAR sind dies € 15.000) für die Ausarbeitung eines endgültigen Projektvorschlags zur Verfügung.

Aufbau einen engen Kontakts zwischen allen Projektpartnern

Schon während der Erarbeitung des Projektvorschlags ist eine regelmäßige Kommunikation mit allen Projektpartnern von Vorteil. Die Erfahrung zeigt, dass informierte Projektpartner sich stärker mit dem Projektvorschlag identifizieren und sich auch stärker einbringen. Diese Empfehlung gilt auch für die Projektabwicklung.

Konsortium mit Partnern bilden, welche ergänzenden Kompetenzen haben

Bei der Auswahl von Projektpartner ist es günstig, wenn die zur Erreichung der Projektziele benötigten Kompetenzen über alle Institutionen gut verteilt sind. Dies verstärkt einerseits den Lerneffekt für alle ProjektmitarbeiterInnen und verringert etwaiges Konkurrenzdenken.

Ausreichend Zeit für Workshops (mit definierten Ergebnissen) einplanen

Für Projekttreffen und Workshops ist genügend Zeit einzurechnen. Projekttreffen sollten regelmäßig und wenn möglich einmal pro Quartal durchgeführt werden (davon mindestens ein physisches Treffen). Weitere Treffen können über Telekonferenzen abgewickelt werden. Projekttreffen und vor allem Workshops sollten sehr gut vorbereitet werden. Die aktive Einbindung aller Projektpartner sollte bei der Planung ebenso sichergestellt werden (z.B. durch vorgesehene Kleingruppenarbeiten). Detaillierte und klar formulierte Zeit- und Zielvorgaben sind Garant für einen verbesserten Output.

Outputs klar definieren - Verwendung von Vorlagen, Best Practice Beispielen

Resultate in Bildungsprojekten sind sehr oft ein gemeinsam erstellter Syllabus, Studienpläne oder Lehr- und Lerninhalte. Zur Gewährleistung von Vollständigkeit und Harmonisierung dieser Dokumente sind Vorlagen mit vorgegebenen Strukturen, eine Kurzbeschreibung des erwarteten Inhaltes sowie Best Practice Beispiele hilfreich.

Einbindung von gemeinsamen Forschungsaktivitäten und Ausarbeitung eines Antrags für ein Folgeprojekt

Die Zusammenarbeit in Bildungsprojekten wird durch die Einbindung von gemeinsamen Forschungsaktivitäten gefördert. Zudem kann die Kontinuität von Partnerschaften durch die Ausarbeitung eines Antrags für ein anschließendes Forschungs- oder Bildungsprojekt abgesichert werden.

Laufende Information über den Projektfortschritt an das Universitätsmanagement, an die lokalen Interessengruppen/institutionen und an die breite Öffentlichkeit

Projektpartner sind in den meisten Fällen Universitätsinstitute. Da bei den meisten Bildungsprojekten Eigenleistungen der Universität eingebracht werden müssen, ist eine laufende Information der Universitätsleitung über den Projektfortschritt von Vorteil. Aber auch Stakeholders sollten über das Projekt von Anfang an informiert sein, da Exkursionen zu Fachinstitutionen jedes Bildungsprojekt bereichern. Eine Information bzw. Bewerbung des Projekts sollte ab Projektbeginn in der breiten Öffentlichkeit erfolgen.

Administration ausreichend im Budget berücksichtigen

Der administrative Aufwand in der Projektabwicklung wird fast immer unterschätzt. Jede Projektschiene hat üblicherweise eigenen Richtlinien für die finanzielle Abrechnung. Zudem haben auch die Projektpartner eigenen Gewohnheiten in ihrem Rechnungswesen. Die Harmonisierung ist oft mit großem Aufwand für Koordinatoren verbunden. In längeren Projekten ist auch mit der Fluktuation von ProjektmitarbeiterInnen zu rechnen. Jeder Wechsel im Personal eines Projektpartners benötigt zusätzliche administrative Arbeiten.

Partnerschaften immer als gleichberechtigt gestalten

Bei Bildungsprojekten gibt es üblicherweise zwei Arten von Projektpartnern: Jene, bei welchen Bildungskapazitäten aufgebaut werden (Nehmer) und jene, welche diese Aktivitäten von außen unterstützten (Geber). Eine erfolgreiche Projektdurchführung ist nur möglich, wenn die Partnerschaften aller im Projekt beteiligten Institutionen gleichberechtigt ist.

7 Fazit und Ausblick

Auch wenn sich das Engagement von Forschern in Bildungsprojekten nicht direkt in den aktuellen akademischen Leistungskennzahlen (Output an wissenschaftlichen Publikationen) widerspiegelt, weisen sie großes Potenzial für die persönliche Entwicklung auf. Durch die üblicherweise internationalen Projekte ist der kulturelle Austausch von großer Bedeutung. Die Möglichkeit zur internationalen Vernetzung ist dabei auch für jüngere WissenschafterInnen vorteilhaft. Darüber hinaus sind auch die gewonnenen fachlichen Erfahrungen für alle Projektbeteiligten enorm. Die bedarfsgerechte Aufbereitung von Lehr- und Lernmaterialen unter Berücksichtigung des Vorwissens der Zielgruppe (z.B. SchülerInnen) und der technischen Gegebenheiten erfordert eine intensive Auseinandersetzung mit der fachlichen Materie.

Aus der Zusammenarbeit mit den Projektpartnern entstehen oft Kooperationen, welche für zukünftige Vorhaben von großem Vorteil sind. Bildungsprojekte geben Auskunft über die wissenschaftliche Qualifikation und Arbeitsweise von ProjektmitarbeiterInnen.

Abschließend kann festgehalten werden, dass Bildungsprojekte keineswegs einen einseitigen Wissenstransfer von Europa zu Partnerländern darstellen, sondern dass alle Projektpartner von Bildungsprojekten profitieren.

8 Danksagung

Die Autoren danken den Fördergebern und Partnern der folgenden Projekte:

EU Jean Monnet Networks Projektnr. 564651-EPP-1-2015-1-SK-EPPJMO-NETWORK "Sustainable Land Management Network (SULANET)".

EU Erasmus+ Projektnr. 561590-EPP-1-2015-1-SK-EPPKA2-CBHE-JP "Enhancing Competencies of Central Asian Universities in Agricultural Policy focused on Environmental Protection & Land Management (ECAP)".

EU Erasmus+ Projektnr. 574056-EPP-1-2016-1-PL-EPPKA2-CBHE-SP, "Lifelong Learning for Sustainable Development (SUSDEV)".

EU Erasmus+ Projektnr. 586037-EPP-1-2017-1-HU-EPPKA2-CBHE-JP, "Innovation on Remote Sensing Education and Learning (IRSEL)".

Appear Projektnr. 113 "Implementation of Academic Land Administration Education in Ethiopia for Supporting Sustainable Development (EduLAND2)".

Appear Projektnr. 176 "Strengthening remote sensing data processing and interpretation capacities for operational use in agricultural system monitoring (EO4Africa)".

FFG Projektnr. 862346 "Netzwerk Wald ∞ analog und digital - Perspektiven für Mensch und Natur".

9 Literatur

- BAUER, T. & MANSBERGER, R., 2006: Zeit f
 ür Ver
 änderung: e-Learning als Hilfsmittel f
 ür effizientere Lehre. 3. Internationale Moodle-Konferenz in Österreich, Hagenberg, September 14.-15.
- MANSBERGER, R., BAUER, T., MICHALEK, C-R., 2008: Communication Tools in E-learning: Experiences in Academic Geomatics Education. FIG International Workshop Sharing Good Practices, Enschede, Netherlands, June 11-13, 2008.
- MANSBERGER, R. & SEHER W., 2018: Projects for Higher Education in Land Management. Experiences, Opportunities, Limits and Recommendations. In: Hepperle, E; Paulsson, J; Maliene, V; Mansberger, R; Lisec, A; Guelton, S (Eds.), Opportunities and Constraints of Land Management in Local and Regional Development. Integrated Knowledge, Factors and Trade-Offs, 55-65; vdf Hochschulverlag AG an der ETH Zürich, Zürich.
- MANSBERGER, R. & PAULUS, CHR., 2015: LLL Course on Land Management in Rural Areas at BOKU University Vienna: The Making Of FIG Working week 2015 - From the Wisdom of the Ages to the Challenges of the Modern World, Sofia, May 17-21. In: International Federation of Surveyors, FIG Working Week 2015 - Proceedings, https://www.fig.net/resources/proceedings/fig_proceedings/fig2015/papers/ts05f/TS05F_mansberger_paulus_7839.pdf.

Internetzugriffe (Dezember 2018)

APPEAR, 2018: https://appear.at/

- ECAP, 2018: http://www.ecap.uniag.sk/
- EU-EACEA, 2018: https://eacea.ec.europa.eu/erasmus-plus_de
- IRSEL, 2018: http://irsel.eu
- MOODLE, 2018: https://moodle.org/
- NETZWERK WALD, 2018: https://www.bundesforste.at/natur-erlebnis/biosphaerenpark-wienerwald/ergaenzende-seiten/ww-archiv/projekte-archiv/netzwerk-wald-archiv/netzwerk-waldu-n-e-n-d-l-i-c-h.html
- SULANET, 2018: http://www.sulanet.uniag.sk/

Methodik des Blended Learnings im Bereich der Aus- und Weiterbildung der mobilen 3D Vermessung

VALENTIN WICH¹ & ANSGAR BRUNN¹

Zusammenfassung: Die Technik der mobilen 3D Vermessung wird immer komplexer und durch die zunehmende Digitalisierung weiter beschleunigt. Ein ähnlicher Trend zeichnet sich auch in verwandten Branchen ab. Vermessungsleistungen müssen daher schneller und in anderen Dimensionen erfüllt werden. Dies stellt gerade kleine und mittlere Unternehmen vor technologische, aber auch personelle Herausforderungen. Im Projekt "Mobile Vermessung für kleine und mittlere Unternehmen (MV-3D4KMU)" wird versucht, die Fortbildung in den Unternehmen zu ergänzen und zu vereinfachen. Das geschieht durch intensivere Nutzung des Mediums Video, mit einer für den Endnutzer einfachen Plattform und durch das spielerische Lernen.

Wissen wird nicht nur aus einer Richtung vermittelt, sondern die Bildungsbedürfnisse werden in der Gruppe erfasst und es wird im ersten Schritt der Grundstein für zielgerichtetes Lehren und gemeinsames Lernen gelegt. Die Erfahrungen aus der Theorie werden in praktische Beispiele umgesetzt. Der Wissensaustausch zwischen der Hochschule und Unternehmen, neben gemeinsamen wissenschaftlichen Arbeiten, schafft eine intensive Rückkopplung und Vernetzung zwischen allen am Bildungsprozess beteiligten Projektparteien.

Autorensoftware für e-Lernmodule ermöglicht eine saubere, einheitlich gestaltete, ansprechende Gestaltung. Webtechnologien begünstigen das einfachere Verteilen von Daten und fördern daneben die digitale Zusammenarbeit im Web, auch über große Entfernungen hinweg.

Bildung soll noch spannender, zielgerichteter und effizienter stattfinden. Die Außenwirkung und Begeisterung für das Feld muss weiter gestärkt werden, um für noch mehr Interesse und Nachwuchs im Feld der mobilen 3D Vermessung, mit ihren Teilbereichen, zu sorgen. Abschließend werden Trends in Bildung und Technik aufgezeigt, um mit der immer höher werdenden Entwicklungsgeschwindigkeit der Technik auch in den Unternehmen mithalten zu können.

1 Einleitung

Die mobile 3D Vermessung hält unaufhaltsam Einzug in die tägliche Arbeit der Geodaten verarbeitenden Unternehmen. Die Datenakquise wird durch die Kombination unterschiedlicher Sensoren (z. B. Kameras, Laserscanner, Satellitenpositionierung, Inertialmesssysteme und Odometer) immer komplexer und durch die zunehmende Digitalisierung weiter beschleunigt. Das Dienstleistungsangebot in der Vermessung, wie auch der Weiterverarbeitung der Daten, ist vorwiegend durch kleine und mittlere Unternehmen (KMU) geprägt, die durch die stark wandelnden Anforderungen vor technologische und personelle Herausforderungen gestellt werden.

Das Projekt "Mobile Vermessung für kleine und mittlere Unternehmen (MV-3D4KMU)" geht besonders auf diese Bedürfnisse, die Fortbildung in den Unternehmen zu ergänzen und zu vereinfachen, ein. Das geschieht durch intensivere Nutzung des Mediums Video, mit einer für den Endnutzer einfachen Plattform, sowie durch spielerisches Lernen.

¹ Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Röntgenring 8, D-97070 Würzburg. E-Mail: [valentin.wich,ansgar.brunn]@fhws.de

Das Thema der Aus- und Weiterbildung, gerade in technischen Berufen, ist sehr komplex. Dabei ist die Qualität, Reichweite und Platzierung der Bildungsinhalte mit erfolgsentscheidend. Akzeptanz und Interesse für das Thema lässt sich durch Anschaulichkeit und eine positive Haltung zum Thema nach innen und außen erzeugen. Die Teilnehmer sollen sich untereinander gut vernetzen, zusammenarbeiten und die so entstehenden Synergieeffekte für ihr Fachgebiet nutzen. Durch die Beschleunigung hoffen wir in der mobilen 3D Vermessung besser geschultes Fachpersonal in kürzerer Zeit ausbilden zu können.

2 Motivation

Die technischen Methoden in der mobilen 3D Vermessung werden immer schnelllebiger, die Produktlebenszyklen werden immer kürzer und die Produkte in ihrer Technik komplexer. Aber die Technologie schafft auch viele neue Möglichkeiten und wird damit noch zukunftsweisender. So lassen sich nun sehr komplexe Ansätze und Ideen einfacher umsetzen, und es kann schneller vom Gedanken hin zur Umsetzung gearbeitet werden. Dadurch, dass auch andere Arbeitsbereiche mit der Digitalisierung beschleunigt wurden, besteht nun der Bedarf das auch die Vermessungsaufgaben entsprechend schneller erfüllt werden. Viele Firmen benötigen somit mehr Personal, oder effizientere, schnellere Technologien. So herrscht der Bedarf neue Fachkräfte in kurzer Zeit anzulernen vor, zum anderen muss der Fortschritt der Technik gesichert sein. Zudem müssen Fachpersonal reintegriert und junge Menschen für das Thema "Mobile 3D Vermessung" begeistert werden.

Der Transformationsprozess von der Information zum Wissen, aber auch der notwendigen Abgrenzung mit Chancen, Möglichkeiten und Notwendigkeiten wird von HOLTEN & NITTEL (2010: 176) beschrieben: "Weitere Faktoren für das Lernen sind: Die gestiegene globale Verfügbarkeit von Daten, Zeit- und Ortsunabhängigkeit beim Lernen, aber auch Bewältigung der gestiegenen Fülle an Informationen [...] Zudem werden sowohl soziale als auch inhaltlich-effektive Faktoren auf neue Art und Weise bedient."

KERRES (2012: 24f) nennt einige weitere Erfolgsfaktoren und begreift das Lernen als Prozess, bei dem es die "*Selbstregulation als innerpsychische Leistung*" mit Lernvorbereitung, Lernhandlung, -hilfen und -kontrollen, die Leistungsbewertung, sowie die Motivation die Konzentration aufrecht zu erhalten gilt. Daneben spielt das bereichsspezifische Vorwissen, die Planung, Zielrichtung, Motivation, Abschirmung und Umgang mit Erfolg und Misserfolg eine entscheidende Rolle.

KLIMSA & ISSING (2011: 31) nutzen eine Problemlösetechnik zur konstruktiven Erarbeitung von Wissen und zum kritischen Denken. Sie geben verschiedene Arten der Problemlösung in ähnlicher Form stichpunkthaltig wieder: "[...] Versuch und Irrtum, Umstrukturierung, Anwendung von Strategien, Kreativität, (vernetztes) Systemdenken. Häufig führt das Eingrenzen der Problemlage, das Zulassen ungewöhnlicher Denkwege, Assoziation, Brainstorming und Entspannung zum Finden der Lösung. Strategien des Problemlösens sind lernbar."

Andererseits ergeben sich im Projekt auch Synergieeffekte zwischen Hochschule und Unternehmen. In der Hochschulbildung wird die Ausbildung praxisnah gestaltet, um die Studierenden für den Arbeitsmarkt vorzubereiten. Eine besondere Intensivierung der überbetrieblichen Zusammenarbeit besteht im Dualen Studium. Unternehmen liefern hier oft Anregungen, oder Impulse für die Lehre an der Hochschule. Auf der anderen Seite wird bei Forschungsarbeiten in Unternehmen

V. Wich, A. Brunn

neues, praxisnah erarbeitetes Wissen gewonnen. Durch Anwendung des neuen Wissens zusammen mit Unternehmen und in Kombination mit deren Fachwissen, können so Best-Practices innerhalb des Projekts verfeinert werden.

Das Heilbronner Modell zum berufsbegleitenden Lernen ist dabei eine Option. Darin wird das theoretische Wissen mit Methodenkompetenz und sog. "On-the-Job-Projekten" verzahnt (CENDON et al. 2016: 89). Das Projekt soll noch darüber hinausgehen und das so erarbeitete Wissen in einem Netzwerk zusammentragen und aufbereiten.

Lernen ist ein vielschichtiger Prozess. Wohingegen sonst im klassischen E-Learning nur Text Bilder, gelegentlich Videos verwendet wurden, werden Lektionen umso mehr ansprechender gestaltet. Indem Interaktion eingebaut wird, mehr auf Videos, und direktes Feedback durch Webkonferenzen, aber auch durch Übungen, gesetzt wird, ist das Wissen einfacher zugänglich.

Außerdem wird mit der Veröffentlichung guter Arbeitsergebnisse auf der Projektseite, geplant sind auch YouTube Videos, eine besondere Außenwirkung erzeugt, um das Berufsbild populärer zu machen. Zudem dienen die neuen Ansätze der besseren Einbindung der Lernenden. Das Medium Video ist für die heutige Generation sehr einfach greifbar, hat eine große Reichweite und ist jederzeit einfach verfügbar. HANDKE (2017: 76) erörtert dazu Parameter zur Relevanz von Inhalten: Kanalinhaber, Videoproduzent, Produktionsdatum, Kommentare zum Video, Länge des Videos, Bewertungen des Videos und die Anzahl der Aufrufe.

In der mobilen 3D Vermessung kommen mehrere Schichten aus Hardware und Software aus der digitalen Photogrammetrie, aus dem Laserscanning, aus der Bildverarbeitung, bis hin zur Navigationstechnologie zum Einsatz. Es sollte dabei ein Bewusstsein und Grundverständnis für alle Bausteine des eingesetzten Systems vorhanden sein. Das bedeutet, wenn die Systeme einfacher beherrschbar werden und besser ineinander integriert sind, dann finden diese mehr Anwendung auf Grund Ihrer Einfachheit. Wir tragen im Projekt auch in Teilen etwas dazu bei, indem wir z.B. Leitfäden zur Bedienung liefern, oder Fachsoftware vereinfachen.

Ein Formvorschlag der Erwachsenenbildung von LIMA & GUIMARÃES (2011: 124) kann in Teilen adaptiert werden. Im Projekt wird dieser über alle Projektbeteiligten hinweg angewendet werden, nicht nur innerhalb einer Institution. LIMA & GUIMARÃES (2011) beschreiben im "European Qualification Framework" (EQF) ein achtstufiges Modell angefangen vom einfachen Teilnehmenden, dem Lernenden, dem Hinterfragenden, der Fachkraft im Thema, der Mitwirkung in Arbeitsgruppen, der Übernahme von Personalverantwortung und Leitung von Mitarbeitern, bis hin zur fachlichen Autorität mit dem Willen die Forschung mit außerordentlichen Beiträgen zu bereichern.

3 Methoden

MEYER & BRUNN (2016) setzten im Projekt "DiPhoBi4KMU" die Online Lernplattform Moodle ein. Einzelne Lektionen sind aus bebilderten Texten im Web und PDFs zusammengesetzt worden, diese konnten über die Lernplattform von den Projektteilnehmern heruntergeladen werden. Die PDFs bieten den Nachteil, dass sie heruntergeladen werden müssen bzw. sich im Browser in einem eigenen Betrachtungsfenster öffnen. Durch die Verwendung einer Autorensoftware, welche Webinhalte im Ergebnis liefert, findet die Integration des Wissens nun nahtlos auf der Webplatt-

form statt. So ist das Nutzererlebnis unterbrechungsfreier gestaltet und es ist keine weitere Software zum Betrachten der Inhalte notwendig. Eine eingehende Beurteilung des Projekts erfolgt in BRUNN (2018).

EULER & HOLZ (1992: 18) beschreiben die Verwendung einer Autorensoftware zur Schaffung technischer Lehrmittel. Die daraufhin gestellte Frage, ob die Pädagogik technisiert wird oder die Technik pädagogisiert werden muss, ist nach wie vor aktuell.

Ein Trend zu beidem ist erkennbar, z.B. durch mehr Technik im Alltag verbessert sich der technische Umgang. Durch die frühe Technikvermittlung der Schulen findet eine Pädagogisierung derselben statt. Mit der Vereinfachung der Technik und den neuen Möglichkeiten bietet sich mehr Raum für die Wissensvermittlung und damit auch für die Pädagogik, so muss nun weniger Zeit in die Beherrschung der Technik selbst, zur Erstellung der Bildungsinhalte investiert werden.

In Leitfäden aus der Forschung, aus eigenen Anwendungen, oder aus Erfahrungen von Teilnehmern werden Erfahrungen gewonnen, die zur schnelleren Wissensvermittlung beitragen.

Die Implementierung eigener Software demonstriert die Machbarkeit. Mittels Apptechnologie wurde eine Augmented Reality Anwendung erstellt. Eine Punktwolkenbetrachtungssoftware für das Web ist ebenfalls kreiert worden. Die wenigsten der Endanwender der Software sind fähig mit Konsolenanwendungen umzugehen. Damit bestand auch hier der Bedarf diesen Bereich zu vereinfachen. Darum wurde eine einfache Bedienoberfläche für den Konverter geschaffen, die aber auch erfahrenen Nutzern Zeit spart. D.h. in diesem Fall wird die Software selbst zu einer Art Best-Practice, d.h. durch die vereinfachte Ausführung wird ein schnelles, brauchbares Ergebnis für den firmeneigenen Arbeitsprozess erzielt.

Viel Wert wurde in der Planung auf online Interaktion gelegt. Online Quizze ermöglichen es den Benutzern spielerisch zu lernen. Die Benutzer werden durch verschiedene Fragestellungen geführt und können so im Wettbewerb mit anderen Teilnehmern Antworten geben. Je schneller und präziser geantwortet wird, desto mehr Punkte werden vergeben. Nach jeder Frage wird die richtige Antwort gezeigt. Das bietet die Möglichkeit auch aus falsch beantworteten Fragen zu lernen. Durch den Wettbewerb werden auch möglichst richtige Antworten gegeben und die Teilnehmer angespornt, d.h. hier wird das Lernen spielerischer mit sog. Gamification gestaltet.

Storytelling kann ebenfalls, als ein wieder in die Mode kommendes Stilmittel in der Bildung verwendet werden, um den Lernenden in den Bann zu ziehen und Spannung zu erzeugen. Schon bei Plato wird die Bedeutung Zöglinge in der Bildungsvermittlung in seinen Bann zu ziehen erläutert. Er schlug vor, auf eine besonders anschauliche, lebhafte Art, mittels spannender Geschichtenerzählung, aber auch mit Ruhe und Zeit zu lehren: *"Wohlan denn, so laß uns in Gedanken unsere Männer erziehen, als wären wir Märchenerzähler und hätten reichlich Zeit."* (PLATON et al. 1989). Dieses Lebhafte wird neben den Präsenzveranstaltungen durch Online Konferenzen, Webinare und interaktive Meetings umgesetzt. Onlineaktivitäten bieten den Vorteil räumlich getrennte Nutzer mit wenig Zeitaufwand zusammen zu bringen. Im Projekt wurde das so umgesetzt, dass die Teilnehmer sich mit dem bzw. den Referenden austauschen, die Teilnehmer untereinander kommunizieren können, aber auch ein Rollenwechsel durch den Moderator ermöglicht werden kann. Jeder der Teilnehmer kann zum Moderator ernannt werden und präsentieren. Oder die Teilnehmer können auch zur Zusammenarbeit in Gruppen befähigt werden. Weiterhin werden Präsenzveranstaltungen abgehalten, sie ermöglichen intensiveres Lernen mit direkteren Rückfragen. Die persönliche Ebene rückt in den Vordergrund und beim Lernen entsteht meist mehr Rückkopplung z.B. durch die Wahrnehmung der Körpersprache des Gegenübers.

Daneben bieten Präsenzveranstaltungen den Teilnehmern, aber auch uns die Möglichkeit zur Netzwerkbildung. Im Projekt versuchen wir zum einen so viel Netzwerkbildung wie möglich zu betreiben, damit sich Teilnehmer untereinander besser austauschen können, damit diese ihre Arbeiten so besser bewältigen können. Zum anderen dient das Networking dem schnelleren Informationsaustausch zwischen allen Projektbeteiligten. Wir bieten dafür in Präsenztreffen einen kleinen Ortswechsel zur meist sitzenden Gruppentätigkeit an: An runden Stehtischen mit Getränken und Snacks kann sich unkompliziert und frei ausgetauscht werden.

In den Räumlichkeiten der Unternehmen wird empfohlen einen Rückzugsort zum Lernen oder zur Teilnahme an Webinaren etc. zu schaffen. Dies sind beispielsweise kleine Projekträume, Büros, Sitzecken usw. (OBERBECK & KUNDOLF 2018: 58).

In Aus- und Weiterbildungsangeboten können nachhaltig Kontakte geknüpft, Hindernisse diskutiert und inhaltliche, didaktische, strukturelle Anpassungen bzw. Ausrichtungen des Weiterbildungsangebotes ermittelt werden. Best-Practices und andere positive Forschungsergebnisse können auch weitere Professoren und den wissenschaftlichen Mittelbau für eine Beteiligung an der Mitwirkung an Weiterbildungen motivieren (OBERBECK & KUNDOLF 2018: 61).

4 Ergebnisse

4.1 Online-Plattform

Zur Zusammenstellung der Inhalte wird die Lernplattform Moodle verwendet. Aus Moodle heraus werden alle Bildungsbausteine verwaltet. Es sind dort Lektionen vorhanden, die Projektwebseite ist verknüpft, Projektbenachrichtigungen werden gelistet, es kann aber auch an interaktiven Online-Meetings mit Adobe Connect oder auch an Abstimmungen teilgenommen werden.

Als zusätzlicher Ansporn wird in Moodle das sog. Level-UP System eingesetzt. Desto mehr Inhalte ein Teilnehmer konsumiert und je mehr Nutzung der Online-Lernplattform stattfindet, umso mehr Punkte werden dem Teilnehmer angewiesen. Damit wird bei viel Aktivität auf der Lehrplattform ein Level-Aufstieg begünstigt. Die Benutzer sehen ihren jeweiligen aktuellen Punktestand anonymisiert im Wettbewerb zueinander in einer Rangliste.

Das E-Learning dient zudem der effizienten Benachrichtigung der Teilnehmer. Bevor Webinare oder andere Veranstaltungen gehalten werden, bedarf es einer Terminabstimmung. Per E-Mail wird drei bis vier Wochen vor der Veranstaltung ein Abstimmungslink verteilt, er enthält eine Kurzbeschreibung der Inhalte der Veranstaltung. Die Frist der Abstimmung beträgt wenige Tage. Nach Beendigung wird allen Teilnehmenden ein Termin mit Uhrzeit und weiterführenden Informationen, wie z.B. der Meeting-URL für das Webinar, genannt. Einen Tag vor der Veranstaltung erhalten alle Teilnehmer eine Erinnerung aus dem E-Learning, nochmals mit einer Agenda, Kurzbeschreibung, Räumlichkeiten und den weiteren Daten zum Termin.

Für die interaktive Zusammenarbeit kommt Adobe Connect zum Einsatz. Teilnehmer können in Adobe Connect präsentieren, es können Gruppen gebildet werden, oder auch komplett frei im Team auf einem virtuellen Whiteboard zusammengearbeitet werden. Dieser virtuelle Raum gibt

Möglichkeiten schnell Ideen zu sammeln. Alle Nutzer können Begriffe auf die weiße Fläche zeichnen. Der Moderator kann z.B. die Begriffe gleich gedanklich sortieren. Anfangs sollten alle Ideen zugelassen werden. Im Nachgang bietet sich eine konstruktive Diskussion zur Einordnung der Begriffe und zu den Begriffen selbst an. Die Ideenentwicklung kann in Adobe-Connect als Video mitgeschnitten werden, oder als Screenshot gespeichert werden.

4.2 Quizze

Das Lernen wird noch spielerischer gestaltet, durch die sog. Gamification. Hier kommen im Bereich von Quiz-Spielen sog. "Kahoots" zum Einsatz. Der Moderator erstellt auf der Internetseite kahoot.it interaktive Quiz Anwendungen. Diese können dann mit einer URL in verschiedene Medien eingebunden werden. Eine andere Möglichkeit wäre den Teilnehmern in einer Präsenzveranstaltung einen Bildschirm mit einem Zugangscode zu zeigen, über den später die Fragen dargestellt werden. Oben im Bild steht auch ein Kennwort, damit sich die teilnehmende Gruppe am Quiz authentifizieren kann. Die Quizteilnehmenden vergeben danach ihren individuellen Spitznamen für das Spiel. Das Quiz beginnt und das Handy dient jetzt als Fernbedienung für das Quiz. Beim Zeigen der Fragen wird pro Frage ein Symbol auf dem Handy jedes Teilnehmenden, als Antwortmöglichkeit gezeigt.



Abb. 1: Simulation eines Kahootquiz mit Darstellung der Fragesituation: Links befindet sich die Präsentationsplattform mit der Frage, der ablaufenden Zeit und den Antwortmöglichkeiten. Rechts wird das Handy des Quizteilnehmers dargestellt. Dieser kann durch das Klicken der Symbole auf dem Handydisplay (rechts) auf die dargestellten Fragen (links) antworten.

In Abhängigkeit der Beantwortungszeit werden Punkte für richtig beantwortete Fragen vergeben. Die drei Teilnehmer mit dem größten Punktestand werden in jeder Runde nach Präsentation der richtigen Lösung gezeigt. Der Teilnehmer mit der maximalen Punktanzahl gewinnt, der zweithöchste belegt den zweiten, der dritthöchste den dritten Platz. Dieser Ansatz verstärkt die positive, aktive Teilnahme erheblich und erlaubt darüber hinaus, aus falsch beantworteten Fragen direkt zu lernen. Zur weiteren positiven Verstärkung können beim ersten Quiz kleine Preise vergeben werden. Pro Quiz sollten nicht mehr als zwölf bis fünfzehn Fragen gestellt werden, damit der Spielspaß

V. Wich, A. Brunn

und die Kurzweiligkeit garantiert bleiben. Eine wohlüberlegte persönliche Moderation zwischen den Fragen kann die Motivation der Teilnehmer weiter erhöhen.

4.3 Autorentools für E-Learning-Module

Lectora (CHEMMEDIA AG 2018) wird als professionelle Autorensoftware verwendet, damit Wissen noch anschaulicher, intuitiver, besser eingebunden in Moodle und ansprechender vermittelt werden kann. Damit werden Inhalte interaktiver, haben mehr Finesse, sind ansprechender und durch die intelligente Nutzung des Mediums Internet schafft das besondere Begeisterung bei technikaffinen Lernenden. Im Internet etablieren sich mehr und mehr vereinfachte, mehrstufige Varianten, zur Deklarierung des Urheberrechts. Mit der Lizenz CC0 erstelltes geistiges Gut kann zum Allgemeinwohl überlassen und frei weiterverwendet werden. (CREATIVE COMMONS 2018)



Abb. 2: Lectora Lektion zum Mobile Mapping (MM). Dargestellt werden unterschiedliche MM Plattformen. Dieses sind Plattformen, welche Sensoren zu Luft, Wasser, Schiene und Land befördern können. Verwendet wurden Bilder, welche als gemeinfrei ausgezeichnet sind.

4.4 Lernvideos

Die Lernvideos sollen auf YouTube gehostet und in die jeweiligen Lektionen mit eingebunden werden. Dieses soll den Projektteilnehmen einen schnelleren Überblick über die einzelnen Teilbereiche verschaffen. Für YouTube spricht auch, dass die heutige Generation eine starke Affinität zum Medium hat. Zudem hat das den positiven Effekt der Außenwirkung, da sehr anschaulich visualisiert werden kann, in welchem technisch-innovativem Feld gearbeitet wird, was sehr viele technikaffine junge Menschen begeistert. Nahezu jeder mit einem Smartphone und Internet kann sehr einfach auf die Lektionen zugreifen. Die große Chance liegt darin, ein großes Publikum mit

verhältnismäßig wenig Aufwand zu erreichen. Außerdem bestünde die Möglichkeit den Kanal zu abonnieren, um auf dem Laufenden zu bleiben.

4.5 Netzwerkbildung

In der Zusammenarbeit mit kleinen und mittleren Unternehmen werden Best-Practice Leitfäden entwickelt. Diese ergeben sich aus dem vorgestellten theoretischen Wissen, der Anwendung des Wissens für den jeweiligen Teilbereich, aus der Erfahrung der Teilnehmer, und die Hochschule für angewandte Wissenschaften (FHWS) trägt ebenfalls dazu bei. Die so entstandenen Best-Practices werden in einem Wiki zusammengefasst und kontinuierlich verbessert. Relevante Forschungen und die praktischen Erfahrungen ergänzen diese.

Der wöchentlichen Newsletter erscheint als Benachrichtigung aus dem E-Learning heraus. Die Benutzer sehen somit direkt eine einfache Wochenübersicht in Textform, als Kurzfassung per E-Mail. Die Langfassung bekommen Sie, wenn Sie den Link zum PDF-Newsletter in Ihrer Benachrichtigung folgen. Darin erhalten Sie ausführliche Informationen bebildert, mit den entsprechenden weiterführenden Artikeln zu den jeweiligen Themen.



Abb. 3: Übersicht über den Newsletter vom 23.11.2018: Test einer Photogrammetriesoftware, ein weiteres nicht abgebildetes Thema dieses Newsletters war die Vorbereitung und Durchführung der Datenerfassung im Mobile Mapping.


Abb. 4: a) Mit Potree (Potree 2018) kann die Punktwolke im Internet auf den unterschiedlichsten Endgeräten im Browser betrachtet, darin gemessen oder digitalisiert werden. b) Target zur Platzierung des 3D Modells. c) 3D Modell der Stadt Würzburg über dem Target, in einer AR Androidapp, erstellt mit der Entwicklungsplattform Unity (Unity Technologies 2018) und der AR-Komponente Vuforia (PTC 2018).

Netzwerkbildung spielt im Projekt eine bedeutende Rolle. Daher werden bereits durchgeführte Veranstaltungen in Kurzform genannt. Diese dienen zudem auch noch zum Treffen der Teilnehmer, d.h. zur Stärkung der Bindung und dem Ausbau der persönlichen Ebene. Dies waren bisher: Ein informelles Treffen mit der VHB, ein Auftakttreffen den Projektteilnehmern aus MV-

3D4KMU und das 11. Anwenderforum Laserscanning wurde zu einem Netzwerktreffen genutzt. Die Treffen wurden immer auch dafür verwendet, um persönliche Kontakte zwischen den Teilnehmern zu knüpfen und so die Netzwerkbildung untereinander zu kräftigen.

Pro Jahr sind ca. drei bis vier Präsenztreffen und drei Webmeetings geplant. Bei Bedarf kann jederzeit telefonischer Kontakt, aber auch Kontakt per E-Mail hergestellt werden, um auf die individuellen Anliegen direkt eingehen zu können.

Die Beschleunigung im Fach setzt voraus, dass die zunehmende Fülle der Aufgaben besser bewältigt werden kann. Teilnehmer, die sich gegenseitig näher kennen und eine gute Einschätzung über die Fähigkeiten der anderen haben, können so entstandene Synergie Effekte aus dem Projekt nutzen und diese Art des Zusammenwirkens optimal für ihre Belange einsetzen. Die Projektteilnehmer kommen aus unterschiedlichen Fachrichtungen: Der klassischen Vermessung, über modernere 3D Vermessung, aber auch aus dem Planungs- und Visualisierungsbereich sind Partner vertreten. ARNOLD ET AL. 2017 (S. 63) zeigen in einem Modellversuch im MINT-Bereich, dass sich gerade Partner mit komplementären Eigenschaften besonders gut in der gemeinsamen Zusammenarbeit ergänzen können, was sich bisher im Projekt z.B. in der Ideenfindung bestätigte.

4.6 Anwendungen

Der ganze Weg von der Datenaufnahme in der mobilen 3D Vermessung bis hin zur Visualisierung mittels Augmented Reality (AR) Anwendung wird Stück für Stück abgebildet. Erste Ergebnisse stehen im Projekt schon jetzt zur Verfügung: Drei Mobile Mapping Befahrungen wurden durchgeführt, eine App mit 3D Stadtmodell und ein Web-Punktwolkenviewer ist implementiert, daneben gibt es eine Software mit der die Projektteilnehmer Ihre Punktwolken einfach in eine Webanwendung konvertieren können.

Die Anwendungen zur Visualisierung mit Entwicklungsprozess und ihrer Verteilungsart sind im Projekt Schritt für Schritt beschrieben: Webtechnologien mit Veröffentlichung im Internet, als auch Apptechnologie zur Verteilung auf Android Endgeräten. Die Teilnehmer haben freien Zugang zu den erarbeiteten Demo-Anwendungen.

Das Thema Aus- und Weiterbildung wird somit vereinfacht und auf den neuesten Stand der Technik gehoben. Sehr viele Inhalte der mobilen 3D Vermessung, stehen den Teilnehmern im Projekt zentral zur Verfügung. Somit müssen sich die Firmen weniger um die Recherche von Inhalten der Fachwissensvermittlung im Thema kümmern. Es stehen Inhalte zur Navigation, der Photogrammetrie, dem Laserscanning und dem Zusammenspiel der Techniken im Mobile Mapping zur Verfügung, damit wird das lehrende Personal in kleinen und mittelständischen Unternehmen entlastet. Dies bringt eine Zeitersparnis für den Ausbilder mit sich, die Ausbildung neuer Mitarbeiter kann damit schneller erfolgen. Ausbilder können mehr Zeit für individuelle Betreuung verwenden. Wir hoffen damit werden Ressourcen in Aus- und Weiterbildung besser genutzt und somit wird dem Fachkräftemangel etwas entgegengewirkt.

5 Fazit & Ausblick

Aus technischer Sicht setzt sich der Trend zur IT, also auch zur Automatisierung und der damit noch einfacher möglichen Erfassung großer Datenmengen weiter fort. Die Entwicklung wird zu Systemen gehen, welche stärker im Bereich der Massendaten angesiedelt sind. Das impliziert wiederum eine Rückkopplung bezüglich der größeren Schnelllebigkeit, stärkeren Komplexität der Systeme aus der Sache selbst heraus.

Ein weiterer Trend zeigt, auch wenn die Systeme im Aufbau komplexer werden, wird erfahrungsgemäß die Bedienung im Zuge neuer technischer Möglichkeiten vereinfacht, bei gleichzeitig mehr verfügbarer Rechenleistung und sinkenden Kosten pro Rechenoperation. Hier liegt die große Chance vieler neuer Möglichkeiten. Für den Endanwender bedeutet das in der Regel mehr Produktivität bei ähnlichem oder weniger Zeiteinsatz für die gleiche Aufgabe und somit eine wirtschaftlichere Arbeitsweise sowie einen Vorteil am Markt.

Das Bildungswesen muss sich dem Trend der Schnelllebigkeit und immer stärkeren IT-Affinität angleichen. Massive Open Online Courses (MOOCs) haben eine große Reichweite, mit der richtigen Anwendung des Konzepts ließen sich so Inhalte nachhaltiger und für ein großes Publikum, mit der entsprechenden Plattform qualitativ hochwertig realisieren. Solche Kurse könnten sich, entsprechend skaliert, finanziell selbst tragen. Das würde den Verantwortlichen die Freiheit bieten sich auf das Kernthema des Kurses zu fokussieren, mit der Verantwortung den Kurs kontinuierlich zu verbessern. Auch VARDI (2012) beschreibt die Bedeutung bei sich wiederholenden Kursen, begleitet von Online Quizzen und von sozialen Netzwerkaktivitäten. Er beleuchtet aber auch, dass Bildungsträger selbst unter finanziellen Druck geraten könnten, wenn sich MOOCs weiter durchsetzen, insbesondere in der weiterführenden Erwachsenenbildung.

Die Qualität und Quantität von insgesamt verfügbaren online Materialien ist gestiegen. Dies sind im Näheren: Hervorragende Webseiten zu wissenschaftlichen Themen, Veröffentlichungen, Daten- und Themenansammlungen und Videomaterial (HANDKE 2017: 36).

Die Wissensverbreitung, Transparenz in der Bildung und Forschung, sowie die Stärkung der Reputation nach außen hin ist möglichst zu begünstigen. Verstärken könnten dies darüber hinaus möglichst einfach verständliche Kurzfassung und Arbeitsbeispiele. Diese sollten den Arbeiten vorangestellt werden. Die öffentliche Verteilung von Kursen ist einfacher, wenn mögliche Urheberrechtsstreitigkeiten gar nicht erst aufkommen. ARNOLD (2011: 367) meint: "Zugleich zeigt sich aber auch, dass die Rechtsgrundlagen entsprechend den vielfältigen neuen Möglichkeiten im Internet, das in der Gesellschaft erarbeitete Wissen öffentlich allen Interessierten zugänglich zu machen, ergänzt, verändert und weiterentwickelt werden müssen. Open Content, Open Access, OER sind solche neuen Möglichkeiten öffentlicher Zugänglichkeit und Bearbeitung, die auch zum Vorteil der Erstellung und Bearbeitung von Onlinebildungsangeboten genutzt werden können."

Hier handelt es sich oft um Sammlungen von frei nutzbaren Inhalten. Bei Lektionen sollte auf Open Educational Resources (OER) in der Erstellung zurückgegriffen werden, um bei freier Veröffentlichung Urheberrechtschwierigkeiten zu vermeiden. So wird Hochwertigkeit und Aktualität, aber auch Bildung in der Fläche, mit neuesten Inhalten der mobilen 3D Vermessung möglich.

6 Danksagung

Die Autoren bedanken sich bei den Projektpartnern aus MV-3D4KMU für die Unterstützung dieses Artikels. Der besondere Dank gilt hierbei der Firma 3D Betrieb GmbH (Würzburg), für die Bereitstellung eines umfassenden 3D Modells der Stadt Würzburg, zur Visualisierung in der Augmented-Reality-App (siehe Abb. 4c).

Dieser Beitrag wurde innerhalb des Projekts MV-3D4KMU durch die Europäische Union und den Freistaat Bayern im Rahmen des ESF-Programms 2014-2020 "Perspektiven in Bayern – Perspektiven in Europa" in Zusammenarbeit mit der Virtuellen Hochschule Bayern (vhb) unterstützt.

7 Literaturverzeichnis

- ARNOLD, M., ZAWACKI-RICHTER, O., HAUBENREICH, J., RÖBKEN, H. & GÖTTER, R., 2017: Entwicklung von wissenschaftlichen Weiterbildungsprogrammen im MINT-Bereich. Waxmann Verlag, New York.
- ARNOLD, P., 2011: Handbuch E-Learning: Lehren und Lernen mit digitalen Medien, 2. erweiterte, aktualisierte und vollständig überarbeitete Auflage. Bertelsmann Verlag, Bielefeld.
- BRUNN, A., 2018: Blended Learning in Photogrammetrie und Bildverarbeitung für regionale KMUs - Ein Erfahrungsbericht. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 27, 38. Wissenschaftlich-Technische Jahrestagung der DGPF, 450-460.

- CENDON, E., MÖRTH, A. & PELLERT, A., 2016: Theorie und Praxis verzahnen: Lebenslanges Lernen an Hochschulen, Ergebnisse der wissenschaftlichen Begleitung des Bund-Länder-Wettbewerbs Aufstieg durch Bildung: Offene Hochschulen. Waxmann Verlag, New York.
- CHEMMEDIA AG 2018: Lectora E-Learning Authoring, https://www.lectora.de, letzter Zugriff: 20.12.2018.
- CREATIVE COMMENS 2018: https://creativecommons.org, letzter Zugriff 21.12.2018.
- EULER, D. & HOLZ, H., 1992: Didaktik des computerunterstützten Lernens: praktische Gestaltung und theoretische Grundlagen, 1. Auflage, Reihe Multimediales Lernen in der Berufsbildung. BW-Verlag, Nürnberg.
- HANDKE, J., 2017: Handbuch Hochschullehre Digital: Leitfaden für eine moderne und mediengerechte Lehre, 2. überarbeitete Auflage. Tectum Verlag, Baden-Baden.
- HOLTEN, R. & NITTEL, D., 2010: E-Learning in Hochschule und Weiterbildung: Einsatzchancen und Erfahrungen, Erwachsenenbildung und lebensbegleitendes Lernen. Bertelsmann, Bielefeld, 176.
- KERRES, M., 2012: Mediendidaktik: Konzeption und Entwicklung mediengestützter Lernangebote, 3. vollständig überarbeitete Auflage. Oldenbourg Verlag, München.
- KLIMSA, P. & ISSING, L.J., 2011: Online-Lernen: Handbuch für Wissenschaft und Praxis, 2. verbesserte und ergänzte Auflage. Oldenbourg Verlag, München.
- LIMA, L.C. & GUIMARÃES, P., 2011: European strategies in lifelong learning: a critical introduction, Study guides in adult education. Barbara Budrich Publishers, Leverkusen.
- MEYER, T. & BRUNN, A., 2016: Digitale Photogrammetrie und Bildverarbeitung für kleine und mittlere Unternehmen durch "Blended Learning". Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **25**, 436-448.
- OBERBECK, H. & KUNDOLF, S., 2018: Mobiles Lernen für morgen: Berufsbegleitende, wissenschaftliche Aus- und Weiterbildung für die Ingenieurwissenschaften. Waxmann Verlag, New York.
- PLATON, 1989: Der Staat. Über das Gerechte. Übersetzt und erläutert von Otto Apelt. Herausgegeben von Karl Bormann. Einleitung von Paul Wilpert. PhB 80. 11., 28 durchgesehene Auflage, Hamburg: Felix Meiner Verlag; zitiert nach: SCHMID-ISLER, S., 2009: Platons Kritik der Mimesis in der Politeia, Seminararbeit der Universität Luzern, <u>https://www.researchgate.net/profile/Salome_Schmid-Isler/publication/259297060_Salome_Schmid-Isler_Platons_Kritik_der_Mimesis_in_der_Politeia/links/0046352ae0cc4a7cc7000000/Salome-Schmid-Isler-Platons-Kritik-der-Mimesis-in-der-Politeia.pdf, letzter Zugriff: 25.12.2018.</u>
- POTREE 2018: https://www.portree.org, letzter Zugriff 21.12.2018.
- PTC 2018: https://www.vuforia.com, letzter Zugriff 21.12.2018.
- UNITY TECHNOLOGIES 2018: https://unity3d.com, letzter Zugriff 21.12.2018.
- VARDI, M.Y., 2012: Will MOOCs destroy academia? Communications of the ACM, November, **55**(11), 5.

E-Learning und Learning Analytics in der universitären Ausbildung der Geowissenschaften -Blended Learning im Masterstudium

MIKE TEUCHER¹ & CORNELIA GLÄßER¹

Zusammenfassung: Im Projekt "Prolearn" wird der Einsatz von E-Learning und E-Teaching Instrumenten in der geowissenschaftlichen Hochschullehre untersucht. Dabei stehen Lernmodule und Learning Analytics Methoden im Fokus, um den Lernerfolg/-fortschritt der Studierenden zu analysieren.

Für die Erstellung der Online-Inhalte, die ergänzend zur Veranstaltung angeboten werden, wird ein System aus 2 Komponenten, bestehend aus der Plattform ArcGIS Online und dem universitätseigenen Learning-Management-System ILIAS verwendet. Durch kurze Tests, kombiniert mit übergeordneten Erhebungen wird der Lernerfolg für die Dozierenden sicht- und durch die direkte Rückmeldung für die Teilnehmer quantifizierbar. Übergeordnete Zielstellung ist, ergänzend zu den fachlich-inhaltlichen Aspekten die Integration eines Empfehlungssystems zur individuellen Reduzierung von Wissensdefizite.

1 Einleitung

Das Angebot digitaler Lehre an deutschen Hochschulen wächst seit den 1990er Jahren stetig. Auch der Stellenwert digitaler Lehre an Hochschulen wird überdurchschnittlich eingestuft (WANNEMA-CHER 2016), weshalb E-Learning-Systeme zunehmend an Bedeutung für Bildungseinrichtungen und Unternehmen gewinnen (WANNEMACHER 2016; PRINZ et al. 2017). Die Fachgruppe Geofernerkundung und Kartographie des Instituts für Geowissenschaften und Geographie der Martin-Luther-Universität Halle-Wittenberg weist mit den Projekten WEBGEO und GEOVLEX im Bereich Blended Learning seit mehr als 10 Jahren Erfahrungen in diesem Bereich auf (THÜRKOW 2009, 2014).

Als konsequente Weiterentwicklung der bestehenden Konzepte, untersucht die Fachgruppe im Projekt "Prolearn" den Einsatz von prototypischen E-Learning und E-Teaching Instrumenten in der geowissenschaftlichen Hochschullehre (TEUCHER 2018). Dabei stehen Lernmodule und Learning Analytics Methoden im Fokus, welche detaillierte Informationen über das Lernerlebnis von Studierenden, u.a. beim Übergang vom Bachelor- zum Masterstudium liefern. Ziel ist es das studentische Vorwissen zu analysieren, Wissenslücken zu identifizieren und neue Lernwege aufzuzeigen.

¹ Martin-Luther-Universität Halle-Wittenberg, Institut für Geowissenschaften und Geographie, Fachgruppe Geofernerkundung & Thematische Kartographie, Von-Seckendorff-Platz 4, D-06120 Halle/Saale, E-Mail: [mike.teucher, cornelia.glaesser]@geo.uni-halle.de

2 Problemstellung

Diese Instrumente werden aktuell für eine ausgewählte Veranstaltung im 1. Fachsemester des Masterstudiums, die durch eine sehr heterogene Zusammensetzung hinsichtlich der Studiengänge und des Vorwissens charakterisiert ist, ergänzend zur Präsenzlehre umgesetzt. Die Inhalte orientieren sich am Wissenstand der Teilnehmer, der zu Beginn des Semesters durch 2 Vorab-Befragungen erhoben wird. Diese umfassen allgemeine Angaben über die Teilnehmer (u.a. Bachelorstudiengang, besuchte themenverwandte Veranstaltungen, EDV-technische Ausstattung und Zugangsmöglichkeiten), sowie eine Überprüfung der fachlichen Kenntnisse in den Bereichen Geoinformationssysteme, Fernerkundung und Kartographie.

Für die Erstellung der jeweiligen Online Lerneinheiten, die parallel zur Vorlesung angeboten werden, wird ein System aus 2 Komponenten eingesetzt. Für die thematische Bearbeitung der Inhalte wird zum Großteil die Plattform ArcGIS Online der Firma ESRI, als Teil des Campus Systems der Universität genutzt (ESRI 2018). Für die Präsentation der Inhalte entsprechend des ECLASS Ansatzes nach GERSON (2000) und um im Laufe des Projektes den individuellen Lernerfolg/-fortschritt der Teilnehmer erfassen zu können, werden die Inhalte und quiz-artige Tests im universitätseigenen Learning-Management-System ILIAS integriert. Durch kurze Tests zum Ende einer jeden Lerneinheit, kombiniert mit zweimaligen übergeordneten Erhebungen im Laufe des Semesters, wird der Lernerfolg für die Dozierenden sicht- und durch die Vergabe von Kompetenzpunkten der Lernerfolg für die Teilnehmer quantifizierbar. Die Entwicklung des Kompetenzmanagements greift auf eine systeminterne Funktion des ILIAS-Systems zurück und ermöglicht die modulübergreifende Verwendung, um den Lernerfolg entsprechend der definierten Wissensbereiche zu erfassen und kumulativ im Laufe des Semesters zu evaluieren.

Weiterhin wird durch diesen Blended-Learning-Ansatz eine individuellere Betreuung der Studierenden durch selbständiges Lernen anhand der entwickelten E-Learning-Module ermöglicht. Der etablierte Ansatz und seine einzelnen Komponenten sind in Abbildung 1 veranschaulicht.



Abb. 1: Verzahnung der E-Learning Komponenten mit Präsenzlehre

3 Fazit & Ausblick

Die Ergebnisse der Vorab-Befragung liefern wichtige Informationen über den aktuellen Wissenstand der Teilnehmer und die Bereitschaft zur Nutzung der Online-Lerneinheiten. Abbildung 2 zeigt als Ergebnis der Vorab-Befragung die Einstellung der Teilnehmer gegenüber Online Lernumgebungen. 74% der Befragten sind an einer Evaluation des eigenen Lernerfolgs/-fortschritt während der Veranstaltung interessiert.



Abb. 2: Erwartungen und Einstellung der Befragten gegenüber Online Lernumgebung

69,23% der Befragten schätzten Ihren Wissensstand zum Thema GIS als "Anfänger(-in) mit geringen Vorkenntnissen" ein, 26,92% wählten die Kategorie "Fortgeschrittene(r) mit sicherer Beherrschung der Grundfunktionen" und lediglich 3,85% sind laut eigener Einschätzung "Anfänger(in) ohne Vorkenntnisse". Die Überprüfung dieser Einschätzung mittels 5 Detailfragen ergab, dass im Bereich GIS Befragte der Kategorie "Anfänger(-in) mit geringen Vorkenntnissen" durchschnittlich 55,56% und 51,19% der erreichbaren Punkte durch die Teilnehmer der Kategorie "Fortgeschrittene(r) mit sicherer Beherrschung der Grundfunktionen" erzielten. Die Aufgabenschwierigkeit nach BORTZ & DÖRING (2006) aller Detailfragen deckte den Bereich von 0 bis maximal 0,88 ab, wobei 53% der Fragen auf den Bereich von 0,2 bis 0,8 entfallen, 40% auf den Bereich von 0,8 bis 1,0 und lediglich 7% den unteren Bereich betreffen. Die gemittelte Schwierigkeit über alle Fragen beträgt 0,5. Für den Bereich GIS beträgt die durchschnittliche Schwierigkeit der Fragen 0,57, mit einem Minimalwert von 0,36 und einem Maximalwert von 0,88. Für die Bereiche Fernerkundung und Kartographie liegen die durchschnittlichen Werte bei 0,31 (0,00-0,88), bzw. 0,57 (0,44-0,88). Tabelle 1 gibt eine Übersicht der gewählten Kategorien für die Bereiche Fernerkundung und Kartographie.

Kategorie	Fernerk	undung	Kartographie		
	Selbst-ein-	Ø Erreichte	Selbst-ein-	Ø Erreichte	
	schätzung	Prozent in	schätzung	Prozent in	
		Detailfragen		Detailfragen	
Anfänger(-in) ohne Vorkenntnisse	80,00%	32,00%	-	-	
Anfänger(-in) mit geringen Vorkenntnis- sen	20,00%	28,00%	56,00%	59,92%	
Fortgeschrittene(r) mit sicherer Beherr- schung der Grundfunktionen	-	-	44,00%	52,53%	
Experte/Expertin: Sie beherrschen auch anspruchsvollere Funktionen	-	-	-	-	

Tab. 1:	Einschätzung	des V	Vissensstandes	s in der	n Bereichen	Fernerkundung	g und l	Kartogra	phie

Die Ergebnisse der ersten übergeordneten Erhebung, die etwa zur Mitte des Wintersemesters durchgeführt wurde, weisen eine gemittelte Schwierigkeit aller Fragen von 0,52 auf, wobei 84% der Fragen im Bereich zwischen 0,2 und 0,8 liegen und jeweils 8% auf den unteren, bzw. oberen Bereich entfallen. Die Erhebung, bestehend aus 25 Fragen, gliedert sich in je 7 Fragen aus dem Bereich GIS, bzw. Fernerkundung und 11 Fragen aus dem Bereich Kartographie. Es zeigt sich, dass es eine Verbesserung im Bereich Fernerkundung gegeben hat, da die Aufgabenschwierigkeit im Vergleich zur Vorab-Befragung von 0,31 auf den 0,52 gestiegen ist. Dies bedeutet, dass gegenüber der Vorab-Befragung die Teilnehmer der ersten übergeordneten Erhebung nun 52,00% der Fragen richtig gelöst haben. Für die Bereiche GIS und Kartographie liegen die Werte bei 0,45 (-0,12), bzw. 0,58 (+0,01). Eine genauere Analyse der Ergebnisse aus dem Bereich GIS ergab, dass vor allem Fragen zum Basiswissen Geographischer Informationssysteme Probleme bereiteten, wohingegen Detailfragen zu den behandelten Inhalten der Präsenz- und Online-Lerneinheiten überwiegend gute Ergebnisse erzielten. Dies ist eine mögliche Erklärung für den Rückgang in diesem Bereich im Vergleich zur Vorab-Befragung. Ähnliche Tendenzen sind für den Bereich Kartographie zu erkennen, da auch hier die Detailfragen bessere Ergebnisse lieferten als Fragen zum Basiswissen Kartographie. Im Gegensatz hierzu wurde deutlich, dass im Bereich Fernerkundung primär die Fragen zum Basiswissen im Zwischentest bessere Ergebnisse erzielten, als Detailfragen, die die Inhalte der Veranstaltung thematisierten.

Zusätzlich zur Vorab-Befragung und einer ersten übergeordneten Erhebung wurden alle Online-Lerneinheiten mit kleineren quiz-artigen Test beendet. Abbildung 3 zeigt die Beteiligung (orange Linie) und die jeweils erreichten Prozent der Teilnehmenden (blaue Säulen).

Es wird ersichtlich, dass über alle Online-Lerneinheiten ca. 56% der Punkte erreicht wurden. Die Maximalwerte von 86%, bzw. 83% wurden bei den Online-Lerneinheiten zu den Themen "Landsat" und "Räumliche Analyse und Interpolation" erzielt, wohingegen die niedrigsten Ergebnisse im Themenkomplex "Veränderungsanalysen mit Fernerkundungsdaten" erzielt wurden. Auch hier zeigt sich, dass im Bereich Fernerkundung grundlegendes Wissen bessere Ergebnisse erzielten als spezifische Anwendungen. Eine weitere Erklärung für die niedrigen Werte der Online-Lerneinheit 6 liegt wohlmöglich in der (technischen) Komplexität der verwendeten Anwendung selbst, da in



dieser Einheit die Teilnehmenden das Thema mit Hilfe einer extern vorkonfigurierten Webanwendung erschließen sollten.

Abb. 3: Ergebnisse der Online-Lerneinheiten

Zusammenfassend lässt sich festhalten, dass anhand der oben dargestellten Ergebnisse das System eine individuelle Identifizierung vorhandener Wissensdefizite und eine Differenzierung des Wissensstands der Teilnehmenden in die Kategorien "Basiswissen" und "erweiterte Kenntnisse" ermöglicht. Die Ergebnisse zeigen aber auch, dass eine solide Bewertung des Lernerfolgs nur bei einer ausreichend hohen und konstanten Beteiligung möglich ist. Diesem Problem konnte im ersten stattgefundenen Durchgang noch nicht entgegnet werden, stellt aber einen zentralen Fokus für die weiteren Projektschritte dar. Des Weiteren sehen 32% keine erhöhte Arbeitsbelastung durch die Verwendung einer Online-Lernumgebung.

Die abschließende Evaluation der Veranstaltung durch die Teilnehmer ergab, dass 60% das digitale Angebot der Veranstaltung regelmäßig nutzten, 80% der Teilnehmer die modularen Online-Lerneinheiten hilfreich zum Verständnis des Lernstoffes empfanden und 60% die technische Handhabung des multimedialen Angebotes gut bis sehr gut einschätzten.

Eine weitere übergeordnete Zielstellung des Projektes "Prolearn" ist ergänzend zu den fachlichinhaltlichen Aspekten, die eingangs aufgezeigt werden, die technische Integration einer umfangreichen Funktion zur Visualisierung des individuellen Lernerfolg/-fortschritt im System ILIAS selbst. Die Möglichkeiten dieser Integration werden in den kommenden Projektschritten analysiert und bewertet. Vorstellbar zum aktuellen Zeitpunkt ist eine "Dashboard"-ähnliche Darstellung des individuellen Fortschrittes. Als zentrale Funktionalität dieses System ist die Erstellung eines semi-/automatisierten Empfehlungssystems von Lerneinheiten oder Literatur zur individuellen Reduzierung der Wissensdefizite der Teilnehmenden.

Weitere Ergebnisse zur Erfassung des Lernerlebnis und Informationen zur Usability und Performanz des neuen Kompetenzmanagementssystems werden zur DLT 2019 vorgestellt.

4 Literaturverzeichnis

- BORTZ, J. & DÖRING, N., 2013: Forschungsmethoden und Evaluation. Berlin, Heidelberg, Springer-Verlag (Hrsg.).
- ESRI, 2018: ArcGIS Online. Unter: http://www.arcgis.com/features/index.html. (abgerufen am 20. Dezember 2018)
- GERSON, S. M., 2000: E-CLASS: Creating a Guide to Online Course Development for Distance Learning Faculty. Online Journal of Distance Learning Administration, **3**(4).
- PRINZ, C.; KREIMEIER, D. & KUHLENKÖTTER, B., 2017: Implementation of a Learning Environment for an Industrie 4.0 Assistance System to Improve the Overall Equipment Effectiveness. Procedia Manufacturing, 9, 159-166.
- THÜRKOW, D.; GLÄBER, C.; SCHEUER, S. & SCHIELE, S., 2009: Visualisation of Hydrological Processes with GEOVLEX: Introduction of an integrated methodical-technical Online Learning Approach. Tools and Techniques for E-Learning, The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 38(6/W7), 5 p.
- THÜRKOW, D.; 2014: H2o EXTREM interaktives Lernen. Moderne Lehre gestalten, anwenden motivieren - vernetzen. Tagungsband zur Jahrestagung für multimediales Lehren und Lernen, Zentrum für multimediales Lehren und Lernen der Martin-Luther-Universität Halle-Wittenberg, 07.11.2014, Halle (Saale).
- TEUCHER, M., GLÄBER, C. & SCHREINER, V., 2018: E-Learning und Learning Analytics in der universitären Ausbildung der Geowissenschaften. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF) e.V., Band 27, Kersten, T.P., Gülch, E., Schiewe, J., Kolbe, T.H., Stilla, U. (Hrsg.), PFGK18 (Photogrammetrie Fernerkundung Geoinformatik Kartographie 2018), 07.-09. März 2018 in München, Tagungsband, 465-470.
- WANNEMACHER, K., 2016: Organisation Digitaler Lehre in den Deutschen Hochschulen. Arbeitspapier Nr. 21, Geschäftsstelle Hochschulforum Digitalisierung (Hrsg.), Berlin, 56 p.

Multidisziplinäre Überlegungen zur nachhaltigen N-Düngung unter Berücksichtigung der Möglichkeiten der satellitengestützten Präzisionslandwirtschaft

FRANCESCO VUOLO¹, LAURA ESSL¹, TARU SANDÉN² & HEIDE SPIEGEL²

Im Forschungsprojekt FATIMA wurde ein integrierter Ansatz für die Optimierung des Stickstoffmanagements gewählt. Neben der technischen Komponente, die die Verarbeitung von Sentinel-2 Satellitendaten umfasste, wurden auch die sozialen und ökonomischen Aspekte eines verbesserten Nährstoffmanagements untersucht. Die Ergebnisse zeigen, dass LandwirtInnen durchaus an neuen Technologien zur teilflächenspezifischen Stickstoffausbringung interessiert sind. Auch wirtschaftlich lässt sich eine Reduktion der Stickstoffmengen argumentieren. Im Zuge des Projekts wurde nicht nur ein Feldversuch für die Validierung der Satellitendaten und als Grundlage einer ökonomischen Analyse angelegt, sondern auch eine funktionierende Online-Anwendung für teilflächenspezifische Düngung basierend auf aktuellen und historischen Satellitenbildern entwickelt und getestet.

1 Einleitung

Eine Optimierung des Stickstoffmanagements hat positive Auswirkungen auf die unmittelbare Umwelt und trägt zum Klimaschutz bei. Die Ergebnisse des aktuellen Grundwassermonitorings des Umweltbundesamtes zeigen, dass gerade in den landwirtschaftlich genutzten Regionen Ostösterreichs die Nitratwerte über dem für Trinkwasser festgesetzten Wert von 50 mg/l liegen und diese Gebiete als voraussichtliche Maßnahmengebiete oder Beobachtungsgebiete ausgewiesen werden (UMWELTBUNDESAMT 2016). Empfohlene Maßnahmen, die den Düngereintrag reduzieren sollen, sind verbesserte Düngeplanung, -ausbringung und Bilanzierung basierend auf Bodenproben.

Neben verstärkter Bewusstseinsbildung und finanziellen Anreizen für ein verbessertes Stickstoffmanagement im Zuge der Gemeinsamen Agrarpolitik der Europäischen Union, ermöglichen auch neue Technologien, die mit Hilfe von Sensoren, Fernerkundungs- oder Bodendaten arbeiten, das Nährstoffmanagement effizienter zu gestalten.

Das H2020-Forschungsprojekt FATIMA (FArming Tools for external Inputs and water Management) hat zum Ziel, die durch den Start der Sentinel-2 Mission der European Space Agency verfügbaren Satellitendaten für verbessertes Ressourcenmanagement im Bereich Düngemanagement anzuwenden. Bisher gibt es nur wenige funktionierende Ansätze, die diese Daten in funktionierende Entscheidungshilfen für Landwirte umwandeln.

¹ Universität für Bodenkultur, Institut für Vermessung, Fernerkundung und Landinformation, Peter-Jordan Straße 82, A-1190 Wien, E-Mail: [francesco.vuolo, laura.essl]@boku.ac.at

² Österreichische Agentur für Ernährungssicherheit, Spargelfeldstraße 191, A-1220 Wien, E-Mail: [taru.sanden, adelheid.spiegel]@ages.at

2 Problemstellung

Bei der Frage nach dem optimalen Stickstoffdüngereinsatz spielen verschiedene Einflussfaktoren eine Rolle: wieviel Stickstoff bringt unter den gegebenen Standortbedingungen den optimalen Ertrag für eine bestimmte Kultur unter Berücksichtigung der Umwelterfordernisse und welche Anwendungsstrategie ist wirtschaftlich und ökologisch sinnvoll?

An diesem Punkt setzte das EU-Forschungsprojekt FATIMA an: neben der Entwicklung von neuen Technologien basierend auf Fernerkundungsdaten der Satelliten Sentinel-2 wurde auch die soziale und ökonomische Komponente der Einzug haltenden "Landwirtschaft 4.0" beleuchtet. Forscher der Universität für Bodenkultur und der AGES stellten sich den Fragen, wie Stickstoffmanagement verbessert werden kann und ob die Landwirte in Österreich bereit für diesen Wandel sind.

3 Methode

Als erste Komponente in dem dreistufigen Prozess (siehe Abbildung 1) wurde die Aufgeschlossenheit und das bestehende Wissen über satelliten- oder sensorunterstützte Düngung abgefragt. Die Ergebnisse bestätigten das Interesse der Landwirte und in einem zweiten Schritt wurden Versuchsflächen in der Region Marchfeld, welches als voraussichtliches Maßnahmengebiet eingestuft ist, angelegt. Der Feldversuch diente der Validierung der Satellitendaten, gab aber auch aufschlussreiche Informationen über die ökonomische Komponente variabler Düngermengen. In einem letzten Schritt wurde eine Anwendung entwickelt, um die teilflächenspezifische Düngung auf zwei unter echten Bedingungen bewirtschafteten Flächen zu testen und den Schritt von der Theorie in die Praxis zu wagen. Die Ergebnisse bieten ein umfassendes, integriertes Bild der Herausforderungen, aber auch Möglichkeiten von neuen Technologien im Bereich Düngemanagement. Die drei methodischen Komponenten werden in den folgenden Kapiteln genauer geschildert.



Abb. 1: Der Weg zu verbessertem Stickstoffmanagement im Projekt FATIMA

3.1 Soziale Komponente

Um herauszufinden, ob Österreich bereit für neue, teilflächenbasierte Lösungen im Bereich Stickstoffmanagement ist, startete Die Freie Universität Amsterdam (IVM) in Kooperation mit der Universität für Bodenkultur (BOKU) eine Umfrage unter niederösterreichischen Landwirten und Landwirtinnen. Die Umfrage wurde in persönlichen Interviews von Studierenden der BOKU im Zeitraum November 2016 – Februar 2017 mit Unterstützung der Landwirtschaftskammer Niederösterreich (NÖ) durchgeführt. Der Fragebogen bestand aus einem allgemeinen Teil mit Fragen zum Betrieb und einem Choice Experiment, in dem die Landwirte und Landwirtinnen aus verschiedenen innovativen Optionen des Düngemanagements wählen konnten.

Den Landwirten und Landwirtinnen wurden drei Kategorien von Technologien für teilflächenspezifischer Stickstoffausbringung vorgestellt:

- Voll automatisierte Technologien (z.B. auf dem Traktor angebrachter Sensor)
- Teilweise automatisierte Technologien (z.B. Applikationskarten zur direkten Verwendung am Bordcomputer)
- Nicht automatisierte Technologien (z.B. Kartenmaterial mit ausgewiesenen Zonen

Den TeilnehmerInnen wurden drei mögliche Optionen unter Berücksichtigung verschiedener Eigenschaften der einzelnen Technologien vorgestellt. So wiesen die drei vorgestellten Technologiegruppen unterschiedliche Kosten, Düngereinsparungspotenzial, Ertragssteigerungspotenzial, Beratungsangebot und Grundwasserverbesserungspotenzial auf. Den Landwirten und Landwirtinnen wurde sogenannte "*choice cards*" (siehe Abbildung 2) präsentiert, auf denen sie für ihren Betrieb am ehesten geeignete Lösung auswählen konnten. Auch die Wahl keines der drei Systeme war möglich. Aus den Antworten konnte die Investitionsbereitschaft, aber auch die Wichtigkeit der einzelnen Kriterien abgeleitet werden. Eine genauere Beschreibung der Methode findet sich bei Blasch (BLASCH et al. 2018).



Abb. 2: Beispiel einer Choice Karte mit den verschiedenen Optionen zum optimierten Düngemanagement An der Studie nahmen 242 niederösterreichische Landwirte und Landwirtinnen teil. Der Großteil der Landwirte und Landwirtinnen stammte aus den Regionen Marchfeld, Weinviertel, Wiener Becken und Tullnerfeld. Die Befragten waren zumeist EigentümerInnen (82%) von Familienbetrieben (95%). Die Betriebsgrößen bewegten sich etwas über dem niederösterreichischen Durchschnitt von 42 ha (Grüner Bericht, 2016).



Abb. 3: Teilnehmercharakteristika der Umfrage

Zwei Drittel der befragten Landwirte und Landwirtinnen bewirtschafteten mehr als 50 ha, 26% sogar mehr als 100ha. Die TeilnehmerInnen der Umfrage wiesen ein jüngeres Durchschnittsalter als der niederösterreichische Durchschnitt auf (siehe Abbildung 3). Die Landwirte und Landwirtinnen waren zudem meist TeilnehmerInnen der Arbeitskreise der Landwirtschaftskammer NÖ. Die Untersuchung der Charakteristika der UmfrageteilnehmerInnen ergibt einen im Vergleich zum Durchschnittsbetrieb (NÖ Agrarstrukturerhebung 2016) jüngeren, mit größerer Fläche ausgestatteten und an Neuigkeiten im Agrarsektor interessierten Betriebsführer oder Betriebsführerin.

3.2 Ökonomische Komponente

Im Laufe des Projekts FATIMA konnte in zwei Jahren der Einsatz verschiedener Düngemengen für Winterweizen auf Probeflächen getestet werden. Feldversuche mit Winterweizen (Sorte "Capo") wurden 2015/16 und 2016/17 auf dem Standort "Engelhartstetten" im Marchfeld (Niederösterreich) (48°11′N 16°55′E) auf zwei benachbarten Feldern durchgeführt. Die Beschreibung des Standortes (Klima und Boden) ist in Tabelle 1 dargestellt. Die Winterweizenversuche bestanden neben der Nulldüngungsvariante N0 (0 kg N) aus drei unterschiedlichen Düngungsstufen mit folgenden Gesamtdüngungsmengen (N1: 60 kg N; N2: 120 kg N; N3: 180 kg N), die als Kalkammonsalpeter gegeben und gleichmäßig auf drei Düngungstermine aufgeteilt wurden. Die Feldversuche bestanden aus einem randomisierten Blockdesign mit drei Wiederholungen, wobei die Parzellen jeweils 100 m x 100 m (1 ha) maßen. Ein Beispiel für die Anlage der Feldversuche (Winterweizen 2016) ist in Abbildung 1 dargestellt. Die Phosphor- und Kaliumdüngung wurde entsprechend den Richtlinien für die sachgerechte Düngung (BMLFUW 2006) durchgeführt. Die Bewirtschaftungsdaten der Anbauperioden 2015/16 und 2016/17 sind in Tabelle 2 dargestellt.

	Einheit	Engelhartstetten
Meter über NN	m	147
Mittlerer Jahresniederschlag (Ø1981-2010)	mm	516
Mittlere Jahrestemperatur (Ø1981-2010)	°C	10.3
Bodentyp (IUSS 2015)		Kalkhaltiger Tscher- nosem
Textur (Sand/Schluff/Ton)	%	(17.6/48.7/33.7)
pH-CaCl ₂		7.54
Karbonatgehalt	%	23.2
Organischer Kohlenstoff	%	2.36
Ntotal	%	0.24

Tab. 1: Beschreibung des Standortes Engelhartstetten

Tab. 2: Bewirtschaftungsdaten der Anbauperioden 2015/16 und 2016/17

Anbauperiode	2015/2016	2016/2017
Bodenbearbeitung	November 2015	9. und 11. November 2016
Saat	November 2015	15. November 2016
1. N Düngung	24. März 2016	16. März 2017
2 N Düngung	21. April 2016	25. April 2017
3 N Düngung	19. Mai 2016	30. Mai 2017
Ernte	26. Juli 2016	20. Juli 2017



Abb. 4: Schema des N Düngungsversuchs in Engelhartstetten (Marchfeld, Österreich) 2016. Hintergrundbild: Sentinel-2A Leaf Area Index (LAI) Karte vom 6. Mai 2016 Der Feldversuch diente neben der Validierung der Satellitendaten als Basis für die Bewertung der Wirtschaftlichkeit der unterschiedlichen Düngemengen. Unter Berücksichtigung der Kosten für Dünger, Arbeitszeit, Maschineneinsatz, den Marktpreisen für Weizen im Juni der beiden Jahre und dem mit dem eingesetzten Dünger erwirtschafteten Ertrag und den erzielten Qualitäten wurde eine ökonomische Analyse der Wirtschaftlichkeit der Düngemaßnahmen durchgeführt.

3.3 Technische Komponente

Die technische Komponente der Studie umfasste die Auswertung von Satellitendaten, um einerseits die wissenschaftlichen Aspekte der Projektaktivitäten zu unterstützen (VUOLO et al. 2017) und andererseits ein praktisch einsetzbares Werkzeug für die Stickstoffdüngung zu entwickeln. Letzteres wird in diesem Paper behandelt.

Für die praktische Anwendung wurden Sentinel-2 Daten verwendet, um den Blattflächenindex (Leaf Area Index LAI) während der gesamten Vegetationsperiode zu berechnen. Der LAI korreliert stark mit der vorhandenen Biomasse und eine Zusammenführung der Werte aus einer Vielzahl von Satellitenbildern zu verschiedenen Zeitpunkten gibt Aufschluss über die Produktivität der Kulturen unter den gegebenen Wachstumsbedingungen. Basierend auf dem LAI wurde ein einheitsloser Indikator für potentielle Produktivität entwickelt und für eine Optimierung der Stickstoffdüngung verwendet.

Außerdem wurde eine Online-Plattform für die automatische Erstellung der Managementzonen mit automatischer Segmentierung und der Ausgabe der jeweiligen Stickstoffdüngermengen entwickelt, in der die gewünschten Parameter der Düngegabe vom Nutzer individuell festgelegt werden können.

Die Satellitendaten und die LAI-Karten wurden aus dem Sentinel-2-Portal des Instituts für Vermessung, Fernerkundung und Landinformation (IVFL) der Universität für Bodenkultur bezogen (VUOLO et al. 2016). Die verarbeiteten Daten werden automatisch in einer Datenbank gespeichert und sowohl einzelne Bilder als auch die Produktivitätskarte stehen für die Erstellung der N-Applikationskarten zur Verfügung. Je nach Einsatzzeitpunkt können daher aktuelle oder historische Karten für die optimierte N-Düngung verwendet werden.

4 Ergebnisse

4.1 Soziale Komponente

Die TeilnehmerInnen wurden nach den für sie zurzeit größten Herausforderungen des Landwirtschaftssektors befragt. Die Marktbedingungen (1) wurden am öftesten genannt, gefolgt von den hohen bürokratischen Anforderungen (2) und extremen Wetterereignissen (3). Im Mittelfeld finden sich Themen mit Bezug zu Ressourcenverfügbarkeit wie Zugang zu Wasser (4) und Land (5). Umweltthemen, wie z.B. Bodendegradation und Stickstoffanreicherung im Grundwasser wurden nur von Wenigen als direkte Herausforderung für den Betrieb gesehen. Dennoch sehen es die Landwirte und Landwirtinnen als Notwendigkeit an, die Umwelt zu schützen und Maßnahmen für Umweltschutz am Betrieb zu treffen.

Um herauszufinden, ob eine generelle Innovations- und Investitionsbereitschaft besteht, wurden die UmfrageteilnehmerInnen gefragt, ob sie innerhalb der letzten fünf Jahre Investitionen getätigt

hatten. Zwei Drittel der Befragten hatten sowohl in verbessertes Management (z.B. Farm Management Software) als auch in bessere technische Ausrüstung investiert. Immerhin 40% konnten des Weiteren auf Innovationen im Bereich ihrer landwirtschaftlichen Praxis (z.B. Umstieg auf biologische Wirtschaftsweise oder reduzierte Bodenbearbeitung) verweisen. Diese Werte zeigen eine generell hohe Innovationsbereitschaft der TeilnehmerInnen.

Die Bereitschaft der TeilnehmerInnen, die bereits vorgestellten Optionen (voll- teilweise- oder nicht-automatisiert) auf ihrem Betrieb einzusetzen, wurde in einem Choice Experiment abgefragt. Alle drei Formen der teilflächenspezifischen Stickstoffdüngung, die vorgestellt wurden, waren den Landwirten und Landwirtinnen gut bekannt. Die Hälfte der Befragten gab an, den Erwerb von teilweise (51%) oder nicht (47%) automatisierten Technologien bereits in Erwägung gezogen zu haben. Das Interesse für die voll automatisierte Lösung war mit 14% weniger stark ausgeprägt.

Die Auswertung des Choice Experiments ergab, dass 75% der Landwirte und Landwirtinnen generell bereit dazu sind, Technologien, die der exakteren Ausbringung von Stickstoff dienen, anzuwenden. Die TeilnehmerInnen haben allerdings unterschiedliche Präferenzen bezüglich der einzelnen vorgeschlagenen Technologien. Die voll automatische Sensorlösung wurde von den TeilnehmerInnen generell abgelehnt. Bei den anderen beiden Technologiegruppen ergab sich ein differenziertes Bild.

Während Technologieoptionen mit höheren Erträgen und Düngereinsparpotenzial von den Landwirten und Landwirtinnen nicht durchgehend bevorzugt wurden, zeigten eine persönliche Beratung sowie eine potenzielle Verbesserung der Grundwasserqualität einen signifikant positiven Einfluss auf die Investitionsbereitschaft. Auch die Kosten der Technologie spielen eine Rolle bei Anschaffung. Ein hoher Anschaffungspreis stößt erwartungsgemäß auf Ablehnung. Auch die Größe des Betriebs spielte eine Rolle bei der Auswahl. Größere Betriebe sind eher an teil- bzw. vollautomatisierten Systemen interessiert.

Eine weitere Beobachtung konnte gemacht werden: Konventionell wirtschaftende Landwirte und Landwirtinnen sind an teilflächenspezifischer Ausbringung interessierter als biologisch arbeitende Landwirte und Landwirtinnen. Dies kann auf die limitierten technischen Möglichkeiten sowie auf alternative Nährstoffversorgung über Leguminosen zurückgeführt werden.

4.2 Ökonomische Komponente

Das Ziel von Düngemaßnahmen ist vorrangig die Erzielung eines höheren Ertrags und eines höheren Deckungsbeitrags. Die im Zuge des FATIMA-Projekt angelegten Versuchsflächen bieten hervorragende Bedingungen für eine Untersuchung der Wirtschaftlichkeit von variierenden Düngemengen.

Beim Vergleich der Stickstoffzufuhr und dem Ertrag, zeigte sich, dass erhöhter N-Einsatz nur bis zu einem gewissen Punkt zu höherem Ertrag führt. Ab einem gewissen Versorgungsgrad ist der Ertragszuwachs mit steigender Düngemenge nur noch gering und der Deckungsbeitrag verringert sich.



Abb. 5: N-Düngung und Ertrag der 4 Düngestufen

Abbildung 5 zeigt, dass die Ertragskurve mit erhöhtem Düngereinsatz abflacht. Ein ähnliches Bild zeigten bereits andere wissenschaftliche Feldversuche (HAWKESFORD 2014; WONG & ASSENG 2007).



Abb. 6: Einkommenseinbußen aufgrund verringerter N-Dünung im Vergleich zu 180kg N/ha

Die ökonomische Analyse zeigt den folgenden Trend: Beim gänzlichen Verzicht auf Stickstoffdünger bewegen sich die Einkommenseinbußen im Vergleich zur höchsten Stufe im Bereich von 192-253 €/ha, sind also durchaus beträchtlich. Anders zeigt sich die Situation beim Vergleich von 120 kgN/ha und 180kgN/ha. Die Unterschiede liegen im Bereich von 25-71€/ha, wobei sich besonders 2016 aufgrund niedriger Abnahme- und höherer Stickstoffpreise der erhöhte Düngereinsatz kaum auszahlte. Die Daten basieren auf der Evaluierung der Versuchsflächen und können daher kein allgemeingültiges Bild geben, dennoch zeigen sie, dass höherer Düngereinsatz nicht immer mit höherem Einkommen gleichzusetzen ist. Des Weiteren machen die Ergebnisse deutlich, dass die Einflussfaktoren eine große Rolle spielen. Nicht nur unterschieden sich Witterungsbedingungen und die aktuellen Marktpreise von Dünger und Weizen in den beiden Jahren, auch der erzielte Proteingehalt variierte innerhalb der beiden Jahre je nach Düngestufe beträchtlich.

Einflußfaktoren Wirtschaftlichkeitsberechnung



Abb. 7: Einflussfaktoren der Wirtschaftlichkeitsrechnung, Vergleich der beiden Jahre 2016 und 2017. Die Daten stammen von den Webseiten der AMA (Dünger) und der Landwirtschaftskammer NÖ (Getreidepreise) sowie aus eigenen Quellen (Niederschlag)

Die aufgrund der Richtlinien für die sachgerechte Düngung (BMLFUW 2017) empfohlene N-Düngermenge beläuft sich auf 110-130kg N/ha bei mittlerer Ertragserwartung. Um den N-Bedarf genau zu bestimmen, empfiehlt sich die Standorteigenschaften wie Boden- und Wasserverhältnisse, das N-Mineralisierungspotenzial, die tatsächliche Ertragserwartung und die Vorfrucht zu berücksichtigen. Ein einfach zu bedienendes Online-Tool (AGES Agrarcommander) wird auf der Website der AGES (www.ages.at/service/service-landwirtschaft/agrar-online-tools) bereitgestellt. Eine genauere Stickstoffdüngungsempfehlung ist auch auf der Basis von Frühjahrs-Nmin – Analysen möglich.

4.3 Technische Komponente

Abbildung 8 zeigt einen Ausschnitt der Produktivitätskarte, die für ganz Österreich erstellt wurde. Die Karte zeigt deutlich die räumlichen Muster, die für die Variabilität der Böden in Ostösterreich typisch sind. Ähnliche Strukturen sind auch Farb-Orthofotos der Region zu erkennen. Eine Validierung mit der Österreichischen Bodenkarte war aufgrund der großen Unterschiede in der Auflösung der beiden Karten nicht möglich. Alternativ dazu wurde eine Validierung auf Feldebene für zwei Felder vorgenommen.

Abbildung 9 zeigt den Vergleich zwischen den Werten der Produktivitätskarte und den Erntemengen der betreffenden Punkte für zwei Felder, die von Landwirten der Region unter konventionellen Bedingungen bewirtschaftet wurden. Die Ergebnisse zeigen eine positive Korrelation für beide Felder. Um noch aussagekräftigere Resultate zu erzielen, wird in Zukunft weitere Validierungsarbeit notwendig sein.

Die Produktivitätskarte wird Testnutzern in einer Online-Anwendung (Abbildung 10) zur Verfügung gestellt. Der Nutzer kann zwischen dem aktuellsten Satellitenbild oder der auf historischen Bildern basierenden Produktivitätskarte wählen, um die Zonen für teilflächenspezifische N-Düngung zu definieren. Die notwendigen Eingabeparameter sind die Mindestfläche der einzelnen Zonen, die Unter- und Obergrenze der Stickstoffgabe sowie die gewählte Düngestrategie.



Abb. 8: Vergleich zwischen der Produktivitätskarte und einem hochauflösendem Geoland Basemap-Orthofoto in einem landwirtschaftlich genutzten Gebiet im Marchfeld.



Abb. 9: Erntemengen (x-Achse) vs. Produktivität (y-Achse) für zwei Felder im Jahr 2017. Die Daten wurden für die einzelnen Managementzonen zusammengefasst.



Abb. 10: Screenshot der Online-Anwendung für die Erstellung der Karte für die teilflächenspezifische N-Düngung.

Abbildung 11 zeigt die Kriterien, die dem Nutzer für die Auswahl der gewählten Datengrundlage dienen. Dies hängt hauptsächlich von der Pflanzenentwicklung zum Zeitpunkt der Erstellung der Karte ab. Typischerweise erfolgt bei Wintergetreide die erste Düngergabe vor dem Aufkommen der Pflanze und die räumliche Variabilität ist noch nicht aufgrund der unterschiedlich entwickelten Biomasse sichtbar. In diesem Fall wird die Produktivitätskarte für die Erstellung der Applikationskarte herangezogen. Für die zweite und dritte Düngung können aktuelle Satellitenbilder verwendet werden, um auf die aktuellen Bedingungen zu reagieren. Zwei verschiedene Strategien sind für die teilflächenspezifische Düngung möglich: mehr Dünger in weniger entwickelten Zonen ("catch up") oder mehr Dünger in besser entwickelten Zonen ("top up"). In einem letzten Schritt können die Daten vom Endnutzer im benötigten Datenformat heruntergeladen werden und dann direkt an das Traktorterminal für eine teilflächenspezifische Düngung übertragen werden.

5 Fazit & Ausblick

Die Ergebnisse des Forschungsprojektes zeigen unter verschiedenen Perspektiven das Potential der teilflächenspezifischen Düngung.

Bei den Landwirten selbst ist die Technologie zumindest bekannt und das potenzielle Interesse für die Anwendung groß. Wie kann der Schritt von der Theorie bzw. dem Hörensagen in die Praxis auch für kleinere Landwirte gelingen?

Die Verwendung der Produktivitätskarten ist weder teuer noch kompliziert. Auch wenn der Traktor nicht mit GPS und Bordcomputer ausgestattet ist, können Landwirte die Karten auf ihre mobilen Endgeräte laden und bei der Fahrt über das Feld manuell mehr oder weniger Dünger ausbringen. Diese Applikation birgt besonders für kleine und mittelgroße Betriebe, die nicht in Sensortechnologie investieren wollen, großes Potential.

Die teilflächenspezifische Stickstoffdüngung basierend auf Produktivitätskarten, die mit Hilfe von Satellitendaten gewonnen werden, kann nicht nur helfen Erträge zu optimieren. Auch werden unerwünschte Verluste von Stickstoff in die Luft und in das Grundwasser mit diesen neuen Technologien möglichst geringgehalten.

Die zukünftigen Aktivitäten werden sich auf eine Validierung der Daten und eine leichtere Anwendbarkeit der Daten konzentrieren, um die wertvollen Informationen unter möglichst vielen Anwendern zu verbreiten.



Abb.11: A. Kriterien für die Auswahl der Datengrundlage für die Definition der Managementzonen und B. Düngestrategien ("catch-up" or "top up") für die Zuordnung der Düngemenge zwischen Oberund Untergrenze

Die Studie wurde im Rahmen des Projektes "FATIMA" (FArming Tools for external nutrient Inputs and water Management, http://fatima-h2020.eu/) durchgeführt. Das Projekt wurde vom Programm Horizon 2020 der Europäischen Union (GA Nr. 633945) finanziert.

6 Literaturverzeichnis

AMT DER NÖ LANDESREGIERUNG, 2016: Der Grüne Bericht 2016. St.Pölten.

- BLASCH, J., VAN DER KROON, B., VAN BEUKERING, P., MUNSTER, R., FABIANI, S., NINO, P. & VA-NINO, S., 2018: The role of social ties and knowledge-sharing for the adoption of precision farming technologies – A farmer survey and choice experiment in Italy. World Congress of Environmental and Resource Economists (WCERE), Gothenburg, Sweden, June 25-29.
- BMLFUW, 2017: Richtlinien für die sachgerechte Düngung im Ackerbau und Grünland. Anleitung zur Interpretation von Bodenuntersuchungsergebnissen. 7. Auflage. Fachbeirat für Bodenfruchtbarkeit und Bodenschutz. Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft, Wien.
- BMLFUW, 2006: Richtlinien für die Sachgerechte Düngung, 2006: Anleitung zur Interpretation von Bodenuntersuchungsergebnissen in der Landwirtschaft. 6. Auflage. Wien, Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft.
- HAWKESFORD, M.J., 2014: Reducing the reliance on notrogen fertilizer for wheat production. Journal of Cereal Science, **59**, 276-283.
- IUSS WORKING GROUP WRB, 2015: World Reference Base for Soil Resources 2014, update 2015. International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports, **106**, FAO, Rome.
- STATISTIK AUSTRIA, 2016: Agrarstrukturerhebung 2016. Statistik Austria StatCube, www.statcube.at.
- UMWELTBUNDESAMT, 2016: Elfter Umweltkontrollbericht. Wien, http://www.umweltbundesamt.at/umweltsituation/umweltkontrollbericht/ukb/.
- VUOLO, F., ŻÓŁTAK, M., PIPITONE, C., ZAPPA, L., WENNG, H., IMMITZER, M., WEISS, M., BARET, F. & ATZBERGER, C., 2016: Data Service Platform for Sentinel-2 Surface Reflectance and Value-Added Products: System Use and Examples. Remote Sensing, 8(11), 938.
- VUOLO, F., ESSL, L., ZAPPA, L., SANDÉN, T. & SPIEGEL, A., 2017: Water and nutrient management: the Austria case studyof the FATIMA H2020 project, Advances in Animal Biosciences, 8(2), 400-405.
- WONG, M. & ASSENG, S., 2007: Yield and environmental benefits of ameliorating subsoil constraints under variable rainfall in a Mediterranean environment. Plant and Soil, **297**, 29-42.

Vergleich unterschiedlicher Methoden zur Schätzung von Blattflächenindex in Dauergrünlandbeständen

ANDREAS KLINGLER¹, ANDREAS SCHAUMBERGER², FRANCESCO VUOLO³ & ERICH M. PÖTSCH²

Zusammenfassung: Der Blattflächenindex (LAI) stellt einen der vielversprechendsten biophysikalischen Parameter zur Beschreibung von Grünlandbeständen dar. Im Gegensatz zu einfachen Vegetationsindizes neigt der LAI nicht zur Sättigung in fortgeschrittenen Vegetationsstadien. In der vorliegenden Arbeit wird die Korrelation von LAI-Zeitreihen, berechnet aus Hyperspektralsignaturen eines Feldspektrometers, Sentinel-2-Daten und AccuPAR-Messungen, untersucht. Dazu wurden wöchentlich feldspektroskopische Erhebungen und AccuPAR-Messungen auf einer vierschnittigen Dauergrünlandwiese durchgeführt. Von derselben Fläche wurden sämtliche wolkenfreie Sentinel-2-Beobachtungen zum Vergleich mit den Bodenmessungen verwendet. Die berechneten LAI-Zeitreihen des Feldspektrometers und der AccuPAR-Methode zeigen generell eine starke Korrelation mit den Sentinel-2-Werten für die gesamte Dauer der Vegetationsperiode. Deutliche Unterschiede sind zwischen den einzelnen Aufwüchsen und Entwicklungsstadien festzustellen.

1 Einleitung

Die klimatischen Veränderungen während der letzten Jahrzehnte haben merkbare Unterschiede in der Wachstumsdynamik von Grünland zur Folge. Dazu kommen klimawandelbedingte Änderungen von Beginn und Dauer der Vegetationsperiode sowie Wetterextreme, welche eine große Herausforderung für die Grünlandwirtschaft darstellen. Um weiterhin ein möglichst optimales Grünlandmanagement gewährleisten zu können, bietet sich der Einsatz von bisher wenig genutzten Technologien zur Erfassung wachstumsrelevanter Kennwerte an. Der Blatt-flächenindex, engl. Leaf Area Index (LAI), definiert als die gesamte einseitige Blattfläche pro Bodenoberfläche (CHEN & BLACK, 1992), stellt einen der vielversprechendsten biophysikalischen Parameter dar, um wichtige Wachstumsprozesse wie Fotosynthese oder Biomasseakkumulation im Grünland beschreiben zu können (CLEVERS et al. 2017). Das Ziel der vorliegenden Untersuchung besteht darin, LAI-Zeitreihen eines Wirtschaftsgrünlandes berechnet aus a) Hyperspektralsignaturen, b) Sentinel-2-Daten und c) AccuPAR LP-80 Messungen zu vergleichen.

¹ Universität für Bodenkultur Wien, Abteilung Pflanzenbau, Konrad-Lorenz-Straße 24, A-3430 Tulln an der Donau, E-Mail: andreas.klingler@raumberg-gumpenstein.at

² HBLFA Raumberg-Gumpenstein, Abteilung Grünlandmanagement und Kulturlandschaft, Altirdning 11, A-8952 Irdning-Donnersbachtal, E-Mail: [andreas.schaumberger, erich.poetsch]@raumberggumpenstein.at

³ Universität für Bodenkultur Wien, Institut für Vermessung, Fernerkundung und Landinformation, Peter-Jordan-Straße 82, A-1190 Wien, E-Mail: francesco.vuolo@boku.ac.at

2 Material und Methoden

2.1 Untersuchungsgebiet

Die Versuchsfläche mit einer Größe von 4,6 ha liegt im steirischen Ennstal (47°30'35.4"N 14°05'03.5"E) auf 643 m Seehöhe. Die langjährige Durchschnittstemperatur (1981-2010) liegt bei 8,2 °C und der durchschnittliche Jahresniederschlag bei 1056 mm. Die Versuchsfläche, welche sich auf einem Augley-Boden befindet, gehört mit einer Nutzungshäufigkeit von vier Schnitten pro Jahr zu den intensiv genutzten Dauergrünlandflächen. Auf den in Abb. 1 gekennzeichneten Bereichen 1-3 wurden in wöchentlichen Intervallen über den Zeitraum einer ganzen Vegetationsperiode Erhebungen durchgeführt. Mithilfe eines GPS-Gerätes mit Genauigkeiten im Submeterbereich (DGPS) wurden die drei Bereiche zu Beginn eines jeden Aufwuchses mit Markern gekennzeichnet. Um möglichst repräsentative Flächen für die Bodenmessungen zu erhalten, wurde besonders auf die botanische Homogenität der Flächen geachtet.



Abb. 1: Versuchsfeld (4,6 ha) mit den Messbereichen 1 – 3 (20 x 20m)

2.2 Eingesetzte Methoden zur Bestimmung des LAI

2.2.1 Feldspektroskopie

Feldspektrometer, ausgestattet mit Hyperspektralsensoren, bieten die Möglichkeit, kontinuierliche Signalkurven über zusammenhängende Bereiche des Spektrums zu generieren. Für konkrete Anwendungen ist es sinnvoll, den Messbereich auf dafür relevante Wellenbereiche abzustimmen. So ist es beispielsweise bei einer Beobachtung der Vegetation erforderlich, sowohl den Bereich des sichtbaren Lichtes (400 bis 700 nm) als auch die Signatur im nahen (700 bis 1400 nm) und fallweise im kurzwelligen (1400 bis 3000 nm) Infrarotbereich (NIR bzw. IR-A und SWIR bzw. IR-B) zu untersuchen.

Die Pflanzen eines Grünlandmischbestandes zeigen im Laufe der Aufwuchsperiode entwicklungsdynamische Veränderungen, die mit Hilfe der Spektroskopie erfasst und in weiterer Folge entsprechend interpretiert werden können. Abb. 2 zeigt am Beispiel ausgewählter Zeitpunkte innerhalb des zweiten Aufwuchses einer Dreischnittfläche die Verläufe von Spektralsignaturen in Abhängigkeit des Entwicklungszustandes. Je stärker sich die Pflanzen zu einem geschlossenen Bestand hin entwickeln, desto mehr nimmt der Kurvenverlauf eine für "vitale" Vegetation typische Form an.



Abb. 2: Unterschiedliche Spektralsignaturen (Feldspektrometer) eines Dauergrünlandbestandes in Abhängigkeit der Entwicklungsphasen während des zweiten Aufwuchses

In der vorliegenden Untersuchung wurde das Feldspektrometer HandySpec Field VIS/NIR 1.7 (tec5) eingesetzt. Der erste der beiden verbauten Sensoren (360 bis 1100 nm) weist eine spektrale Auflösung von 10 nm und der zweite Sensor (1000 bis 1690 nm) 5 nm auf. Mit höheren Auflösungen können mehr Informationen aus den Spektren ausgelesen werden, indem einzelne Elemente mit individuellen Reflexions- und Absorptionseigenschaften besser differenzierbar sind (GOVENDER et al. 2007). Die Ausgabe der Spektralsignatur erfolgt nach einem Resampling einheitlich in Ein-Nanometer-Schritten, d. h. geringere Auflösungen werden interpoliert und ergeben damit eine mehr oder weniger starke Glättung der Signatur. Beobachtungen unter suboptimalen atmosphärischen Bedingungen können dank der integrierten Referenzmessung einigermaßen effizient erledigt werden.

Eine Kalibrierung mittels Graureferenzadapter wurde in der Regel nur zu Beginn einer Messkampagne vorgenommen. Das HandySpec erfasst zunächst die Reflektanz des Pflanzenbestandes und unmittelbar darauf mit dem nach oben gerichteten Cosinus-Korrektor die Umgebungsstrahlung. Solange die Strahlungsbedingungen während des kurzen Zeitabstandes zwischen den beiden zusammengehörenden Messungen stabil sind, können korrekte Messungen erwartet werden, auch wenn sich die Atmosphäre während der Messkampagne insgesamt verändert. Die besten Ergebnisse sind zu erwarten, wenn konstant hohe Strahlungsintensitäten, also wolkenlose Bedingungen über den gesamten Messzeitraum bestehen. Aufgrund der Tatsache, dass die direkte Strahlung in Abhängigkeit des Sonnenstandes einen erheblichen Einfluss auf die Schattenbildung innerhalb des Bestandes und damit die Reflektanz hat, wurde bei den Messungen darauf geachtet, dass die Azimuthwinkel von einfallender und reflektierter Strahlung möglichst gering sind (MILTON, 1987). Messungen am Feld wurden deshalb vorzugsweise um die Tagesmitte vorgenommen.

In Abstimmung mit dem 10 Meter Tiling Grid von Sentinel-2 wurden auf der viermähdigen Dauergrünlandfläche in wöchentlichem Abstand in den in Abb. 1 dargestellten drei Bereichen Spektren an jeweils vier Positionen in je vierfacher Wiederholung (48 Einzelmessungen) erhoben und zu drei Mittelwertspektren zusammengefasst.

2.2.2 AccuPAR-Ceptometer

Mit dem AccuPAR PAR/LAI Ceptometer Model LP-80 wird die fotosynthetisch aktive Strahlung, engl. Photosynthetically Active Radiation (PAR), im Wellenlängenbereich von 400-700 nm gemessen. Die gemessene PAR stellt eine Kombination aus transmittierter und durch Blätter gestreute Strahlung dar. Durch Inversion eines PAR-Modells für Vegetation wird in weiterer Folge der LAI berechnet. Die genauen Gleichungen können im Operators Manual (DECAGON DEVICES INC., 2014) eingesehen werden.

Die Messungen erfolgten in denselben Bereichen wie die feldspektroskopischen Erhebungen (Abb. 1). Um den Einfluss der Messungen auf den Bestand möglichst gering zu halten wurden in nur einem der vier Punkte in jedem der drei Bereiche mit wöchentlicher Rotation Erhebungen durchgeführt. Ein LAI-Wert pro Messpunkt setzt sich aus 24 Einzelmessungen zusammen, die in einem Radius von ca. 2 m um den Referenzsensor, jeweils um 15° versetzt erhoben wurden.

2.2.3 Sentinel-2

Die Sentinel-2-Satelliten zählen zur neuesten Generation von Erdbeobachtungssatelliten der Europäischen Raumfahrtbehörde (ESA). Durch die hohe Wiederholrate von fünf Tagen am Äquator und die hohe räumliche Auflösung, eröffnen sich vor allem für den landwirtschaftlichen Sektor vielfältige Anwendungsmöglichkeiten (DRUSCH et al. 2012). Mithilfe der Sentinel-2-Satelliten können im Bereich der Erdbeobachtung Landnutzungskarten, Landnutzungsänderungen, geo- und biophysikalische Variablen und viele weitere Parameter erstellt werden (ESA, 2018b). Um die Sentinel-2-Daten mit den Bodenmessungen vergleichen zu können, wurden die Werte aus dem Zentrum des nächstliegenden Pixels der Satellitenaufnahme herangezogen und den Bodenmessungen gegenübergestellt.

Die Aufbereitung der Sentinel-2-Daten erfolgt durch das Institut für Vermessung, Fernerkundung und Landinformation (IVFL) der Universität für Bodenkultur, Wien. Sämtliche verwendete

Sentinel-2-Daten wurden über das Datenportal Sentinel-2 Value Adder bezogen. Eine genaue Beschreibung der Datenaufbereitung sowie der erhältlichen Produkte findet sich in VUOLO et al. (2016).

2.3 Transformationsschritte und Berechnung des LAI

Die im Feld erhobenen Hyperspektralsignaturen wurden mit Hilfe von Spectral Response Functions (ESA, 2018a) in multispektrale Bänder, entsprechend den Sentinel-2-Definitionen, umgeformt und anschließend mit zeitnahen Sentinel-2-Reflexionsdaten (\pm max. 4 Tage) verglichen (Abb. 3). Das fehlende Band 12, welches außerhalb des spektralen Aufnahmebereichs des Feldspektrometers liegt, wurde anhand der empirischen Gleichung1 berechnet.

Gleichung 1: Berechnung des Sentinel-2 Bandes Nummer 12

B12= -1.774155 + B2 * (-0.367618) + B4 * 0.706318 + B6 * (-0.036889) + B11 * 0.557938

Aufgrund der leichten Überschätzung des Feldspektrometers im hohen Reflexionsbereich (Abb. 3), wurden sämtliche Reflexionswerte des Feldspektrometers mit Gleichung 2 an die Sentinel-2-Reflexionen angepasst.



Abb. 3: Vergleich der Sentinel-2 Reflexionswerte mit den umgeformten Feldspektrometer-Reflexionswerten vor der Anpassung mit Gleichung 2

Gleichung 2:

RFspec_{adj} = 1,06301245+ RFspec * 0,95792555

RFspec_{adj} entspricht den angepassten und RFspec den ursprünglichen spektralen Reflexionswerten des Feldspektrometers.

Mithilfe eines neuronalen Netzes nach BARET et al. (2010) wurde aus den umgeformten Multispektraldaten des Feldspektrometers und aus den Sentinel-2-Daten der Blattflächenindex identisch berechnet. Das Neuronale Netz wurde speziell an die Daten von Sentinel-2 angepasst und mit Werten aus den physikalischen Strahlungstransfermodellen PROSPECT und SAIL (JACQUEMOUD et al., 2009, VERHOEF, 1984) trainiert.

Die LAI-Werte der drei Sensoren wurden in R 3.5.1. (R CORE TEAM, 2018; ZEILEIS & GROTHENDIECK, 2005) auf Tagesbasis interpoliert und wöchentlich am gleichen Tag miteinander verglichen. Um den linearen Zusammenhang der Sensoren bewerten zu können, wurde der Pearson Korrelationskoeffizient berechnet.

3 Ergebnisse und Diskussion

Für die LAI-Werte, berechnet aus Feldspektrometer und Sentinel-2, konnte für die Dauer der gesamten Vegetationsperiode ein Korrelationskoeffizient (r) von 0,97 berechnet werden. Zwischen Sentinel-2 und AccuPAR (r = 0,88) sowie zwischen Feldspektrometer und AccuPAR (r = 0,84) bestehen ebenfalls starke lineare Zusammenhänge (Abb. 4).



Abb. 4: Streudiagramm zwischen Bodenmessungen (links: AccuPAR, rechts: Feldspektrometer) und den satellitenbasierten LAI-Messungen der Versuchsfläche für das Jahr 2018

Zu Beginn des ersten Aufwuchses wird der LAI durch das AccuPAR deutlich unterschätzt (Abb.5). Eine derartige Unterschätzung des AccuPAR im Vergleich zu einer destruktiven Erhebungsmethode konnte auch von HE et al. (2007) festgestellt werden. Die unterschiedlich schnelle morphologische und phänologische Entwicklung der im Grünlandbestand vorkommenden Arten sowie die generelle Artenzusammensetzung beeinflussten die LAI-Erfassung erheblich. Im vorliegenden Versuch wurde beim zweiten Aufwuchs ein um mehr als 50 Gewichtsprozent höherer Anteil an Blättern am Gesamtpflanzengewicht im Vergleich zum ersten Aufwuchs gemessen. Daraus kann ein Teil der Abweichungen der Messmethoden zwischen den einzelnen Aufwüchsen erklärt werden. In den bereits weiter entwickelten Pflanzenbeständen des ersten, zweiten und vierten Aufwuchses kam es vor der Ernte zu einem sprunghaften Anstieg des AccuPAR-LAI. Aufgrund der fortgeschrittenen Vegetationsentwicklung trat zu diesen Zeitpunkten eine Lagerung der Grünlandbestände auf. Diese hat sehr geringe PAR-Werte unterhalb des Bestandes zur Folge, woraus sich der sehr hohe LAI erklären lässt. Aufgrund dieses Phänomens kommt es schließlich zu einer Sättigung der LAI Werte, da der PAR-Wert unterhalb des Bestandes nicht mehr weiter absinken kann.



Abb. 5: LAI-Zeitreihen der unterschiedlichen Sensoren für die einzelnen Aufwüchse des Jahres 2018

Vor allem zu Beginn der Folgeaufwüchse wurde eine gute Übereinstimmung der AccuPAR-Methode mit den optischen Sensoren festgestellt. Es liegt die Vermutung nahe, dass sich die homogenere Struktur des Grünlandes in den Folgeaufwüchsen einheitlicher erfassen lässt, wohingegen die deutlich heterogene Struktur des primären Grünlandaufwuchses zu erheblichen Differenzen zwischen den Erhebungsmethoden führt.

4 Fazit & Ausblick

In der vorliegenden Arbeit wurden LAI-Zeitreihen einer Wirtschaftsgrünlandfläche aus AccuPAR-Messungen, Feldspektrometer- und Sentinel-2-Reflexionsdaten berechnet, miteinander verglichen und deren Zusammenhänge untersucht. Zwischen allen drei Sensoren konnte ein starker linearer Zusammenhang beobachtet werden. Die gute Übereinstimmung der drei Sensoren, vor allem zwischen Feldspektrometer und Sentinel-2, bietet die Möglichkeit, über eine virtuelle Konstellation LAI-Zeitreihen mit mehreren Sensoren zu generieren.

In Gebieten mit einem hohen Bewölkungsgrad kann die Kombination von Sensoren zu einer verbesserten Beschreibung der Entwicklungsdynamik von Grünlandbeständen führen. In strukturell komplex aufgebauten Grünlandbeständen kommt es aufgrund von sich ändernder Blattarchitektur, unterschiedlichem Blattstängel-Verhältnis, ungleicher phänologischer Entwicklung der bestandesbildenden Arten und weiterer Parameter zu Differenzen zwischen den unterschiedlichen Messmethoden. Um valide Messergebnisse biophysikalischer Variablen von komplexen und artenreichen Grünlandbeständen zu erhalten, müssen deren vegetationsspezifische Eigenheiten beachtet werden. Die Ergebnisse dieser Forschungsarbeit sind von wachsender Bedeutung für das Monitoring von Grünlandflächen. Regionale und auch überregionale

Anwendungsmöglichkeiten, wie etwa der Einsatz in Grünlandwachstumsmodellen stellen eine umfangreiche und zusätzliche Informationsquelle für die Landwirtschaft dar.

5 Danksagung

Die Arbeit wurde durch das FFG-Projekt Farm/IT: "Innovative Technologien für eine smarte Landwirtschaft" und über die HBLFA Raumberg-Gumpenstein durch das DaFNE-Projekt SatGrass: "Nutzung von Fernerkundungs- und Klimadaten zur Beschreibung von Ertrags- und Qualitätsdynamiken im Grünland" gefördert.

6 Literaturverzeichnis

- BARET, F., WEISS, M., BICHERON, P. & BERTHELOT, B., 2010: Sentinel-2 MSI Products WP1152 Algorithm Theoretical Basis Document for Product Group B. INRA-EMMAH, Avignon, France.
- CHEN, J.M. & BLACK, T.A., 1992: Defining leaf area index for non-flat leaves. Plant, Cell & Environment, 15 (4), 421-429.
- CLEVERS, J., KOOISTRA, L. & VAN DEN BRANDE, M., 2017: Using Sentinel-2 Data for Retrieving LAI and Leaf and Canopy Chlorophyll Content of a Potato Crop. Remote Sensing 9 (5), 405.
- DECAGON DEVICES INC. 2014: Operator's Manual. AccuPAR model LP-80 PAR/LAI Ceptometer Pullman, WA, 78.
- 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., MARCHESE, F. & BARGELLINI, P., 2012: Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. Remote Sensing of Environment, 120, 25-36.
- ESA, 2018a: Sentinel-2 Spectral Response Functions (S2-SRF), 5. https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/document-library/ /asset publisher/Wk0TKajiISaR/content/sentinel-2a-spectral-responses, (10.11.2018).
- ESA, 2018b: Sentinel Online. SENTINEL-2, https://sentinel.esa.int/web/sentinel/missions/sentinel-2, (11.12.2018).
- GOVENDER, M., CHETTY, K. & BULCOCK, H., 2007: A review of hyperspectral remote sensing and its application in vegetation and water resource studies. Water SA, **33** (2), 145-152.
- HE, Y., GUO, X. & WILMSHURST, J.F., 2007: Comparison of different methods for measuring leaf area index in a mixed grassland. Canadian Journal of Plant Science, **87** (4), 803-813.
- JACQUEMOUD, S., VERHOEF, W., BARET, F., BACOUR, C., ZARCO-TEJADA, P.J., ASNER, G.P., FRANÇOIS, C. & USTIN, S.L., 2009: PROSPECT + SAIL models: A review of use for vegetation characterization. Remote Sensing of Environment 113, Supplement 1 (0), 56-66.
- MILTON, E.J., 1987: Principles of field spectroscopy. International Journal of Remote Sensing, **8**(12), 1807-1827.
- R CORE TEAM, 2018: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/.
- VERHOEF, W., 1984: Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. Remote Sensing of Environment, **16**(2), 125-141.

- VUOLO, F., ŻÓŁTAK, M., PIPITONE, C., ZAPPA, L., WENNG, H., IMMITZER, M., WEISS, M., BARET, F. & ATZBERGER, C., 2016: Data Service Platform for Sentinel-2 Surface Reflectance and Value-Added Products: System Use and Examples. Remote Sensing, 8 (11), 938.
- ZEILEIS, A. & GROTHENDIECK, G., 2005: zoo: S3 Infrastructure for Regular and Irregular Time Series. Journal of Statistical Software, 14 (6), 1-27.

Wörthersee 3D – Empirischer Vergleich von Flachwasserbathymetriedaten: UAS Photogrammetrie versus Single Beam Echolot

GERNOT PAULUS¹, KARL-HEINRICH ANDERS¹, MELANIE ERLACHER¹, PETER MAYR² & RUDI SCHNEEBERGER³

Zusammenfassung: 2018 fand in Kärnten mit der Befliegung des gesamten Wörtherseeufers durch ein professionelles unbemanntes Luftfahrzeug ("UAS- Unmanned Aerial System", "Drohne") eines der flächenmäßig größten wissenschaftlichen drohnenbasierten Geodatenerfassungsprojekte in Österreich statt. Im Rahmen einer interdisziplinären Forschungskooperation wurde das gesamte Wörtherseeufer mit einer Länge von fast 50 km mit einem zertifizierten Multirotorsystem ausgerüstet mit einem hochauflösendem Kamerasystem vollständig beflogen und photogrammetrisch ausgewertet. Im Rahmen dieses Konferenzbeitrags fokussieren wir auf den empirischen Vergleich von Bathymetriedaten der Flachwasserbereiche des Wörthersees. Single Beam Echolot Profile, die im Rahmen der Echolotvermessung des Wörthersees (Herbst 2017) erfasst wurden, werden mit UAS-basierten photogrammetrisch abgeleiteten Oberflächenmodellen (Befliegung Frühjahr 2018) von ausgewählten Flachwasserbereichen quantitativ verglichen und die ersten Ergebnisse präsentiert.

1 Einleitung

Der Wörthersee steht wie viele Seen in Europa im Zentrum des gesellschaftlichen und wirtschaftlichen Interesses. Die anthropogene Nutzung der Seen und ihrer Ufer steht oft im Gegensatz zum Interesse, den guten ökologischen Zustand von Gewässern nachhaltig für kommende Generationen zu bewahren. Besonders die Uferbereiche, die als Puffer zwischen den terrestrischen und aquatischen Zonen fungieren, stellen aufgrund ihrer Diversität ökologisch hochwertige Lebensräume dar. Neben den chemisch-physikalischen bzw. biologischen Parameter muss bei der Umsetzung der EU Wasserrahmenrichtlinie eine Bewertung der strukturellen Gegebenheiten der Seeufer durchgeführt werden.

Im Rahmen einer interdisziplinären Forschungskooperation zwischen der Abteilung 8 Umwelt, Wasser und Naturschutz des Amtes der Kärntner Landesregierung, den Industriepartnern flussbau iC GesmbH und ViewCopter e.U. sowie dem Studiengang Geoinformation und Umwelttechnologien an der FH Kärnten als wissenschaftlichen Partner wurde das gesamte Wörtherseeufer mit einer Länge von fast 50 km mit einem zertifizierten Multirotorsystem ausgerüstet mit einem hochauflösendem Kamerasystem vollständig beflogen und photogrammetrisch ausgewertet.

Ziel dieses Projektes ist die Erstellung eines sehr genauen digitalen dreidimensionalen Geländeund Oberflächenmodells des Uferstreifens des Wörthersees unter besonderer Berücksichtigung des Wasser- Landübergangsbereiches. Die erwarteten Ergebnisse sind der "missing digital link" –

¹ FH Kärnten, Geoinformation und Umwelttechnologien, Europastrasse 4, A-9524 Villach, E-Mail: [g.paulus, k.anders, m.erlacher]@fh-kaernten.at

² flussbau iC GesmbH, 10. Oktoberstrasse 23, A-9500 Villach, E-Mail: p.mayr@ic-group.at

³ ViewCopter e.U., Lilienweg 2, A-9560 Feldkirchen, E-Mail: rudi@vcopter.net

der fehlende digitale Baustein zu einem nahtlosen, durchgängigen und vollständigen 3D Modell zwischen den bereits vorliegenden Laserscan-Geländedaten der Seeumgebung und der im Winter 2017 durchgeführten Echolotvermessung des Wörthersees.

2 Problemstellung

Herausforderungen in diesem angewandten Forschungsprojekt sind: (1) die Flugplanung zur Durchführung der Drohnenbefliegung gegliedert in einzelne Teilmissionen von ca. 1 km Länge, um den gesetzlichen Sichtflugvorgaben zu entsprechen sowie die enge Kommunikation mit der Austro Control am Tower des Flughafens Klagenfurt; (2) die Sicherstellung der Lage- und Positionsgenauigkeit der erfassten Daten durch Referenzmessungen im gesamten Projektgebiet; (3) das Management und die photogrammetrische Prozessierung der riesigen anfallenden digitalen Datenmenge und, last, but not least, (4) die Integration der verschiedenen 3D Gelände- und Oberflächendaten der Seeumgebung, des Übergangsbereiches und des Seeuntergrundes zu einem ganzheitlichen, lückenlosen dreidimensionalen digitalen 3D Modell des Wörthersees.

Im Rahmen dieses Konferenzbeitrags fokussieren wir auf den empirischen Vergleich von Bathymetriedaten der Flachwasserbereiche des Wörthersees. Single Beam Echolot Profile, die im Rahmen der Echolotvermessung des Wörthersees (Herbst 2017) erfasst wurden, werden mit UASbasierten photogrammetrisch abgeleiteten Oberflächenmodellen (Befliegung Frühjahr 2018) von ausgewählten Flachwasserbereichen quantitativ verglichen und die ersten Ergebnisse präsentiert.

3 Methodik

Im Zuge einer primären Datenerfassungskampagne wurde im Zeitraum 27. März bis 18. April 2018 innerhalb von 11 Flugmissionstagen der gesamte Wörtherseeuferstreifen mit einem professionellen, Klasse C zertifizierten UAS-Multirotorsystem beflogen und die dabei erfassten Einzelbilder photogrammetrisch ausgewertet. Als Referenzdaten dienten einerseits für die photogrammetrische Prozessierung mittels DGPS erfasste Passpunkte und andererseits für den empirischen Vergleich der Flachwasserbereiche Single Beam Echolot Punktmessungen, die im Rahmen der Echolotvermessungskampagne des Wörthersees im Herbst und Winter 2017 erfasst wurden.

3.1 UAS Datenerfassung und photogrammetrische Prozessierung

Als Grundlage für die UAS Flugmissionsplanung wurden vom der Abteilung 8 Umwelt des Amtes der Kärntner Landesregierung ausgehend von der Uferlinie (45.935 km) des Wörthersees ein seeseitiger und landseitiger ca. 80m breiter Puffer mit einer Gesamtflächenausdehnung von ca. 5.5km² vorgegeben.

Die Befliegung wurde von einem professionellen Dienstleistungsunternehmen (ViewCopter e.U.) mit einem für die Befliegung von besiedelten Gebieten zugelassenen UAS Multirotorsystem (Viewcopter V6H - Adaptierter und erweiterter Hexacopter auf Basis DJI Matrice 600 Pro) mit einer Flugzeit von ca. 25 Minuten durchgeführt. Als Kamerasensor wurde eine SONY ILCE-A7RM2 Vollformatkamera mit 44 MP und einem 28mm Objektiv verwendet. Abbildung 1 zeigt das UAS System im Einsatz.



Abb. 1: UAS Multirotorsystem ViewCopter V6 H mit SONY A7R im Einsatz.

Auf Grund der aktuellen gesetzlichen Rahmenbedingungen sind in Österreich nur UAS Befliegungen mit direkter Sichtverbindung des Piloten zum UAS und einer maximalen Flughöhe von 150m AGL möglich (AUSTRO CONTROL 2018). Auf Grund dieser gesetzlichen Rahmenbedingungen wurde das gesamte Projektgebiet in 54 Flugmissionssegmente unterteilt. Jedes Segment wurde mit 4-5 Flugstreifen mit 70% Überlappung in Flugrichtung und 70% Überlappung zwischen den einzelnen Flugstreifen geplant. Besonderes Augenmerk bei der Flugplanung wurde dabei sowohl auf eine Überlappung von mindestens 100m zwischen angrenzenden Segmenten als auch auf eine ausreichende Erfassung über die vorgegebenen Projektgebietsgrenzen hinaus, um eine stabile photogrammetrische Prozessierung der Randbereiche zu gewährleisten. Ausgehend von der vorgegebenen 5.5.km² großen Projektgebietsgröße beträgt die Gesamtabdeckung der UAS Befliegung bei einer Flughöhe von 150m und einer Bodenauflösung von ca. 2cm 14.5 km² (Tab. 1).

Die photogrammetrische Prozessierung wurde mit AGISOFT Photoscan Pro (Version 1.4.1.) durchgeführt. Dabei wurden die UAS Daten der 54 Flugmissionssegmente zu insgesamt 9 Prozessierungsregionen zusammengefasst. Als Serverinfrastruktur wurde ein XEON Server mit 24 Kernels, 4 GPU NVIDIA 20N Graphikkarten und 128 GB RAM Arbeitsspeicher eingesetzt. Die Gesamtmenge der zu prozessierenden Daten umfasst 17589 Einzelbilder mit einem Datenvolumen von 605.54 GB. Tabelle 1 gibt einen Überblick über die erfassten Bilddaten und die Ergebnisse der photogrammetrischen Auswertung.

UAS Survey Lake Woerthersee 2018 27.3.2018-18.4.2018 11 Flight Missions Photogrammetric Processing	Number Flight Segments	Number Images Total	Number Images Aligned	Flying Altitude m AGL	Ground Resolution cm/pix	Coverage Mission Area km²	Dense Point Cloud # points	# Control Points	RMSE X Control Points (cm)	RMSE Y Control Points (cm)	RMSE Z Control Points (cm)	RMSE Total Control Points (cm)
Mission Area 01	6	1899	1811	153	2.33	1.39	469,875,189	12	5.66	4.23	4.1	8.17
Mission Area 02	6	1893	1883	154	2.34	1.61	666,340,903	20	1.34	1.15	0.85	1.97
Mission Area 03	6	1952	1924	149	2.27	1.76	865,558,108	23	1.26	1.69	0.85	2.27
Mission Area 04	5	1667	1618	152	2.28	1.75	840,883,589	25	1.26	0.97	0.95	1.85
Mission Area 05	5	1686	1684	153	2.31	1.41	586,423,132	15	0.71	1.39	0.99	1.85
Mission Area 06	6	1907	1208	156	2.33	1.02	390,318,553	12	0.69	1.43	0.56	1.68
Mission Area 07	7	2592	2534	154	2.27	2.07	883,664,917	23	1.47	1.46	0.74	2.2
Mission Area 08	6	2105	2099	153	2.26	1.74	802,798,064	24	2.07	2.26	1.81	3.57
Mission Area 09	7	1888	1888	147	2.23	1.71	816,115,969	26	1.24	0.99	1.33	2.07
	54	17589	16649	152	2.29	14.46	6,321,978,424	180	1.74	1.73	1.35	2.85
	Total Sum	Total Sum	Total Sum	Mean	Mean	Total Sum	Total Sum	Total Sum	Mean	Mean	Mean	Mean

Tab. 1: Überblick über die erfassten Bilddaten und die Ergebnisse der photogrammetrischen Auswertung

3.2 Referenzmessungen

Als Referenzdaten wurden einerseits für die photogrammetrische Prozessierung mittels DGPS erfasste Passpunkte und andererseits für den empirischen Vergleich der Flachwasserbereiche Single Beam Echolot Punktmessungen, die im Rahmen der Echolotvermessungskampagne des Wörthersees im Herbst und Winter 2017 erfasst wurden, verwendet.

3.2.1 Terrestrische DGPS Passpunkte

Die Erfassung der terretsrischen Passpunkte wurde von der Firma flussbau iC GsmbH mittles TRIMBLE R8S DGPS (RTK-APOS) unter Verwednung einer lokalen Transformation erfasst. Eine besondere Herausforderung der Erfassung der Passpunkte bei diesem Projekt ist der sehr hohe Verbauungsgrad des Wörtherseeufers verbunden mit der sehr eingeschränkten freien Zugänglichkeit des Uferbereiches. Als Passpunkte wurden daher landseitig gut erkennbare markante Punkte (Kanaldeckel) entlang der Uferstrasse verwendet. Zusätzlich wurden mittels Boot markante Punkte an Stegen seeseitig eingemessen. Insgesamt wurden mit diesem Zugang 417 Passpunkte erhoben.

3.2.2 Echolotmessungen

• Positionierung und Höhenbestimmung

Die gesamte Positionierung und Höhenbestimmung der Echolotvermessung erfolgte mittels RTK-DGNSS, wobei die Korrekturdaten von APOS bezogen wurden. Die Kontrolle erfolgt über die amtlichen Festpunkte. Die Kontrolle des Wasserspiegels wurde zusätzlich am Pegel Pörtschach (HZB-Nummer: 212985) durchgeführt.

• Fächerecholotmessung

Das eingesetzte Fächerecholot arbeitet mit einer Frequenz von 400kHz. Der Öffnungswinkel variiert zwischen 150° und 165°. Die entsprechende Beamöffnung beträgt quer 0,5° und längs 1°. Die Tiefenauflösung kann mit 6mm angenommen werden. Die maximale Ping Rate beträgt 50Hz. Als Inertialsystem wird die Messeinheit Octans V mit einer Richtungsgenauigkeit von 0,1° und von 0,01° bei Roll, Pitch und Yaw sowie einer Auflösung von 0,001° verwendet. Zur integrativen
Stützung kam ein GNSS Heading System der AsteRx-U Marine von Septentrio zum Einsatz. Die Temperatur und das Wasserschallgeschwindigkeitsprofil wurde mit der VALEPORT Swift SVP Sonde gemessen.

• Sedimentecholotmessung "Sub Bottom Profiling"

Die Sedimentecholotung wurde mit dem SES-2000 compact durchgeführt. Das SES-2000 arbeitet mit einer hohen Frequenz (85 - 115kHz) und mit einer niedrigen Frequenz (2 - 22kHz). Die Schwingeröffnung ist 4° x 4°. Die maximale Ping Rate beträgt dabei 40 Hz.

• Ergänzungsmessung mit Singlebeam Echolot

Das eingesetzte Singlebeam Admodus Sonar arbeitet mit einer Frequenz von 210kHz und 33kHz oder 15kHz mit einem Öffnungswinkel von +/- 4.5° und +/-9° oder +/-13.5°. Die Auflösung am Boden beträgt 1cm. Die maximale Ping Rate beträgt bis 20 Hz. Insgesamt liegen für die Flachwasserbereiche des Wörthersees 8295 Single Beam Echoloteinzelpunktmessungen vor. Davon wurden in einer ersten empirischen Vergleichsuntersuchung 1504 Single Beam im Ostteil des Wörthersees im Bereich von Krumpendorf bis Klagenfurt (Uferlänge ca. 5.5 km) verwendet.

4 Erste Ergebnisse

Nachfolgend werden erste Ergebnisse der empirischen Vergleichmessungen zwischen photogrammetrisch berechneten Oberflächenmodellen digitalen von ausgewählten Flachwasserbereichen und Single Beam Echolot Einzelpunktmessungen aus dem Ostteil des Wörthersees zwischen Krumpendorf und Klagenfurt vorgestellt. Abbildung 2 zeigt einen Detailausschnitt eines Orthophotos eines ausgewählten Flachwasserbereiches in Krumpendorf. Die Orthophotoauflösung beträgt 2.37 cm pro Pixel. Abbildung 3 zeigt das photogrammetrisch berechnete digitale Oberflächemodell (4.55 cm/pix) denselben Ausschnitts. In Abbildung 4 sind für den Abschnitt Krumpendorf - Klagenfurt (Uferlänge von ca. 5.5 km; 1505 Single Beam Referenzpunkte) die Absoluthöhen aus Echolotmessung und photogrammetrisch berechnetem digitalen Oberflächenmodell vergleichend gegenübergestellt. Die Messpunkte sind nach aufsteigenden Absoluthöhen (von tiefen Wasserbereichen zu Flachwasserbereichen) angeordnet. Die Höhe des Wasserspiegels wurde visuell auf Grundlage des hochauflösenden Orthophotos an gut erkennbaren flachen Uferbereichen identifiziert und dann im digitalen Oberflächenmodell gemessen. Ein Vergleich mit dem Pegel in Pörtschach zum Befliegungszeitpunkt wurde noch nicht durchgeführt.



Abb. 2: Detailausschnitt Orthophoto (2.27 cm/Pixel) & Single Beam Referenzpunkte im Bereich Krumpendorf



Abb. 3: Detailausschnitt Digitales Oberflächenmodell (4.55 cm/Pixel) & Single Beam Referenzpunkte im Bereich Krumpendorf



Abb. 4: Vergleich der Absoluthöhen von Single Beam Echolotmessung ("*SingleBeamZ*") und photogrammetrisch berechnetem digitalen Oberflächenmodell ("*UAS_DSM_Z*") für einen ca. 5.5. km Uferabschnitt zwischen Krumpendorf und Klagenfurt. Die Messpunkte sind von links nach rechts nach zunehmender Absoluthöhe geordnet. Eine Refraktionskorrektur wurde nicht durchgeführt.

5 Fazit & Ausblick

Das hier vorgestellte Projekt stellt aktuell eines der flächenmäßig größten wissenschaftlichen drohnenbasierten Geodatenerfassungsprojekte in Österreich statt. Im Rahmen einer interdisziplinären Forschungskooperation zwischen der Abteilung 8 Umwelt, Wasser und Naturschutz des Amtes der Kärntner Landesregierung, den Industriepartnern flussbau iC und ViewCopter sowie dem Studiengang Geoinformation und Umwelttechnologien an der FH Kärnten als wissenschaftlichen Partner wurde das gesamte Wörtherseeufer mit einer Länge von fast 50 km mit einem zertifizierten Multirotorsystem ausgerüstet mit einem hochauflösendem Kamerasystem vollständig erfolgreich beflogen und photogrammetrisch durch die Verwendung von Passpunkten mit der notwendigen Genauigkeit erfolgreich berechnet. Um die Datenmenge zu reduzieren und für die Einbindung in das Geodatenportfolio des Kärnten Geographischen Informationssystems KAGIS zu optimieren, wurden sowohl die Orthophotos als auch die digitalen Oberflächenmodelle der 9 Teilprozessierungsgebiete zu einem Gesamtdatensatz fusioniert. Dabei wurden das finale Orthophoto auf 5cm Auflösung pro Pixel und das digitale Oberflächenmodell auf 10cm pro Pixel resampled und auf Basis der beim KAGIS verwendeten Luftbildblattschnitte organisiert. Aktuell arbeitet das KAGIS daran, die UAS Orthophotos in die bestehenden KAGIS Daten zu integrieren und via Kärnten Atlas (<u>http://gis.ktn.gv.at</u>) im 1. Quartal 2019 der Öffentlichkeit bereitzustellen.

Es wurden desweiteren erste Ergebnisse eines empirischen Vergleichs der hochauflösenden UAS digitalen Oberflächendaten mit Single Beam Echolot Referenzpunktmessungen im Flachwasser für einen ausgewählten Bereich im Ostteil des Wörthersees gezeigt. In weiterer Folge werden diese vergleichenden Untersuchungen für den gesamten Wörthersee durchgeführt.

6 Danksagung

Die photogrammetrische Auswertung dieses Projekts wurde im Rahmen einer Forschungskooperation zwischen der Abteilung 8 Umwelt und Naturschutz des Amtes der Kärntner Landesregierung und dem Studiengang für Geoinformation und Umwelttechnologien an der FH Kärnten durchgeführt. Die Single Beam Echolot Referenzdaten wurden vom Kärntner Geographischen Informationssystem KAGIS zur Verfügung gestellt.

7 Literaturverzeichnis

AUSTRO CONTROL, 2018: LBTH Nr. 67 - Lufttüchtigkeits- und Betriebstüchtigkeits-anforderungen für unbemannte Luftfahrzeuge der Klasse 1. https://www.austrocontrol.at/jart/prj3/ac/data/dokumente/LTH_LFA_ACE_067_2018-12-03_1212958.pdf, letzter Zugriff 27.12.2018.

Vergleich von Bathymetriedaten aus luftgestützter Laserscanner- und Kameraerfassung

CHRISTIAN MULSOW¹, GOTTFRIED MANDLBURGER^{2,3}, CAMILLO RESSL³ & HANS-GERD MAAS¹

Zusammenfassung: In diesem Beitrag werden die Verfahren Laserbathymetrie und Photogrammetrie für die Erfassung von Gewässerböden verglichen. Anhand eines kombinierten Laserscanner- und Bilddatensatzes des Autobahnsees in Bayern werden beide Verfahren ergebnismäßig einander gegenübergestellt sowie Vor- und Nachteile angeführt. Ein besonderer Fokus liegt dabei auf der Güte der Geländemodelle sowie der Vollständigkeit und räumlichen Auflösung. Praktische Aspekte bei der Auswertung sowie die Komplexität der jeweiligen Verfahren werden vergleichend betrachtet.

1 Motivation

Die Aufnahmesysteme Laserscanner und Kamera konkurrieren im Aufgabengebiet der luftgestützten Topographieerfassung. Bis zum Aufkommen der Laserscannertechnologie war hier die klassische Photogrammetrie das exklusive Verfahren der Wahl. Über die letzten Jahrzehnte entwickelte sich Laserscanning hin zum effektiven Werkzeug zur Geländeerfassung, insbesondere durch die First-Pulse/Last-Pulse Technologie bin hin zur Full-Waveform Auswertung (PFEIFER et al. 2015). Mit dem Aufkommen von effektiven automatisierten Verfahren der Bildmessung- und Zuordnung (z.B. Structure-from-Motion) konnte die Photogrammetrie wieder einigen Boden gut machen. Beide Verfahren können einander ergänzen, wie diverse Anwendungen zeigen (MANDLBURGER et al. 2017).

Die luftgestützte Laserbathymetrie hat sich in den vergangenen Jahren als leistungsfähiges Verfahren zur Vermessung der Topografie von Gewässerböden im Flachwasserbereich durchgesetzt. Ebenfalls lassen sich für diesen Zweck klassisch erhobene Bilddaten verwenden. Bei beiden Verfahren müssen Refraktionseffekte beim Medienübergang von Luft in Wasser berücksichtigt werden. Eine Voraussetzung dafür ist die Kenntnis über die Wasserspiegellage (Trennfläche) und den Brechungsindex des Gewässers. Letzter kann aus empirischen Werten in Abhängigkeit von der Gewässerart (Salz vs. Süßwasser) und der Temperatur geschätzt werden (QUAN & FRY 1995). Schwieriger gestaltet sich die Bestimmung der Wasseroberfläche, insbesondere wenn diese durch Wellenbewegungen beeinflusst wird. Zudem müssen bei Vorliegen eines Fließgewässers lokale Stauungen und Senkungen sowie die allgemeine Neigung der Wasseroberfläche mitberücksichtigt werden. Bei der Erfassung von stehenden Gewässern ohne signifikante Wellenbewegung kann die

¹ TU Dresden, Institut für Photogrammetrie und Fernerkundung, Helmholtzstraße 10, D-1062 Dresden, Deutschland, E-Mail: [christian.mulsow, hans-gerd.maas]@ @tu-dresden.de

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

³ TU Wien, Department für Geodäsie und Geophysik, Gußhausstraße 25, A-1040 Wien, E-Mail: camillo.ressl@geo.tuwien.ac.at

Problematik entsprechend vereinfacht werden. Die Wasseroberfläche lässt sich dann als waagerechte Ebene annehmen. Entsprechend wird nur noch die Wasserspiegelhöhe benötigt. Diese kann zum einen durch örtliche Pegelmessungen erfasst oder simultan innerhalb der Auswertung bestimmt werden.

Anhand eines Testdatensatzes sollen beide Technologien hinsichtlich Anwendbarkeit, Komplexität und Datenqualität einander gegenübergestellt werden.

2 Testdatensatz

Als Testdatensatz standen die Bild- und Laserscannerdaten einer Befliegung des Autobahnsees in Bayern nahe Augsburg zur Verfügung (Abb. 1). Es handelt sich dabei um einen künstlich angelegten, Grundwasser-gespeisten Baggersee in der Überflutungsebene des Lech Flusses. Das vorherrschende Bodensubstrat ist daher grobkörniger Kies. Die zeitgleiche Datenerfassung von Laserund Bilddaten erfolgte am 9. April 2018 im Rahmen eines Experiments, bei dem neben einem topo-bathymetrischen Laserscanner (Riegl VQ-880-G) auch zwei 100 Mpix Kameras zum Einsatz kamen. Neben der im vorliegenden Beitrag verwendeten RGB Kamera (IGI DigiCAM 100), kam auch eine pan-chromatische Kamera desselben Typs zum Einsatz, deren Empfindlichkeitsbereich mittels vorgeschaltetem Filter auf das mono-chromatische Coastal Blue (λ =400-460 nm) eingeschränkt wurde (MANDLBURGER et al. 2018). Insgesamt wurden vier Flugstreifen mit einer Fluggeschwindigkeit von 100 Knoten (50 m/s) erfasst.



Abb. 1: Aufnahmeobjekt Autobahnsee

2.1 Bilddaten

Der Bildblock besteht aus insgesamt 125 Bildern, welche aus zwei unterschiedlichen Höhen (64 Bilder aus 610m [Ground Sampling Distance (GSD): 5.6cm] und 61 Bilder aus 450m [GSD: 4.2cm]) aufgenommen wurde. Pro Höhenstufe erfolgte die Erfassung mit jeweils zwei Flugstreifgen mit rund 90% Längsüberlappung und 60% Querüberlappung (Ab. 2). Daraus folgt eine überaus hohe Redundanz der Bilddaten. Die Blockmitte ist in rund 40 Aufnahmen enthalten. Wie bereits erwähnt, kam als bildgebender Sensor eine IGI DigiCAM 100 Mediumformat-Kamera mit einer Auflösung von 10608 × 8708 Pixel (100 Mpx) und einem 50mm Objektiv zum Einsatz. Die Kamera basiert dabei auf einer PhaseOne iXU-RS 1000 (MANDLBURGER et al. 2018). Es wurden für die Untersuchungen ausschließlich die RGB-Bilddaten verwendet, da diese eine gegenüber den Coastal-Blue-Daten eine bessere Auswertbarkeit versprachen (MANDLBURGER et al. 2018). Für die Georeferenzierung standen die während des Bildfluges erfassten GNSSS/INS Daten sowie 10 am Boden signalisierte und mittels RTK-GNSS eingemessene Passpunkte zur Verfügung (siehe Abb. 2). Zudem wurden vier Targets für die radiometrische Kalibrierung des Laserscaners und des Bildmaterials eingemessen.



Abb. 2: Blockstruktur mit Passpunkten in Gelb über der in Pix4D generierten Punktwolke. (Screenshot Pix4DMapper)

2.1.1 Orientierung Bildblock

Zunächst wurde der gesamte Block in Pix4D Mapper⁴ prozessiert. Die Absolutorientierung erfolgte auf Basis der Passpunkte. In der gleichen Software wurde eine georeferenzierte Punktwolke der Oberfläche (siehe Abb. 2) sowie ein Orthophotomosaik generiert. Da innerhalb der verwendeten Software keine Modellierung von Refraktionseffekten implementiert ist, ist das extrahierte Höhenmodell (DHM) nur für die Landbereiche gültig. Bei Nichtberücksichtigung des Refraktionseinflusses weisen Unterwasserbereiche im digitalen DHM tendenziell zu geringe Tiefen auf (MAAS 2015). Für eine möglichst korrekte Bestimmung der Unterwasserbereiche ist eine Kompensation der Refraktionseffekte unabdingbar. Kommerzielle Softwareprodukte beinhalten die entsprechenden Funktionalitäten nicht. Daher wurde für die weitere Prozessierung eine spezielle Software, welche am Institut für Photogrammetrie und Fernerkundung an der TU-Dresden für Mehrmedien-Szenarien entwickelt wurde (MULSOW 2010), eingesetzt.

Für die Vorverarbeitung wurde der Block in MatchAT (Trimble/Inpho) prozessiert. Im Ergebnis standen automatisch generierte Verknüpfungspunkt-Messungen, manuell gemessene Passpunkte sowie Näherungswerte für die Orientierungen zur Verfügung. Eine Trennung der Verknüpfungspunkte nach Land- und Unterwasserpunkten erfolgte anhand der berechneten Punkthöhen. Aus dem orientierten Bildblock wurde anhand von manuellen Ufermessungen eine Wasserspiegelhöhe im See von 509,05 m ermittelt. Alle Punkte unterhalb dieses Schwellwertes konnten als Unterwasserpunkte angesehen werden. Circa 17000 Bildmessungen (~1000 Verknüpfungspunkte) wurden der weiteren Verarbeitung im Mehrmedien-Bündel übergeben, wobei ~1900 Bildmessungen (~120 Verknüpfungspunkte) auf den Unterwasserbereich entfielen.

Da Erfahrungen bei ähnlichen Projekten gezeigt hatten, dass eine integrierte Bestimmung der Wasseroberfläche auf Basis von Senkrechtaufnahmen in einer Bündelblockausgleichung nicht zuverlässig möglich ist (MULSOW 2018), wurde die Wasseroberfläche als Ebene (Normale in Lotrichtung) und der oben angeführten Wasserspiegelhöhe festgelegt. Eine gute Bestimmbarkeit der Trennflächenparameter bedingt große Einfallswinkel der Bildstrahlen relativ zur Oberflächennormale, da sich hier die Refraktionseffekte stärker auf die Bildkoordinaten auswirken. Projekte im Nahbereich haben gezeigt, dass selbst komplexe Trennflächen aus konvergent erstellten Bildblöcken mit hoher Genauigkeit bestimmt werden können (MULSOW 2016).

Die Bündeltriangulation erfolgte in zwei unterschiedlichen Parameterkonfigurationen, welche sich hinsichtlich der Behandlung der Unterwasserpunkte unterschieden:

- I. Simultane Bestimmung aller Unbekannten (sämtliche Orientierungen und Kameraparameter), Unterwasserpunkte als Ein-Medium-Punkte behandelt (ohne Refraktionskorrektur)
- II. Simultane Bestimmung aller Unbekannten (sämtliche Orientierungen und Kameraparameter), Unterwasserpunkte als Mehrmedien-Punkte behandelt

⁴ www.pix4d.com

Genauigkeitsparameter	I - Ohne Refraktions- modellierung	II - Mit Refraktions- modellierung		
Rückprojektionsfehler s ₀ [px]	0.11	0.11		
RMS Bildpunkte Land x' y' [px]	0.09/0.10	0.10/0.10		
RMS Bildpunkte Wasser x' y' [px]	0.13/0.14	0.14/0.14		
Kamerakonstante c _k [mm]	51.532	51.529		
Hauptpunkt x _H [mm]	0.0042	0.0041		
Hauptpunkt y _H [mm]	0.0287	0.0286		
Radialsymetrische Verz. A1	-1.486e-05	-1.490e-05		
Radialsymetrische Verz A2	3.933e-09	3.953e-09		
Mittlere Genauigkeit (RMS) Verknüpfungspunkte Land X/Y/Z [cm]	0.1/0.2/2.6	0.1/0.2/2.6		
Mittlere Genauigkeit (RMS) Verknüpfungspunkte Wasser X/Y/Z [cm]	0.2/0.3/3.9	0.2/0.3/5.4		

Tab. 1: Ergebnisse Bündelblockausgleichung – mit und ohne Refraktionsmodellierung (Ein-Medium vs. Mehrmedien)



Abb. 3: Vergleich der Höhen (bezogen auf Wasserspiegel) der Verknüpfungspunkte aus Bündelblockausgleichung mit und ohne Refraktionsmodellierung

Aus Tabelle 1 ist ersichtlich, dass die Modellierung der Refraktion praktisch keinen Einfluss auf die Genauigkeitskennwerte der Bündelblockausgleichung hat (z.B. s0 und RMS der Bildpunktmessungen). Ein Vergleich der Verknüpfungspunktkoordinaten zeigt den hohen Grad der Übereinstimmung für die Lagekomponente (RMS 0.1cm in X und Y). Einzig die Höhen der Unterwasserpunkte differieren, wie erwartet, signifikant (siehe Abb. 3). Hohe Übereinstimmungen konnten auch für die Bildorientierungen sowie die Kameraparameter beobachtet werden (siehe Tabelle 1). Hier differieren die Werte im Bereich ihrer Standardabweichungen. Daraus lässt sich für diesen Bildblock schlussfolgern, dass eine konventionelle Orientierung des Blockes ohne Refraktionsmodellierung zulässig ist (wenn keine Höhenpasspunkte unter Wasser verwendet werden). Inwiefern sich diese Erkenntnis auf andere Bildblöcke übertragen lässt, ist Gegenstand zukünftiger Untersuchungen. Grundsätzlich kann dies jedoch für flache Gewässer (relativ zur Flughöhe) und regelmäßige Bildblöcke (Senkrechtaufnahmen, ähnliche Überlappungen) angenommen werden. Grund ist die daraus folgende Homogenität des Einflusses der Refraktion, welche primär Einfluss auf die Punkthöhen hat. Demzufolge können Standard-Programme für die Blockorientierung verwendet werden und bei der Geländemodellextraktion ist der Vorwärtsschnitt mit Mehrmedienmodellierung durchzuführen (MANDLBURGER 2018). Wie in MULSOW (2010) gezeigt, lässt sich die Modellierung der Refraktion für den Vorwärtsschnitt relativ einfach realisieren - im Gegensatz zur Bündelausgleichung oder dem Rückwärtsschnitt.

Die aus der Bündelblockausgleichung ermittelten RMS-Werte der Unterwasserpunkte sind nur um ca. Faktor 2 schlechter als die der Landpunkte. Diese Werte sind jedoch zu optimistisch, da aufgrund der festgehaltenen Wasseroberfläche und des Refraktionsindex diese Komponente des Bildstrahlweges im Modell als fehlerfrei angesehen werden. Dies ist in der Praxis jedoch nicht der Fall, womit mit einer weitaus schlechteren Höhegenauigkeit der Unterwasserpunkte gerechnet werden muss. Aufgrund des Fehlens von Kontrollmessungen im See kann die Güte der Verknüpfungspunktkoordinaten nicht durchgreifend geprüft werden. Aus vergleichbaren Projekten (MULSOW 2018) kann mit einer um Faktor 4 gegenüber von Landpunkten reduzierten Höhengenauigkeit gerechnet werden, also ~10cm statt ~5cm.

2.1.2 Punktwolkengenerierung aus Bilddaten

Die Punktwolkengenerierung besteht im Wesentlichen aus zwei Schritten –Bildmessung und Berechnung der Objektkoordinaten über (multiplen) Vorwärtsschnitt. Analog zu etablierten Verfahren zur Geländemodellextraktion wurden zunächst die Bilder paarweise in Stereonormalbilder umgewandelt. Die Festlegung der Paarungen erfolgte unter dem Gesichtspunkt möglichst großer Schnittwinkel. Daher wurden nicht direkt benachbarte Bilder als Paare definiert, sondern Paarungen mit rund 60% Überlappung. Die von Y-Parallaxen weitestgehend befreiten Bildpaare erleichtern die Suche von homologen Punkten in den Bildpaaren. Im Mehrmedienfall werden Kernlinien nicht zwangsläufig wie im Ein-Medienfall als Geraden abgebildet, sondern eher als Kurven. Aufgrund der nahezu senkrechten Aufnahmen und der relativ zum Aufnahmeabstand geringen Wassertiefe kann die Krümmung der Kernlinien als hinreichend klein angenommen und daher für die Punktzuordnung vernachlässigt werden.

Angelehnt an etablierte Matchingverfahren wurde der Suchprozess hierarchisch gestaltet (Bildpyramiden-Ansatz). Ausgangspunkt ist ein in der Auflösung um den Faktor 5 verringertes Bildpaar. Das Referenzbild wurde zunächst in ein 75×50 Zellenraster unterteilt. Für jede Zelle erfolgte dann die Bestimmung des besten Merkmalpunktes (Harris-Operator). Die extrahierten Harris-Punkte wurden dann über Least Squares Matching (LSM) im Stereopartner identifiziert und subpixelgenau vermessen (Patch-Größe 21×21 Pixel, Reduzierter Parametersatz: Verschiebung in x und ein Maßstabsparameter). Aus den Parallaxendifferenzen wurde eine Disparitätskarte der aktuellen Auflösungsstufe gebildet. Diese diente als Näherungswert für die nächst höhere Auflösungsstufe. Dieser Vorgang wurde bis zum Erreichen der vollen Auflösungsstufe iterativ durchgeführt. Die transformierten Bildkoordinaten der homologen Punktpaare dienten dann in einem finalen Matching in den originalen Bildern als Startwerte. Die über LSM mit vollem Parameterumfang gewonnenen Bildkoordinaten können als von etwaigen Refraktionseffekten unbeeinflusst angesehen werden.

Die automatische Trennung von Land- und Unterwasserpunkten erfolgte über die Analyse der Höhenwerte, welche über konventionellen Vorwärtsschnitt (ohne Refraktionskorrektur) ermittelt wurden. Punkte unterhalb des definierten Wasserspiegels wurden über Mehrmedien-Vorwärtsschnitt (MULSOW 2018) nochmalig bestimmt. Die Berechnung erfolgte auf Basis der in der Mehrmedien-Bündelblockausgleichung ermittelten Orientierungsdaten (innere und äußere) sowie der festgelegten Trennflächenparameter (Wasserebene) und des Brechungsindex von Wasser.



Abb. 4: Gematchte Höhenpunkte auf Land (schwarz) und im Wasser (blau). Die Uferlinie ist rot markiert

2.1.3 Ergebnisanalyse-Bilddaten

Der Fokus bei den Untersuchungen lag auf dem Gewässerboden, daher konzentrieren sich die Betrachtungen der Oberflächengüte auf diese Bereiche. Wie Abb. 4 zeigt, wurden primär Unterwasserpunkte in Bereichen mit Struktur gefunden und bestimmt. Diese liegen konzentriert in den Ufer-

regionen, sowie an Kanten von Bewuchszonen. Dies verdeutlicht den Hauptnachteil der Oberflächenerfassung aus Bilddaten – die eklatante Abhängigkeit von auswertbaren Strukturen. Homogene Objektbereiche zeigen keine bzw. fehlerhafte Messungen. Es konnten Punkte bis zu einer Wassertiefe von ca. 4m (bei einer max. Gewässertiefe von ca. 4.5m) hinreichend zuverlässig bestimmt werden (siehe auch Abb. 5).



Abb. 5: Interpoliertes Höhenmodell aus Daten des Bildblockes mit unterlegtem Orthophotomosaik. Die Höhen sind relativ zum Wasserspiegel (Höhe 0) in [m] angegeben

Die Genauigkeit kann über das Tiefenspektrum hinweg keineswegs als homogen angesehen werden. Wie erwartet, fällt die Abbildungsqualität mit steigender Wassertiefe ab. Daraus lässt sich auch ein Abfallen der Genauigkeit ableiten. In Ermangelung von Referenzmessungen im Wasser kann dieser Effekt nicht hinreichend quantifiziert werden.

Die hohe Überlappung (längs und quer) im Bildblock und somit der Punktwolken aus den Stereobildpaaren erlaubt eine Abschätzung der relativen Genauigkeit innerhalb der Gesamtpunktwolke. Für diesen Zweck wurden die Punktwolken zweier Bildpaare aus verschiedenen Flugstreifen verglichen (2915/2921 und 3088/3093). In beide Punktwolken wurde zunächst jeweils eine kubische Spline-Oberfläche eingepasst und Höhenwerte auf ein 0.5m Raster gerechnet. Über den Vergleich der Höhenwerte in jedem Rasterpunkt konnte eine hohe Übereinstimmung beider Modelle festgestellt werden (RMS 4.4 cm, Abweichungen im Mittel im mm-Bereich, also kein Offset). Einzig Bereiche mit wenigen Messungen (kontrastarm, südlicher Bereich) zeigen deutlich höhere Abweichungen (max. 22 cm). Die in Tab.1 angegebene Höhengenauigkeit für Unterwasserpunkte konnte damit bestätigt werden. Die Schnittgeometrie der homologen Bildstrahlen von Unterwasserpunkten ist theoretisch schlechter als von vergleichbaren Landpunkten. Aufgrund der Brechung der Strahlen hin zum Oberflächenlot (höherer Brechungsindex von Wasser) verkleinert sich dementsprechend der Schnittwinkel im Medium Wasser.

Zur visuellen Kontrolle des Modells wurde das DHM mit dem Wasserspiegel verschnitten und in den Bildraum projiziert. Wie Abb. 7 zeigt, folgt die berechnete Schnittlinie der realen Uferlinie im Bereich weniger Dezimeter.



Abb. 6: Analyse der relativen Genauigkeit anhand von Höhenabweichungen (in [m]) von überlappenden Punktwolken aus Stereobildpaaren (2915/2921 [Sterne], 3088/3093 [Kreise])

3 Laserbathymetrie

3.1 Orientierung Laserblock

Die zeitgleich mit den Bildern aufgezeichneten Messungen des topo-bathymetrischen Laserscannersystems wurden separat ausgewertet. Zur Verfügung standen neben der Trajektorie der 4 Flugstreifen (GNSS/INS), die mittels Kalman Filterung in der Applanix Sofware Pospack⁵ ausgewertet

⁵ https://www.applanix.com/products/pospac-mms.htm

wurde, die Scannermessungen Entfernung, Auslenkwinkel, sowie zusätzliche Attribute wie kalibrierte Amplitude und Reflektivität pro Echo, die aus der online Analyse der aufgezeichneten vollen Wellenform ermittelt wurden (PFENNIGBAUER et al. 2014). In Analogie zur Bündelausgleichung des Bildblocks wurde zunächst eine Laserscanning Streifenausgleichung mit dynamischer Trajektorienkorrektur (GLIRA et al. 2016) durchgeführt. Die im Anschluss an die Streifenausgleichung durchgeführte Qualitätskontrolle (RESSL et al. 2008) des Flugblocks ergab eine robust abgeschätzte Standardabweichung (σ_{MAD}) von 1.0.cm. Diese wurde aus den Höhendifferenzen in glatten Bereichen überlappender Flugstreifen bestimmt.

Der Vorteil des robusten Streuungsmaßes σ_{MAD} gegenüber der normalen Standardabweichung (σ =6.5cm) ist die Unempfindlichkeit gegenüber Ausreißern. Da die Maskierung der glatten Bereiche, welche für die Qualitätsbeurteilung maßgeblich ist, nicht perfekt ist, verfälschen Höhendifferenzen im rauen Geländebereich (z.B. am Rande von Vegetation oder Gebäuden) die Standardabweichung durch das Quadrieren der Residuen ungebührend stark. Es sei ferner noch angemerkt, dass die Höhendifferenzen nach der Streifenausgleichung im Mittel Null betragen, die Schätzung daher unverfälscht (un-biased) ist. Da keine für die absolute Georeferenzierung des Laserblocks erforderlichen Passflächen vorlagen, wurde dieser nach der Streifenausgleichung durch eine Starrkörpertransformation auf den bereits über die Passpunkte absolut orientierten Bildblock eingepasst.

3.2 Bestimmung Wasserspiegel und Refraktionskorrektur

Die Laserstrahlen werden an der Wasseroberfläche beim Eintritt in das optisch dichtere Medium Wasser zum Lot hin gebrochen. Gleichzeitig verringert sich die Ausbreitungsgeschwindigkeit. Beide Phänomene sind durch das Brechungsgesetz von Snellius festgelegt und hängen von den Brechnungsindizes in Luft ($n_L \approx 1.0$) und Wasser ($n_W \approx 1.33$) ab. Die unkorrigierten Messstrahlen erscheinen dabei im Gegensatz zur Mehrmedienphotogrammetrie zu tief.

Für die Refraktions- und Laufzeitkorrektur ist daher ein digitales Modell der Wasseroberfläche erforderlich. Laserbathymetrie weist dabei den Vorteil auf, dass Reflexionen des Lasersignals an der Wasseroberfläche genauso empfangen werden wie vom Gewässergrund. Aufgrund des hohen Grads an gerichteter Reflexion, ist die Signalantwort grüner Laserstrahlung von der Luft-Wasser-Grenzschicht meist eine Mischung aus Reflexionen von der Oberfläche und Volumenstreuung aus den ersten cm der Wassersäule (GUENTHER et al. 2000). Für die Festlegung der Wasseroberfläche wurde daher der statistische Ansatz von MANDBURGER et al. (2013) verwendet. Dabei werden alle Laserechos in einem Höhenband von bis zu 50cm unterhalb einer grob festgelegten Näherungshöhe analysiert und in ein Raster einsortiert. Die Wasseroberfläche wird anschließend je nach Gewässertrübe als hohes Quantil (95%-99%) der einzelnen Zellpunkte berechnet. Für die relativ klaren Wasserbedingungen des Autobahnsees wurde das 99% Quantil und eine Zellgröße von 10m verwendet. Das derart berechnete Wasseroberflächenmodell weist eine maximale Undulation von 15cm auf, wobei die Wasseroberfläche im Bereich südlich der Insel tendenziell etwas höher liegt als der umgebende Wasserspiegel. Eine Anhebung der Wasserstandshöhe ist auch bei den Randpixeln im Übergang vom benetzten zum trockenen Bereich zu beobachten und dort auf eine Überschätzung der Wasseroberfläche durch gelegentliche Landpunkte zurückzuführen. Die maximale Wasserspiegelhöhe wurde daher mit h=509.10m begrenzt, wobei die mittlere aus den Laserdaten ermittelte Höhe bei 509.00m liegt. Der Mittelwert ist damit um 5cm tiefer als die für den Bildblock als konstant angenommene Höhe von 509.05cm. Die eigentliche Refraktionskorrektur erfolgte schließlich für jedes Echo durch Verschnitt des Laserstrahlvektors (ray tracing) mit dem Wasseroberflächenmodell und anschließender Refraktions- und Laufzeitkorrektur des Laserstrahlverlaufs im Wasser. Dadurch erfährt jeder rohe Unterwasser-3D-Punkt sowohl eine Lage- als auch eine Höhenkorrektur. Die praktische Durchführung erfolgte mit dem Programmsystem OPALS (PFEI-FER et al. 2014).



Abb. 6: Interpoliertes Höhenmodell aus Laserscannerdaten mit unterlegtem Orthophotomosaik. Die Höhen sind relativ zum Wasserspiegel (Höhe 0) angegeben.

3.3 Filterung und DGM Berechnung:

Nach den oben beschriebenen Vorbereitungsschritten erfolgte die Filterung der Punktwolke in Boden (trocken und benetzt) und Nicht-Boden (niedrige, mittlere, hohe Vegetation, Gebäude, etc.) mittels hierarchischer robuster Interpolation (PFEIFER et al. 2015). Eine zusätzliche Klassifizierung in Seeboden, Seeoberfläche und Punkten aus der Wassersäule erfolgte im Nachgang auf Basis des Wasseroberflächenmodells. Aus den land- und wasserseitigen Bodenpunkten wurde schließlich ein regelmäßiges DGM Raster mit einer Gitterweite von 50cm berechnet (Abb. 6). Während wegen der hohen Punktdichte von mehr als 25 Punkten/m² prinzipiell auch eine noch kleinere Gitterweite möglich gewesen wäre, wird die effektiv erzielbare räumliche Auflösung vom Durchmesser des Laserstrahlkegels (50-60cm) beschränkt.

Für die Abschätzung des Genauigkeitspotenzials der 3D Punktwolke ist neben der Genauigkeit der Trajektorie (GNSS/INS) und der Entfernungs- und Auslenkwinkelmessung auch die Unsicherheit in der Bestimmung der genauen Lage und lokalen Neigung der Wasseroberfläche und des Brechungsindexes zu berücksichtigen. Aus der Erfahrung bei Projekten ähnlicher Messtiefe und

Gewässertrübung kann ein Höhenfehler von ≤10cm als konservative Schätzung angegeben werden. Da für die DGM Berechnung praktisch flächendeckend Punkte aus allen vier Flugstreifen vorhanden sind, kann für die relative DGM Genauigkeit (precision) von einem Wert von etwa 5cm ausgegangen werden.

4 Vergleich der Resultate

Der Vergleich der Höhenmodelle aus Bildauswertung und Laserscanning offenbart eine hohe Übereinstimmung für Uferzonen ohne Bewuchs sowie für Gewässerbereiche mit ausreichend auswertbarer Textur im Bildmaterial (siehe Abb. 7). Abweichungen über ±20cm wurden aus genannten Gründen vom Vergleich ausgeschlossen. Der Abdeckungsgrad des aus Laserbathymetriedaten gewonnenen DHM ist klar höher als das aus Bilddaten generierte DHM. Der Vorteil des aktiven Messverfahrens Laserscanning gegenüber der passiven Bildaufnahme bei texturschwachen Messobjekten wird hier abermals deutlich.

Die Absolutgenauigkeit kann leider in Ermangelung von Referenzmessungen im See nicht durchgreifend analysiert werden. Eine hinreichend aussagekräftige Einschätzung ist jedoch über die Berechnung der jeweiligen Uferlinien durch Schnitt aus DHM und Wasserspiegel möglich. Die Projektion der berechneten Uferlinie in das Bildmaterial erlaubt die visuelle Beurteilung der Übereinstimmung mit der realen Uferlinie. Aus Abb. 8 ist ersichtlich, dass die Schnittlinie aus Laserdaten (rote Linie) der realen Wasserlinie im Bereich von wenigen Dezimetern folgt. Bei der Beurteilung gilt es jedoch zu bedenken, dass das Gelände im Uferbereich recht flach verläuft und daher der Schnitt mit der Wasseroberfläche schleifend ist.

Die Uferlinie aus dem Bild-DHM folgt in den Uferbereichen ohne Bewuchs der realen Wasserlinie und somit der Linie aus Laserdaten ebenfalls im hohen Maße. In Uferbereichen mit Bewuchs wandert die Linie in Richtung See, da dort durch Vegetation verdeckte Wasserbereiche nicht in den Bildern erfasst werden konnte. Hier zeigt sich ein weiterer Vorteil der Lasereraufnahme, da diese durch Echoauswertung (Full-Waveform Analyse) auch teilweise verdeckte Bereiche erfassen kann.



Abb. 7: Höhenunterschiede im Spektrum -± 20cm zwischen den DHM's aus Laserscannerdaten und Bildblock. Grau hinterlegte Bereiche zeigen höhere Abweichungen. Die rote Linie kennzeichnet den Uferverlauf.



Abb. 8: Uferlinien generiert aus Schnitt der jeweiligen Höhenmodelle mit dem Wasserspiegel (Rot - DHM aus Laserdaten; Blau - DHM aus Bilddaten).

5 Fazit & Ausblick

Mit den Methoden Laserbathymetrie und Mehrmedienphotogrammetrie stehen dem potentiellen Anwender zwei leistungsfähige Optionen zur Erfassung von Gewässerböden zur Verfügung. Beide unterscheiden sich hinsichtlich des Messprinzips grundlegend voneinander und weisen daher Stärken und Schwächen in den verschiedenen Charakteristika auf. Beiden gemein ist die Beeinflussung durch das Medium Wasser. Sei es durch Refraktion beim Trennflächendurchgang zwischen Luft und Wasser oder durch Dispersion und Trübung. Eine klare Empfehlung für die eine oder andere Methode kann nicht ausgesprochen werden, da die erreichbaren inneren Genauigkeiten ähnlich sind (5cm vs. 4.4cm, Streifenvergleich Laser vs. Streifenvergleich Stereobildpaare). Wie bereits festgestellt, hat die Laserbathymetrie den entscheidenden Vorteil der Unabhängigkeit von photogrammetrisch auswertbarer Textur. Dem gegenüber steht der deutlich geringere technische Aufwand für die Luftbilderfassung. Vor allem auch im Hinblick auf flexibel einsetzbare unbemannte Flugsysteme, die mit kostengünstigen Kameras ausgestattet sind, besteht mit der Mehrmedienphotogrammetrie ein neues Anwendungsfeld für die Erfassung von Flachwasserbereichen. Im Rahmen der Bildblock-Auswertung wurden diverse Möglichkeiten zur Optimierung identifiziert. Bereits bei der Flugplanung sollte den speziellen Anforderungen an eine Unterwasserpunktbestimmung Rechnung getragen werden. Für die Realisierung größerer Triangulationswinkel sollten Objektive mit größerem Öffnungswinkel zum Einsatz kommen. Denkbar wäre auch die Nutzung von Schrägaufnahmen. Hier ist jedoch zu beachten, die Aufnahmerichtung nicht zu sehr aus der Lotrechten zu schwenken da ansonsten Welleneffekte und Totalreflektionen die Abbildungsqualität zu stark abfallen lassen würden. Für eine stabilere Lagerung des Blockes sowie für eine durchgreifende Kontrolle der Ergebnisse sollten Passpunkte am Gewässerboden signalisiert und eingemessen werden. Dies ist auch für die Laserbathymetrie zu empfehlen.

Weiterhin sollten die Möglichkeiten der automatischen Extraktion der Wasserlinie in den Bildern untersucht werden. Wie bereits in Abb. 8 verdeutlicht, kann anhand der Helligkeits- und Farbunterschiede auf die Grenzlinie zwischen Wasser und Ufer geschlossen werden (siehe auch KRÖNERT & MEICHSNER 2017 und MULSOW et al. 2014). Diese könnte extrahiert und als zusätzliche Zwangsbedingung (Höhenpasspunktlinie) in der Bündelblockausgleichung integriert werden. Besonders gut geeignet sind in diesem Zusammenhang die Infrarotkanäle von Multispektralkameras wegen des sehr ausgeprägten Grauwertunterschiedes zwischen Wasser und Land infolge der hohen Signalabsorption infraroter Strahlung im Medium Wasser.

6 Danksagung

Der Beitrag von Gottfried Mandlburger wurde durch Mitteln der Deutschen Forschungsgesellschaft (DFG) im Rahmen des Forschungsprojektes "Bathymetrievermessung durch Fusion von Flugzeuglaserscanning und multispektralen Luftbildern" unterstützt.

7 Literaturverzeichnis

- GLIRA, P., PFEIFER, N. & MANDLBURGER, G., 2016: Rigorous Strip Adjustment of UAV-based Laserscanning Data Including Time-Dependent Correction of Trajectory Errors. Photogrammetric Engineering & Remote Sensing, 82(12), 945-954.
- GUENTHER, G.C., CUNNINGHAM, A.G., LAROQUE, P.E. & REID, D.J., 2000: Meeting the accuracy challenge in airborne lidar bathymetry. In: Proceedings of the 20th EARSeL Symposium: Workshop on Lidar Remote Sensing of Land and Sea, Dresden.
- KRÖHNERT, M. & MEICHSNER, R., 2017: Segmentation of environmental time lapse image sequences for the determination of shore lines captured by hand-held smartphone cameras. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial In-formation Sciences IV-2/W4 (2017), 1-8.
- MAAS, H.-G., 2015: On the Accuracy Potential in Underwater/ Multimedia Photogrammetry. Sensors, **15**(8), 18140-18152.
- MANDLBURGER, G., 2018: A Case Study on Through-Water Dense Image Matching. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **42**(2), 659-666.
- MANDLBURGER, G., PFENNIGBAUER, M. & PFEIFER, N., 2013: Analyzing near water surface penetration in laser bathymetry - A case study at the River Pielach. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **II**-5/W2, 175-180.
- MANDLBURGER, G., WENZEL, K., SPITZER, A., HAALA, N., GLIRA, P. & PFEIFER, N., 2017: Improved Topographic Models via Concurrent Airborne LIDAR and Dense Image Matching, ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-2/W4, 259-266.
- MANDLBURGER, G., KREMER, J., STEINBACHER, D. & BARAN, R., 2018: Investigating the use of coastal blue imagery for bathymetric mapping of inland water bodies. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(1), 275-282. 10.5194/isprs-archives-XLII-1-275-2018.
- MULSOW, C., 2010: A flexible multi-media bundle approach. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, **38**(5).
- MULSOW, C., 2016: Ein universeller Ansatz zur Mehrmedien-Bündeltriangulation. In: Photogrammetrie - Laserscanning - Optische 3D-Messtechnik, Beiträge der Oldenburger 3D-Tage 2016, Th. Luhmann/Ch. Schumacher (Hrsg.), Verlag Herbert Wichmann, 266-273.
- MULSOW, C., 2018: Digital elevation models of underwater structures from UAV imagery. In: Hydrographische Nachrichten, Journal of Applied Hydrography, HN110, 14-19.
- MULSOW, C., KOSCHITZKI, R. & MAAS, H.-G., 2014: Photogrammetric monitoring of glacier margin lakes. Geomatics, Natural Hazards and Risk
- PFEIFER, N., MANDLBURGER, G. & GLIRA, P., 2015: Laserscanning. In C. Heipke (Ed.), Photogrammetrie und Fernerkundung (1st ed., pp. 1–51). Berlin Heidelberg: Springer.
- 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.

- PFENNIGBAUER, M., WOLF, C., WEINKOPF, J. & ULLRICH, A., 2014: Online waveform processing for demanding target situations. In Proc. SPIE, 90800J.
- RESSL, C., KAGER, H. & MANDLBURGER, G., 2008: Quality checking of ALS projects using statistics of strip differences. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol XXXVII, 253-260.
- QUAN, X. & FRY, E., 1995: Empirical equation for the index of refraction of seawater, Appl. Opt. 34, 3477-3480

Analyse der Auswirkung von Wellen auf konventionelle Refraktionskorrekturmethoden in der Laserbathymetrie

KATJA RICHTER¹, DAVID MADER¹, PATRICK WESTFELD² & HANS-GERD MAAS¹

Zusammenfassung: Die geometrisch korrekte Erfassung des Gewässerbodens in der Laserbathymetrie erfordert die Korrektur der Refraktion auf Basis des Snellius'schen Gesetzes. Voraussetzung für eine strenge Refraktionskorrektur sind genaue Informationen über die Geometrie der lokalen wellen-induzierten Wasseroberfläche. Da diese Informationen in der Regel nicht verfügbar sind, basieren konventionelle Methoden zur Refraktionskorrektur auf der Einführung bestimmter Vereinfachungen. Meist werden Wellenmuster und Strahldivergenz vernachlässigt und eine horizontale oder eine lokal geneigte ebene Wasseroberfläche sowie ein infinitesimal schmaler Laserpuls angenommen. Abweichungen von den getroffenen Annahmen können zu einem signifikanten Lage- und Höhenversatz am Gewässerboden führen.

Zur Prognose der wellen-induzierten Koordinatenverschiebungen für beliebige Wellenverhältnisse wurde ein Simulationsansatz entwickelt. Das Kernstück der Simulation bildet die streng differentielle Modellierung der Refraktion eines Laserstrahls mit endlichem Durchmesser an einer bewegten Wasseroberfläche. Außerdem wird die Refraktion eines infinitesimal schmalen Laserstrahls an einer horizontalen und an einer lokal geneigten Wasseroberfläche modelliert. Der Vergleich der resultierenden Gewässerbodenpunktkoordinaten erlaubt eine Analyse der Auswirkung von Wellen auf die Refraktionskorrektur. Zur Validierung stehen die Ergebnisse einer experimentellen Untersuchung in einer kontrollierten Umgebung in einem Wellenbad zur Verfügung.

1 Einleitung

Für die Korrektur der Refraktion auf Basis des Snellius'schen Gesetzes ist die Kenntnis des Einfallswinkels des Laserstrahls auf die lokale Wasseroberfläche erforderlich (Abb. 1, αlokal). Die wellen-induzierte Neigung der Wasseroberfläche ist allerdings in der Regel unbekannt und stellt somit eine Herausforderung für die Refraktionskorrektur in der Laserbathymetrie dar. Konventionelle Korrekturmethoden überwinden die Problematik der fehlenden Wasseroberflächeninformationen durch die Einführung verschiedener Vereinfachungen. Am simpelsten ist die Annahme einer horizontalen ebenen Wasseroberfläche, an welcher der Laserstrahl gebrochen wird (Abb. 1, lila). Schon bei kleineren Wellen kann die Abweichung von der Ebenheit zu einem signifikanten Lageund Höhenversatz am Gewässerboden führen (dXY_{hz und} dZ_{hz}). Komplexere Korrekturmethoden versuchen daher, die tatsächliche Geometrie der Wasseroberfläche zu berücksichtigen, indem eine lokale Wasseroberflächenneigung für die Refraktionskorrektur verwendet wird (Abb. 1, rot). Grundlage für die Bestimmung der lokalen Wasseroberflächenneigung ist ein Wasseroberflächenmodell, das auf den detektierten Wasseroberflächenpunkten basiert, die beispielsweise zu einem

¹ Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Helmholtzstraße 10, D-01069 Dresden, E-Mail: [katja.richter1, david.mader, hans-gerd.maas]@tu-dresden.de

² Bundesamt f
ür Seeschifffahrt und Hydrographie, Sachgebiet Geod
ätisch-hydrographische Verfahren und Systeme, Neptunallee 5, D-18057 Rostock, E-Mail: Patrick.Westfeld@bsh.de

Dreiecksnetz vermascht werden (ULLRICH & PFENNIGBAUER 2011). Durch Verschneidung des einfallenden Laserstrahls mit dem Dreiecksnetz kann die lokale Wasseroberflächenneigung abgeleitet



Abb. 1: Refraktion an der lokalen wellen-induzierten Wasseroberfläche (blau), der angenommenen horizontalen Wasseroberfläche (lila) und der angenommenen lokal geneigten Wasseroberfläche (rot), die im Lageversatz dXY_{hz} und dXY_{tilt} und Höhenversatz dZ_{hz} und dZ_{tilt} resultiert.

werden. Auch bei dieser Korrekturmethode können Abweichungen zwischen der wahren Wasseroberfläche und der Repräsentation durch das Dreiecksnetz zu Koordinatenfehlern am Gewässerboden führen. Der Versatz in Lage (dXY_{tilt}) und Höhe (dZ_{tilt}) fällt jedoch geringer aus als bei der ersten Methode. Neben der Vereinfachung der Wasseroberflächengeometrie führen beide Korrekturansätze Annahmen zur Geometrie des Laserstrahls ein, welcher als infinitesimal schmale Linie betrachtet wird.

In vorangegangenen Studien wurde ein Simulationsansatz entwickelt, mit dessen Hilfe die welleninduzierten Koordinatenverschiebungen für beliebige Wellenverhältnisse prognostiziert werden können (WESTFELD et al. 2017). Es hat sich gezeigt, dass die laterale Verschiebung am Gewässerboden in Abhängigkeit vom Seegang bis zu 10% der Wassertiefe beträgt. Darüber hinaus muss je nach Gewässertiefe mit Höhenabweichungen im Zentimeter- bis Dezimeterbereich gerechnet werden. Die Auswirkungen der Wellenbewegungen auf die Koordinaten am Gewässerboden sind hauptsächlich stochastisch, enthalten aber auch signifikante systematische Anteile.

Dieser Beitrag zeigt eine Weiterentwicklung des bestehenden Simulationsansatzes, welche noch realistischere Prognosen der durch Wellen verursachten Koordinatenverschiebungen ermöglicht. Bisher war die Simulation auf die Betrachtung eines statischen Laserstrahls beschränkt, unter dem sich die Wellen hinwegbewegen. Für ein realitätsnahes Simulationsergebnis wird im weiterentwickelten Ansatz die Bewegung der Plattform sowie der Scanmechanismus des Laserstrahls in die Simulation integriert.



Abb. 2: Schwimmbecken während der Datenerfassung mit dem bathymetrischen Flugzeuglaserscanner RIEGL VQ 820G (a), resultierende ALB Punktwolke mit Wasseroberflächenpunkten in blau und Gewässerbodenpunkten in grau (b), Erfassung der Beckengeometrie mit dem terrestrischen Laserscanner RIEGL LMS-Z420i (c) und resultierende Referenzpunktwolke (d).

Zur Validierung der Simulationsergebnisse stehen Laserbathymetriedaten und terrestrische Referenzdaten eines Wellenfreibads zur Verfügung, welche die experimentelle Analyse der Auswirkung von Wellen auf konventionelle Refraktionskorrekturmethoden ermöglichen (Abb. 2). Im Beitrag wird untersucht, inwieweit die in der Simulation vorhergesagten Koordinatenfehler zu den experimentell ermittelten Werten korrespondieren.

Das Freibad bietet kontrollierte Untersuchungsbedingungen mit zuverlässig produzierbaren Wellen, geringer Gewässertrübung und einer festen und präzise messbaren Gewässerbodengeometrie. Allerdings erreicht der schräg abfallende Boden des 12 m x 50 m großen Wellenbeckens bei horizontaler Wasseroberfläche nur eine maximale Wassertiefe von 1,6 m. Die Laserbathymetriedaten wurden mit dem Flugzeuglaserscannersystem RIEGL VQ-820-G in verschiedenen Flughöhen (500 m, 600 m, 700 m) und Flugrichtungen aufgenommen. Bei einer Strahldivergenz von 1 mrad ergibt sich auf der Wasseroberfläche ein Laserfootprint mit einem Durchmesser von 0,5 m bis 0,7 m. Die Erfassung der Referenzdaten erfolgte mit dem terrestrischen Laserscanner Riegl LMS-Z420i am abgelassenen Schwimmbecken. Abbildung 2 (d) zeigt die resultierende Referenzpunktwolke.

Platt	form	Sensor	
Position [x y z] in m [0 0 500/600/700]		Scanwinkel	20°
Orientierung [ω φ κ] in ° [0 0 0]		Winkelschrittweite	0,1°
Fluggeschwindigkeit	58 m/s	Strahldivergenz	1 mrad
Wasserbecken		Pulsrepetitionsrate	520 kHz
Brechungsindex Wasser 1,33		Wellen	
Grundfläche 12 x 50 m		Windgeschwindigkeit	10 m/s
Wassertiefe 1,60 m		Gitterweite der Fourier	600
Atmosphäre		Transformation	000
Brechungsindex Luft	1,000292	Windrichtung	45°

Tab. 1: Übersicht der für die Simulation genutzten Parameter

2 Numerische Simulation

Konventionelle Refraktionskorrekturmethoden vereinfachen die Geometrie der Wasseroberfläche und des Laserstrahls, um die fehlenden Informationen über die wellen-induzierte Neigung der Wasseroberfläche zu kompensieren. Schon bei leichtem Seegang treten Abweichungen von den getroffenen Annahmen auf, die zu einem Lage- und Höhenversatz am Gewässerboden führen. Ziel der numerischen Simulation ist die Vorhersage der durch Wellen verursachten Koordinatenverschiebungen für beliebige Wellenmuster. Zu diesem Zweck wird zunächst eine streng differentielle Modellierung der Refraktion eines Laserstrahls mit endlichem Durchmesser an einer welleninduzierten Wasseroberfläche durchgeführt (Abschnitt 2.1). Außerdem wird die Refraktion entsprechend der in den konventionellen Korrekturmethoden eingeführten Vereinfachungen modelliert (Abschnitt 2.2). Der Vergleich der resultierenden Gewässerbodenpunktkoordinaten erlaubt eine Analyse der Auswirkung von Wellen auf die Refraktionskorrektur.

2.1 Streng differentielle Modellierung der Refraktion

Die möglichst realistische Simulation der Refraktion eines Laserstrahls mit endlichem Durchmesser an einer zeitlich veränderlichen wellen-induzierten Wasseroberfläche erfordert die Modellierung der Wasseroberfläche, des Gewässerbodens, der Plattformbewegung, des Scanmechanismus und der Strahldivergenz. Die Wahl aller notwendigen Parameter (z.B. Amplitude der Wellen, Wassertiefe, Fluggeschwindigkeit, etc.) orientiert sich an den Experimenten im Wellenbad. Damit wird die Vergleichbarkeit von Simulation und experimenteller Analyse gewährleistet. In Tabelle 1 sind die wichtigsten Parameter zusammengefasst.

Für die Modellierung der wellen-induzierten Wasseroberfläche wird das von TESSENDORF (2001) entwickelte Modell für Ozeanwellen verwendet. Mit Hilfe des Modells wird für jeden Zeitpunkt *t* ein Höhenfeld in Form eines dichten regelmäßigen Rasters generiert. Das Höhenfeld repräsentiert eine realistische Ozeanoberfläche und kann über verschiedene Parameter (z. B. Gitterweite der Fourier Transformation, Windgeschwindigkeit, Windrichtung) modifiziert werden. Die Parameterwahl erfolgt manuell, d.h die Parameter werden nicht durch Inversion des Tessendorfmodells aus den realen Messdaten abgeleitet. Bei der Parameterwahl wird angestrebt, das Wasseroberflächenmodell möglichst gut an die Wellensituation des Experimentes anzupassen. Ziel ist eine wirklichkeitsgetreue Reproduktion des realen Wellenmusters in der Simulation.



Abb. 3: Profil der Wasseroberfläche in Messdaten (a) und Simulation (b).

Zunächst erfolgt eine Analyse des realen Wellenmusters in den Messdaten. Zu diesem Zweck werden jeweils 50 cm breite Schnitte durch die Punktwolke der Wasseroberflächenpunkte zu Profilen zusammengefasst und mit Spline-Funktionen approximiert. Aus den lokalen Maxima und Minima können die maximale Amplitude und Wellenlänge des Wellenmusters abgeleitet werden. Abbildung 3(a) zeigt ein typisches Profil. Der Ursprung der Wellen ist auf der rechten Seite, wobei die Wassertiefe nach links hin abnimmt. Die horizontale Linie repräsentiert die mittlere Wasserhöhe. Anschließend wird durch die Wahl geeigneter Simulationsparameter eine Wasseroberfläche mit möglichst ähnlichen Eigenschaften modelliert. Das entsprechende Wellenprofil ist in Abbildung 3(b) dargestellt.

Die Modellierung des Gewässerbodens orientiert sich an der ebenen Charakteristik des realen Gewässerbodens. Auf das Gefälle des Beckenbodens wird verzichtet, da die Auswirkung der Neigung im Verhältnis zur beleuchteten Fläche nicht signifikant ist. Der simulierte Gewässerboden wird demnach als horizontale Ebene generiert. Die Plattformbewegung wird als gleichförmige geradlinige Translation mit konstanten Kurswinkeln modelliert. Flugrichtung, Fluggeschwindigkeit und Flughöhe orientieren sich an der Erfassung der Laserbathymetriedaten im Wellenbad.

Der Scanmechanismus wird entsprechend den Spezifikationen des verwendeten Flugzeuglaserscanners RIEGL VQ-820-G als Palmer-Scan simuliert. Die rotierende Scanachse wird in Bezug zur Flugrichtung um 20° geneigt und erzeugt am Boden ein elliptisch-bogenförmiges Scanmuster. Scangeschwindigkeit, Winkelschrittweite und Pulsrepetitionsrate werden den Herstellerangaben entsprechend gewählt. Zur Modellierung der Strahldivergenz wird der Laserstrahl in eine große Anzahl von Teilstrahlen aufgespalten, die auf der Wasseroberfläche einen Laserfootprint mit endlichem Durchmesser repräsentieren. Der Intensitätsverlauf innerhalb des einfallenden Laserpulses folgt einer gaußförmigen Intensitätsverteilung, d. h die Intensität eines einzelnen Teilstrahls ergibt sich aus seiner Lage im finiten Laserpuls.

Die Modellierung der Refraktion erfolgt differentiell, indem jeder Teilstrahl separat betrachtet wird. Zunächst werden die Schnittpunkte mit dem Wasseroberflächenmodell und die lokalen Wellenneigungen bestimmt. Anschließend wird für jeden Teilstrahl die Refraktion an der Wasseroberfläche nach dem Snellius'schen Brechungsgesetz modelliert und der Auftreffpunkt auf dem simulierten Gewässerboden berechnet. Der intensitätsgewichtete Schwerpunkt der Gewässerbodenpunkte aller Teilstrahlen repräsentiert die finale Bodenpunktkoordinate für den gesamten Laserstrahl. Die Simulation basiert auf der Annahme eines identischen Hin- und Rückwegs. Die Auswirkungen von Mehrfachstreuung und Dispersion werden vernachlässigt.

2.2 Konventionelle Modellierung der Refraktion

Ergebnis der streng differentiellen Modellierung der Refraktion sind neben den Gewässerbodenkoordinaten auch die zugehörigen Wasseroberflächenpunkte. Je nach Punktdichte repräsentieren sie die tatsächliche Wasseroberfläche mehr oder weniger gut. Wie bei einer realen Messung können die Wasseroberflächenpunkte für die konventionelle Refraktionskorrektur verwendet werden. In der Praxis existieren verschiedene Korrekturansätze, die auf unterschiedlichen Annahmen basieren. In der Simulation wurden folgende Methoden umgesetzt:

- Modellierung der Refraktion eines infinitesimal schmalen Laserstrahls an einem horizontalen Wasseroberflächenelement (Methode 1)
- Modellierung der Refraktion eines infinitesimal schmalen Laserstrahls an einem lokal geneigten Wasseroberflächenelement (Methode 2)

Für die Modellierung der Refraktion nach der ersten Methode wird an jedem Wasseroberflächenpunkt ein lokales horizontal orientiertes Wasseroberflächenelement definiert, an welchem der infinitesimal schmale Laserstrahl nach Snellius gebrochen wird (Abb. 4). Die Höhe des Wasseroberflächenelements entspricht der Höhe des zugehörigen Wasseroberflächenpunktes. Mit dem berechneten Brechungswinkel wird die Richtung des aus der strengen Refraktionskorrektur resultierenden Vektors zwischen Wasseroberflächenpunkt und Gewässerbodenpunkt angepasst und die finale Bodenpunktkoordinate berechnet.

Für die zweite Methode wird zunächst aus allen Wasseroberflächenpunkten mittels Delaunay-Triangulation ein Dreiecksnetz erstellt (Abb. 5, grau). Die Dreiecke repräsentieren die geneigten Wasseroberflächenelemente. Da an den Eckpunkten der Dreiecke keine lokalen Neigungen bestimmt werden können, wird das Dreiecksnetz in einem regelmäßigen Raster hochauflösend abgetastet (Abb. 5, rot). Die Koordinaten der Rasterpunkte werden über lineare Interpolation im Dreiecksnetz bestimmt. Anschließend wird der einfallende Laserstrahl mit dem gesampelten Wasseroberflächenmodell verschnitten. Der für das Snellius'sche Brechungsgesetz benötigte Einfallswinkel wird in Bezug zur Oberflächennormalen des geschnittenen Dreiecks berechnet.



Abb. 5: Modellierung der Refraktion eines infinitesimal schmalen Laserstrahls an einem horizontalen Wasseroberflächenelement (a) und an einem lokal geneigten Wasseroberflächenelement (b).

3 Ergebnisse

Die numerische Simulation wurde entsprechend der Aufnahme der Laserbathymetriedaten im Wellenbad für verschiedene Flughöhen durchgeführt. Die Ausmaße des simulierten Gebietes orientieren sich an der Größe des Wellenbeckens. Insgesamt wurden zwischen 1400 und 2200 Laserpulse analysiert. Die Ergebnisse der numerischen Simulation sind in Tabelle 2 zusammengefasst. Der Koordinatenversatz am Gewässerboden setzt sich aus den lateralen Komponenten dX und dY sowie der Tiefenkomponente dZ zusammen. Da die Auswirkungen der Wellen auf die Refraktion linear mit der Wassertiefe zunehmen, werden alle Ergebnisse in Prozent der Wassertiefe angegeben.

	Flug-	dX		dY			dZ			
	höhe	min.	max.	RMS	min.	max.	RMS	min.	max.	RMS
Methode 1	500 m	-4,50	5,59	1,16	-5,21	5,09	1,18	-1,59	2,15	0,40
	600 m	-4,41	3,82	1,16	-3,95	4,01	1,14	-1,45	1,96	0,43
	700 m	-4,24	3,71	1,13	-3,91	4,46	1,18	-1,70	1,86	0,46
Methode 2	500 m	-4,54	3,25	0,78	-2,96	3,16	0,71	-1,31	1,43	0,32
	600 m	-3,16	2,76	0,68	-3,76	3,50	0,72	-1,31	1,79	0,34
	700 m	-3,26	3,11	0,72	-3,65	3,51	0,79	-1,75	1,55	0,40

Tab. 2: Koordinatenversatz am Gewässerboden in Prozent der Wassertiefe

Bei einer Flughöhe von 500m ergibt sich für die Modellierung der Refraktion mit horizontalen Wasseroberflächenelementen (Methode 1) ein RMS-Wert (root mean square) von 1,16% (max. 5,59%) für die X-Komponente dX_{hz} und 1,18% (max. 5,09%) für die Y-Komponente dY_{hz}. Die Modellierung der Refraktion mit lokal geneigten Wasseroberflächenelementen (Methode 2) resultiert in einem RMS-Wert von 0,78% (max. 3,25%) für dX_{tilt} bzw. 0,71% (max. 3,16%) für dY_{tilt}. In Relation zur Wassertiefe des Wellenbades von 1,6 m ergibt sich für Methode 2 ein RMS-Wert von 1,3 cm (max. 5,2 cm) bzw. 1,1 cm (max. 5,1 cm). Die Werte zeigen, dass sich der laterale Koordinatenversatz mit steigender Komplexität der Wasseroberflächenrepräsentation verringert. Das gilt auch für die anderen beiden Flughöhen, wobei die Unterschiede in der Footprint-Größe kaum Auswirkung auf das Ergebnis haben. Die Tiefenkomponente dZ fällt grundsätzlich kleiner aus als die beiden Lagekomponenten. Für eine Flughöhe von 500 m variiert der RMS-Wert je nach Korrekturmethode zwischen 0,32% (dZ_{tilt}) und 0,40% (dZ_{hz}) (0,5 cm bzw. 0,6 cm). Mit steigender Flughöhe nimmt der Koordinatenversatz in Z-Richtung zu. Die Zahlenwerte beziehen sich auf einen ebenen Gewässerboden. Bei bewegtem Gelände kann der Einfluss auf die Z-Komponente deutlich höher ausfallen.

In Abbildung 6 ist der absolute Koordinatenversatz am Gewässerboden für beide Refraktionskorrekturmethoden getrennt nach den beiden Lagekomponenten und der Tiefenkomponente dargestellt. Die obere Zeile zeigt den Koordinatenversatz für die Modellierung der Refraktion an horizontalen Wasseroberflächenelementen. Unten sind die Ergebnisse für die lokal geneigten Wasseroberflächenelemente visualisiert. Die Teilabbildungen zeigen einen 20 m x 10 m großen Ausschnitt der simulierten Daten. Sowohl die Ausbreitungsrichtung der Wellen als auch die Flugrichtung liegen entlang der x-Achse. Jeder Punkt repräsentiert einen simulierten Wasseroberflächenpunkt. Die Farbcodierung entspricht dem Koordinatenversatz am Gewässerboden. Die simulierte Wasseroberfläche beinhaltet im visualisierten Ausschnitt einen Wellenberg, der schräg durch den Ausschnitt verläuft. Der Farbverlauf in der Abbildung zeigt deutlich die Auswirkung der Welle auf die Modellierung der Refraktion sowie die resultierenden Bodenpunktkoordinaten. Der Einfluss auf die Lagekomponenten ist dabei größer als der Einfluss auf die Tiefenkomponente. Der Vergleich zwischen beiden Korrekturverfahren zeigt, dass der Koordinatenversatz am Gewässerboden durch Verwendung des komplexeren Wasseroberflächenmodells in Methode 2 verringert wird.



Abb. 6: Koordinatenversatz am Gewässerboden bei einer Flughöhe von 500 m für die Modellierung der Refraktion eines infinitesimal schmalen Laserstrahls an horizontalen Wasseroberflächenelementen (oben) und an lokal geneigten Wasseroberflächenelementen (unten).

4 Validierung

Die in der numerischen Simulation prognostizierten Koordinatenverschiebungen am Gewässerboden werden zur Validierung mit experimentell ermittelten Werten verglichen. Die Grundlage für die experimentellen Untersuchungen bilden im Wellenbad erfasste Laserbathymetriedaten sowie terrestrische Referenzdaten. Die Laserbathymetriedaten liegen als unkorrigierte 3D-Punktwolke vor, die in Wasseroberflächenund Gewässerbodenpunkte klassifiziert ist. Für jeden 3D-Punkt ist ein exakter Zeitstempel verfügbar. Außerdem liegen Informationen über die Sensortrajektorie sowie Herstellerangaben bezüglich der Brechungsindizes von Luft und Wasser vor. Auf Grundlage dieser Informationen wird die Refraktion mit den beiden konventionellen Methoden modelliert (s. Abschnitt 2.2). Dabei wird auch die geringere Ausbreitungsgeschwindigkeit des Laserpulses im Wasser berücksichtigt. Zur experimentellen Untersuchung der wellen-induzierten Koordinatenverschiebungen werden die Abweichungen zwischen refraktionskorrigierter ALS-Punktwolke und TLS-Referenzpunktwolke berechnet. Die Punkte am Boden des Beckens eignen sich zur Bestimmung der Tiefenkomponente dZ. Aufgrund der Ebenheit des Bodens kann die laterale Komponente dXY nur an senk-

rechten Wänden ermittelt werden. Die Punkte an den Wänden des Beckens können nicht verwendet werden, da sie durch den schrägen Einfallswinkel in Kombination mit dem hochreflektierenden Material von starkem Rauschen beeinträchtigt sind, welches die geometrischen Effekte überlagert. Aus diesem Grund werden nur die Punkte am Betonsockel der Wasserrutsche genutzt.

In Abbildung 7 ist der Vergleich von korrigierter ALS-Punktwolke und TLS-Referenzpunktwolke am Boden des Beckens für einen Flugstreifen mit einer Flughöhe von 500 m dargestellt. Die Abweichungen zwischen den Punktwolken zeigen deutlich die Auswirkung des lokalen Wellenmusters auf die Bodenpunktkoordinaten, die Wellen pausen sich quasi zum Boden durch.



(a)



(b)

Abb. 7: Abweichung dZ zwischen refraktionskorrigierter ALS Punktwolke (farbig) und TLS Referenzpunktwolke (grau) am Beckenboden für Methode 1 (a) und Methode 2 (b).



Abb. 8: Punktdichte und Verteilung der Wasseroberflächenpunkte in Punkte pro Quadratmeter.

Im Vergleich zur Dichte und Verteilung der Wasseroberflächenpunkte in Abbildung 8 wird deutlich, dass die größten Abweichungen in Bereichen auftreten, in denen nur wenige oder keine Informationen über die Wasseroberfläche vorliegen. Die Untersuchung der lateralen Abweichungen zwischen korrigierter ALS Punktwolke und TLS Referenzpunktwolke wird durch die geringe Anzahl an verwertbaren ALS Punkten am Betonsockel der Wasserrutsche limitiert. In Tabelle 8 sind die Ergebnisse für beide Refraktionskorrekturmethoden dargestellt. Die Abweichungen variieren zwischen 1,8 cm und 15,6 cm. In Tabelle 3 sind die Ergebnisse für alle Flugstreifen zusammengefasst. Der RMS-Wert des lateralen Koordinatenversatzes dXY variiert zwischen 8,24 % und 11,01 %, während die Werte für die Tiefenkomponente vergleichsweise klein ausfallen (RMS 1,08 % bis 2,20 %). Eine Auswirkung der verschiedenen Flughöhen ist nicht sichtbar.

Insgesamt werden die in der numerischen Simulation prognostizierten Effekte durch die experimentelle Validierung bestätigt. Im Vergleich zu den experimentellen Ergebnissen ist die Simulation allerdings zu optimistisch. Die aus der experimentellen Untersuchung abgeleiteten Koordinatenfehler sind bis zu 10-mal größer als die in der Simulation vorhergesagten Werte. Die Ursachen dafür liegen sowohl in der numerischen Simulation als auch in der experimentellen Bestimmung der Koordinatenfehler begründet. Die Simulation basiert auf der Annahme, dass der Hin- und Rückweg des Laserstrahls identisch ist, unter Vernachlässigung von Mehrfachstreuung und diffuser Reflektion am Gewässerboden. Effekte, die auf dem Rückweg noch wirken könnten, werden demnach nicht modelliert. Außerdem wird die Dispersion aktuell noch nicht berücksichtigt. Durch diese Limitierungen fallen die in der numerischen Simulation bestimmten Koordinatenverschiebungen tendenziell zu klein aus. Ein weiterer Grund sind die Unterschiede zwischen realer und simulierter Wasseroberfläche. Wie in Abbildung 3 zu erkennen, zeichnet sich das simulierte Wellenmuster durch glattere Wellenberge aus. Außerdem wird der künstliche Charakter der maschinell erzeugten Wellen nicht optimal durch das Wellenmodell von TESSENDORF (2001) reproduziert. Das betrifft insbesondere die Reflektionen der Wellen an den Wänden des Schwimmbeckens.

	dXY RMS			dZ RMS			
	500 m 600 m 700 500 m 600 m 7				700		
Methode 1	8,30	-	-	2,02	2,33	1,83	
Methode 2	8,24	-	-	1,80	2,20	1,73	

Tab. 3: RMS der Abweichungen zwischen refraktionskorrigierter ALS Punktwolke und TLS Referenzpunktwolke in Prozent der Wassertiefe

Darüber hinaus unterscheidet sich die Verteilung der Wasseroberflächenpunkte in Simulation und Experiment. Die Simulation basiert auf einer homogenen Verteilung der Scanpunkte in einem bogenförmigen Scanmuster. Im Gegensatz dazu ist die Repräsentation der Wasseroberfläche in der experimentellen Untersuchung durch eine inhomogene Verteilung der Wasseroberflächenpunkte charakterisiert. Abbildung 7 zeigt ein Beispiel für eine typische Punktverteilung im Datensatz. Die Punktdichte variiert von 0 bis 8 Punkten pro Quadratmeter. Die Ergebnisse der experimentellen Untersuchung bleiben somit Mittelwerte für das gesamte Untersuchungsgebiet, in denen Bereiche mit hoher und niedriger Punktdichte gemittelt werden.

Eine weitere Ursache für die unterschiedlichen Ergebnisse von Simulation und experimenteller Untersuchung können systematische Fehler bei der Bestimmung der Wasseroberfläche aus den Messdaten sein. Da das grüne Laserlicht in die Wassersäule eindringt, repräsentiert die erste Interaktion des emittierten Laserpulses nicht exakt die Grenzfläche zwischen Luft und Wasser (GUEN-THER 1985). In ihrer Studie konnten MANDELBURGER et al. (2013) eine mittlere Eindringtiefe von 10 cm – 25 cm nachweisen. Die näherungsweise Ableitung der Wasseroberfläche aus den nahe der Wasseroberfläche detektierten Echos erfolgt mittels statistischer Analysen. Tendenziell liegt die geschätzte Wasseroberflächenhöhe etwa 3 cm – 6 cm zu tief, was bei der Refraktionskorrektur sowohl Auswirkung auf die Lage als auch auf die Höhe der Gewässerbodenpunkte hat.

5 Zusammenfassung und Ausblick

Der Beitrag analysiert die Auswirkungen von lokalen Wellenmustern auf die Genauigkeit der Gewässerbodenkoordinaten in der Laserbathymetrie. Der Einfluss der Wellen auf die Modellierung der Refraktion wurde sowohl in einer numerischen Simulation als auch in einer experimentellen Validierung untersucht. Die in der Simulation vorhergesagten Koordinatenfehler und die aus den realen Messdaten abgeleiteten Fehler zeigen eine konsistente Tendenz. Allerdings fallen die experimentell ermittelten Werte signifikant größer aus als die in der Simulation bestimmten Werte. Ursachen dafür sind Limitierungen in der numerischen Simulation und in der experimentellen Bestimmung der Koordinatenfehler.

Die Ergebnisse zeigen, dass der Einfluss von Wellen in konventionellen Refraktionskorrekturmethoden nicht ausreichend berücksichtigt wird. Im nächsten Schritt soll daher die Wasseroberfläche auf Basis der detektierten Wasseroberflächenpunkte durch ein parametrisiertes Wasseroberflächenmodell approximiert werden. Damit wird die strenge Berücksichtigung von Welleneffekten auf die Refraktion ermöglicht und die geometrische Refraktionskorrektur verbessert. Letztendlich kann ein signifikant höheres Genauigkeitspotential in der Laserbathymetrie erreicht werden.

6 Danksagung

Die durchgeführten Forschungsarbeiten werden durch die Deutsche Forschungsgemeinschaft (DFG) gefördert. Wir danken der Firma Milan Geoservices für die Erfassung der bathymetrischen Laserscannerdaten. Außerdem geht unser Dank an Stadtbäder und Freizeitanlagen GmbH Radebeul für den Zugang zur Schwimmbadanlage und die Steuerung der historischen Wellenmaschine.

7 Literatur

- GUENTHER, G. C., 1985: Airborne laser hydrography: System design and performance factors. National Oceanic and Atmospheric Administration Rockville MD.
- MANDLBURGER, G., PFENNIGBAUER, M. & PFEIFER, N., 2013: Analyzing near water surface penetration in laser bathymetry - A case study at the River Pielach. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 5, 175-180.
- TESSENDORF, J., 2001: Simulating ocean waters. ACM SIGGraph Course Notes.
- ULLRICH, A. & PFENNIGBAUER, M., 2011: Laser-Hydrographieverfahren. Patent WO 2011137465 A1. Riegl Laser Measurement Systems GmbH.
- WESTFELD, P., MAAS, H. G., RICHTER, K. & WEIB, R., 2017: Analysis and correction of ocean wave pattern induced systematic coordinate errors in airborne LiDAR bathymetry. ISPRS Journal of Photogrammetry and Remote Sensing, **128**, 314-325.

Vermessung hochalpiner Bergseen mittels moderner Technologien

DANIEL INNERHOFER¹, GOTTFRIED MANDLBURGER^{2,3}, FRANK STEINBACHER⁴ & MARKUS AUFLEGER¹

Zusammenfassung: Im Sommer 2017 fand eine Messkampagne im Tiroler Stubaital statt mit dem Ziel, verschiedene moderne Methoden der Gewässervermessung in Hinblick auf deren Anwendbarkeit im hochalpinen Bereich zu überprüfen, zu kombinieren und im Überlappungsbereich zu vergleichen. Die dabei verwendeten Messmethoden waren ein hochauflösendes Fächerecholot, Mehrmedienphotogrammetrie und bathymetrisches Flugzeuglaserscanning. Als Untersuchungsgebiet wurden zwei Seen in einer sehr exponierten Lage der Stubaier Alpen ausgewählt, der Grünausee (2328 müA) und die Blaue Lacke (2289 müA) im Bereich der Sulzenauhütte.

1 Einleitung

Die bathymetrische Erfassung von Gebirgsseen ist speziell für geomorphologische Fragestellungen von Bedeutung. Da solche Seen oft von relativ gesehen kurzer Lebendsauer sind oder, aufgrund hydrologischer und geologischer Prozesse, einem starken Wandel unterworfen sein können, bedarf es einer Vermessungsmethode die es erlaubt in kurzer Zeit die Gewässersohle zu erfassen und ggf. mehrere Messungen innerhalb eines kurzen Zeitraums durchzuführen. Für die Vermessung von Seen stehen im Allgemeinen verschiedene technische Lösungen zur Verfügung. So hat sich in der Vergangenheit die Sonartechnik sehr stark weiterentwickelt und moderne Echolote erlauben es, großflächige Bereiche in relativ kurzer Zeit zu erfassen. Aber auch andere Fernerkundungsmethoden wie das Airborne Laserscanning mit grünen Lasern und die Mehrmedienphotogrammetrie bieten diesen Vorteil und ermöglichen die Vermessung von sonst schwer zugänglichen Gewässern. Ziel der vorgestellten Untersuchung ist es, die Möglichkeit zur Anwendung und Kombination dieser Methoden zu überprüfen sowie die jeweiligen Ergebnisse miteinander zu vergleichen. Spezielles Augenmerk wird dabei auf die Anwendung im hochalpinen Bereich, zur Vermessung von Gebirgsseen, gelegt. Die größten Herausforderungen, die sich dabei ergeben sind die schwierige Erreichbarkeit der Seen (u.U. Transport der Messausrüstung zu Fuß) und die klimatischen Bedingungen (kurze Zeitspannen mit günstigen Messbedingungen, starke Witterungsabhängigkeit und Temperaturschwankungen in den Bergen).

¹ Universität Innsbruck, Institut für Infrastruktur, AB Wasserbau, Technikerstr. 13, A-6020 Innsbruck, E-Mail: daniel.innerhofer@uibk.ac.at

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

³ TU Wien, Department für Geodäsie und Geoinformation, Gußhausstraße 27-29/E120 A-1040 Wien

⁴ AirborneHydroMapping GmbH, Feldstraße 1b, A-6020 Innsbruck, E-Mail: info@ahm.co.at

2 Datenerfassung

In diesem Abschnitt wird das Untersuchungsgebiet vorgestellt und die durchgeführte Messkampagne (Echolot, Messflug) beschrieben. Die Datenerfassung erfolgte einerseits mittels Echolot (Abschnitt 2.3) und andererseits mittels Laserscanning und Luftbildauswertung (Abschnitt 2.4). Für letztere wurden photogrammetrische Passpunkte signalisiert und eingemessen (Abschnitt 2.2).

2.1 Untersuchungsgebiet

Das Untersuchungsgebiet wurde unter dem Gesichtspunkt ausgewählt, dass es sich um ein wissenschaftliches Projekt handelt und somit möglichst aussagekräftige und übertragbare Ergebnisse geliefert werden sollen. Dazu wurden gewisse Anforderungen an die zu untersuchenden Bergseen gestellt: der See soll eindeutig alpinen Charakter haben, er soll trotzdem gut erreichbar sein und er muss für eine Befliegung geeignet sein (Sicherheitsaspekte beim Manövrieren im Gebirge). Des Weiteren sollte er nicht im direkten Einzugsgebiet eines Gletschers liegen, um einen fortlaufenden Eintrag von Schwebstoffen auszuschließen. Diese Kriterien werden von den Seen im Bereich der Sulzenau im Tiroler Stubaital erfüllt: dem Grünausee und der Blauen Lacke. Beide Seen liegen in wenigen Kilometern Entfernung zu einander auf rund 2.300 m Seehöhe, oberhalb der Sulzenauhütte, nahe des Alpenhauptkamms.



Abb. 1: Lage des Untersuchungsgebiets in den Tiroler Alpen.



Abb. 2: Links der Grünausee mit Blick nach Süden und rechts die Blaue Lacke von Osten aus gesehen.

Bei beiden Seen handelt es sich um Moränenstauseen, die durch Ufermoränen vom Haupttal abgetrennt werden und somit außerhalb der derzeitigen Gletschervorfelder liegen (METZ 1977). Durch ihre räumliche Nähe war es möglich, mit relativ geringem Mehraufwand, zwei Seen mit unterschiedlicher Charakteristik zu vermessen. Während die Blaue Lacke oft noch eine deutlich sichtbare Trübung aufweist, hat der Grünausee meist klares Wasser (PATZEL 2008). Bei längerfristig günstigen Witterungsbedingungen, wie dies z.B. 2016 der Fall war, besteht auch für die Blaue Lacke Bodensicht bis zur maximalen Tiefe von ca. 12 m.

Abb. 3 gibt einen Überblick über das Untersuchungsgebietes und die vor Ort durchgeführten Vermessungsarbeiten.



Abb. 3: Messkampagne Sulzenau: (a) Flugplanung und Passpunktanordnung, (b) Messboot im Einsatz (Grünausee), (c) Messboot vom Passpunkt 206 aus im Einsatz (Blaue Lacke), (d) Sichttiefenmessung Grünausee mit Secchi Scheibe, (e) radiometrische Kontrollflächen am Ufer des Grünausees, (f) Passpunktmessung beim Günausee, (g) Langzeitpasspunktbeobachtung Grünausee.

2.2 Passpunktmessung

Die Georeferenzierung der Luftbilder erfolgte anhand von photogrammetrischen Passpunkten. Abgestimmt auf die Größe der Bildpixel am Boden, die je nach Flughöhe 8-20 cm betrug, wurden zur Signalisierung quadratische weiße Zieltafeln mit einer Kantenlänge von 30 cm verwendet. Für jeden der beiden Seen wurden im Uferbereich jeweils zehn Passpunkte mittels in den Boden eingeschlagenen Steckeisen vermarkt. An diese wurde eine Gewindestange montiert und die Zieltafeln daran über eine zentrische Bohrung befestigt.

Die koordinative Einmessung erfolgte mit zwei Leica RTK GNSS Empfängern. Eine Antenne fungierte dabei als Basisstation, welche jeweils über einen gesamten Messtag betrieben wurde. Diese Langzeitbeobachtungen ermöglichte im Zusammenhang mit den Daten der naheliegenden
EPOSA Referenzstation Matrei in Tirol eine zentimetergenaue Koordinierung vor Ort. Die weiteren Passpunkte wurden mit einem Roverempfänger eingemessen, wobei die Beobachtungsdauer auch in diesem Fall zumindest 10-15 Minuten betrug. Die Auswertung der Koordinaten erfolgte einerseits in der Leica Software sowie mit der Open Source Programm RTKLIB⁵. Probleme bei der Datenauswertung ergaben sich vor allem durch Mehrwegeeffekte im Hochgebirge bzw. wegen der Abschattung durch die umliegenden Bergketten vor allem in südlicher Richtung. Trotz verhältnismäßig langer Beobachtungszeiten betrug die 3D-Puntklagegenauigket der photogrammetrischen Passpunkte lediglich 1-2 cm. Im Vergleich zur Bodenpixelgröße (8-20 cm) bzw. dem Laserfootprint (60 cm) ist dies aber dennoch eine Größenordnung genauer und daher als ausreichend zu betrachten.

2.3 Echolotaufnahme

Die Echolotung erfolgte am 11. + 12. Juli 2017 mit dem System GeoSwath Plus Compact der Firma Kongsberg. Dabei handelt es sich um ein hochauflösendes Fächerecholot, welches speziell für Flachwasseranwendungen konzipiert ist. Bei einem Fächerecholot wird aus dem empfangenen Echo des Schallimpulses (Ping) nicht nur die Entfernung sondern auch die Richtung bestimmt. Dadurch können mit einem Ping Punkte über die gesamt Fächerbreite gemessen werden. Das verwendete Gerät erreicht dabei eine Auflösung von bis zu 40 Punkten pro Meter Fächerweite. Die Vermessung des Sees erfolgt streifenweise, wobei die Messrichtung orthogonal zur Fahrtrichtung ist. So können mit parallelen Fahrwegen überlappenden Streifen aufgemessen werden.

max. Reichweite unter Wandler	max. Fächerweite	max. Überdeckung am Seeboden	Tiefenauflösung	Öffnungswinkel (Horizontal)	Transmit Pulselänge	Ping Rate
50 m	190 m	bis zu 12x facher Wassertiefe	1,5 mm	0,5 °	32 bis 224 µs	bis 30 Hz

Tab. 1: Gerätespezifikationen GeoSwath Plus Compact mit 500 kHz-Wandler (KONGSBERG 2018)

Das Messsystem bietet den Vorteil, dass es bei einem Gesamtgewicht von ca. 50 kg in mehrere Einheiten zerlegt werden kann. Als Trägersystem für das Echolot wurde ein Schlauchboot-Katamaran verwendet. Somit war es möglich, sowohl die Messausrüstung als auch das Boot soweit zu zerlegen und verpacken, dass die Teile jeweils von einer einzelnen Person, mit Hilfe einer Lastenkraxe, auch durch schwieriges Gelände transportiert werden konnten. Die Georeferenzierung während der Echolotung erfolgt über ein RTK-GNSS-System, wobei die benötigeten Korrekturdaten normaleweise über eine Ntrip-Verbindung von einem spezialisierten Datendienst abgerufen werden. Da jedoch im alpinen Gelände kein ausreichender Mobil-Funk-Empfang vorhanden war, musste für Zeit der Vermessung eine lokale Referenzstation installiert werden. Dazu wurde eine Trimble R8s Basistation in Sicht- und Funkreichweite der Seen über einem Fixpunkt aufgestellt und über Funk mit einem R10 Rover verbunden. Dieser konnte an die Prozessoreinheit des Echolots gekoppelt werden und ermöglichte somit den störungsfreien RTK-Betrieb. Dabei werden alle Daten des Wandlers und der Peripheriesysteme (GPS, Kompass, MRU, SVS) zeitsynchronisiert über die zentrale Recheneinheit erfasst und lokal gespeichert. Wichtig ist

⁵ RTKLIB: An Open Source Program Package for GNSS Positioning, http://www.rtklib.com/

ebenso die Erfassung der Temperatuverteilung, respektive der Schallausbreitungsgeschwindigkeit, über die Tiefe. Dazu wird mittels einer SVP-Sonde ein Schallgeschwindigkeitsprofil der Seen zum Messzeitpunkt aufgenommen.

Ein Punkt, der bei der bathymetrischen Vermessung mit dem Echolot zum Tragen kommt, ist der Zeitbedarf einer solchen Vermessung. Die Messung selbst kann bei einem See dieser Größe (Grünausee ca. 50.000 m²) innerhalb von zwei Stunden erfolgen. Nimmt man jedoch die benötigte Zeit für Auf- und Abbau der Messtechnik, Verladen und Transport hinzu, ergibt sich ein Zeitbedarf von bis zu eineinhalb Tagen pro See. Da es sich insgesamt um 350kg Ausrüstung handelt, die zu Fuß durchs Gelände transportiert werden müssen, sind auch entsprechende Anforderungen an das Personal gestellt.



Abb. 4: Echolotung des Grünausees.

2.4 Befliegung

Zur Vermessung des Seebodens mit den Mitteln der aktiven und passiven Fernerkundung wurde am 22.08.2017 eine Befliegung mit einem topo-bathymetrischen Laserscanner (Riegl VQ-880-G) sowie einer RGB Luftbildkamera (IGI H39) realisiert. Die Datenaufnahme wurde von der Firma AHM durchgeführt und war ursprünglich zeitgleich mit der Echolotkampagne geplant. Witterungsbedingt musste diese allerdings um fünf Wochen verschoben werden. Mehrere Starkregenereignisse Ende Juli und Anfang August 2017 führten allerdings zu einer verstärkten Trübung beider Seen. Während die Secchi Tiefe zum Zeitpunkt der Echolotkampagne beim Grünausee noch mehr als 10 m betrug und damit eine Gesamterfassung dieses Sees mittels Laserbathymetrie erlaubt hätte, wurde am Tag der Befliegung für beide Seen eine Secchi-Tiefe von ca. 2.5 m gemessen. Da jegliche optische Messmethode zur Bestimmung von Wassertiefe wesentlich durch die Gewässertrübung limitiert ist, konnte im konkreten Fall lediglich der Flachwasserbereich bis zu einer Tiefe von ca. 3.5 m erfasst werden. Dies entspricht der 1.5-fachen Secchi Tiefe und damit genau der Tiefenmessperformance des eingesetzten Lasersystems (RIEGL 2018). Die prinzipielle Flugplanung ist in Abb. 3a dargestellt und erfolgte für beide Seen in getrennten Fluglinien aus dem jeweiligen Seitental, welche in Y-Form im Bereich des Sulzenau-Talbodens zusammen trafen. Die Erfassung stellte fliegerische Herausforderungen aufgrund der hochalpinen Lage, dem steilen Abfall des Geländes talauswärts und der komplexen Thermik dar. Daher erfolgte die Befliegung für jeden See in mehreren Fluglinien unterschiedlicher Höhe, beginnend bei einer Flughöhe von etwa 1000 m über Grund bis zu der Zielflughöhe von etwa 500 m über Grund. Für jeden See wurden derart drei Flugstreifen im sogenannten "terrain following mode", d.h. jeweils im Sinkflug talauswärts. Die Fluggeschwindigkeit betrug etwa 100 Knoten (ca. 50 m/s). Zur Stützung des Flugblocks wurden zwei Fluglinien quer über beide Seen aus entsprechend großer Flughöhe durchgeführt.

Der eingesetzte topo-bathymetrische Laserscanner weist eine Scanrate von 550 kHz auf. Die Strahldivergenz beträgt 1.1 mrad, was je nach Flughöhe zu einem Durchmesser des Laserabtastflecks am Boden zwischen 0.55-1.10 m entspricht. Obwohl die Punktdichte mit etwa 20 Punkten/m² im Einzelstreifen auch die Ableitung von Geländemodellen mit einer Rasterweite von etwa 25 cm ermöglicht hätte, ist die effektive räumliche Auflösung durch die Größe des Abtastflecks limitiert. Der Scanner zeichnet für jeden Laserpuls die gesamte Echowellenform auf. Diese wird einerseits für offline Analyse im post processing abgespeichert und andererseits online ausgewertet (PFENNIGBAUER et al. 2014).

Zeitgleich zu den Laserdaten wurden Luftbilder mit einer 39 MPix RGB Kamera aufgezeichnet. Der Sensor hat 7216 × 5412 Pixel und die Pixelgröße am Sensor beträgt 6.8 µm. Beim eingesetzten 35 mm Objektiv ergibt sich damit eine Pixelgröße am Boden (Ground Sampling Distance, GSD) von 10-20 cm. Das RGB Bildmaterial wurde für die Tiefenbestimmung durch Mehrmedienphotogrammetrie (MAAS 2015; MANDLBURGER 2018) und Analyse der Spektralinformation (LYZENGA et al. 2006) benutzt. Die Bilder einer ebenfalls auf der Plattform befindlichen Multispektralkamera mit Infrarotkanal wurden in dieser Untersuchung nicht verwendet.

3 Methoden

3.1 Echolot

Die Datenerfassung und –aufzeichnung erfolgt mit der von Kongsberg gelieferten Software wie in Kap. 2.3 beschrieben. Die Datenbearbeitung und –export erfolgen mit demselben Softwarepaket. Die Daten der einzelnen Peripheriesysteme werden von der Software im Zuge der Rohdatenprozessierung automatisch verschränkt. Dabei werden die aufgezeichneten Echos entsprechend der Temperaturverteilung und Schallausbreitung im Wasser korrigiert und mithilfe der GPS-Koordinaten und Lagedaten des Messbootes georeferenziert. In einem zweiten Schritt werden Filter aufgebracht um störende Echos aus der Wassersäule (z.B. Gasblasen) zu entfernen. Dazu wird sowohl der Bereich eingegrenzt aus dem Echos vom Boden erwartet werden sowie die Amplitude des Signals als Kriterium herangezogen. Die so erhaltenen Punktwolken können anschließend interpoliert und in ein Polygonnetz umgewandelt und exportiert werden.

Ein weiterer Vorteil des verwendeten Echolots liegt darin, dass auch die Amplitudenstärken des Echosignals ausgewertet werden können, entsprechend einer klassischen Sidescan-Auswertung. Durch entsprechende Schattierung der Ergebnisse erhält man somit einen guten Eindruck über die Morphologie der Gewässersohle (Kap. 4.1, Abbildung 6).

3.2 Laserbathymetrie

Die Laserdaten wurden einerseits mit der Herstellersoftware RiProcess und darüber hinaus mit der wissenschaftlichen Laserscanning Software OPALS (PFEIFER et al, 2014) prozessiert. Nach der initialen direkten Georeferenzierung in RiProcess wurde für den gesamten Flugblock eine Streifenausgleichung mit dynamischer Trajektorienkorrektur (GLIRA et al., 2016) durchgeführt. Die Standardabweichung der Höhendifferenzen gemessen in glatten Bereichen überlappender Flugstreifen betrug dabei nach der Streifenausgleichung 3.5 cm. Während bei mäßig bewegtem Gelände und einfacheren Flugbedingungen mit demselben Scanner in anderen Projekten auch bereits bessere Passgenauigkeiten von ca. 2 cm erreicht wurden, sind die Abweichungen angesichts des komplexen Geländes und der schwierigen Flugbedingungen mit der damit verbundenen verminderten Genauigkeit der Flugtrajektorie als zufriedenstellend zu betrachten.

Für die erforderliche Refraktions- und Laufzeitkorrektur wurde zunächst für jeden See getrennt ein Wasseroberflächenmodell nach dem statistischen Ansatz von MANDLBURGER et al. (2013) berechnet. Jeder Laserpuls im Wasserbereich wurde anschließend mit diesem Modell verschnitten und der wasserseitige Strahlverlauf entsprechend dem Brechungsgesetz von Snellius korrigiert (MAAS, 2015, MANDLBURGER, 2018a). Die korrigierten 3D-Punkte wurden anschließend mittels hierarchischer robuster Interpolation in Boden- und Nichtbodenpunkte klassifiziert und aus ersteren schließlich das DGM des Gewässerbodens sowie des umgebenden Geländes berechnet.

3.3 Bild-basierte Bathymetriebestimmung

Neben den Laserdaten wurden die RGB Bilder zur Tiefenbestimmung herangezogen. Dazu wurde der gesamte Bildblock in Agisoft Photoscan⁶ und Pix4D Mapper⁷ orientiert. Die absolute Georeferenzierung erfolgte dabei auf Basis der 22 mittels GNSS gemessenen Passpunkte. Der Rückprojektionsfehler belief sich dabei auf ca. 0.15 Pixel und die mittleren Abweichungen an den Passpunkten (RMS) betrugen 2 cm in der Lage und 4 cm in der Höhe.

Im Anschluss an die Aerotriangulation wurden mit dem Programmsystem SURE⁸ mittels Dense Image Matching (DIM) eine dichte Punktwolke sowie ein DSM abgeleitet und darauf aufbauend ein True Orthophoto der beiden Seen berechnet. Da eine korrekte Bildzuordnung vor allem für tiefe Seebereiche nicht möglich war, wurde zur Rektifizierung der Luftbilder die Wasseroberfläche aus der Laservermessung in das DIM DEM integriert. Für den Flachwasserbereich hingegen ist eine bildbasierte Tiefenbestimmung möglich. Hierbei kamen zwei verschiedene Ansätze zur Anwendung: (i) Mehrmedienphotogrammetrie und (ii) Tiefenbestimmung aus spektraler Information. Beide Ansätze sind in MANDLBURGER (2018b) detailliert beschrieben. Die erforderlichen Arbeitsschritte werden daher im Folgenden nur grob skizziert.

Bei der Mehrmedienphotogrammtrie ist beim Vorwärtsschritt der Punkte des Gewässerbettes die Brechung an der Wasseroberfläche gemäß dem Snellius'schen Brechungsgesetzt zu berücksichtigen. Als Input werden (i) die unkorrigierten 3D-Objektpunkten, (ii) ein Wasseroberflächenmodell, (iii) der Brechungsindex von Wasser (~1.33) und (iv) für jeden Punkt die zugehörigen Bildpositionen, aus denen der Punkt vorwärtsgeschnitten wurde, benötigt. Mit diesen Informationen kann

⁶ Webseite Agisoft: <u>http://www.agisoft.com/</u>

⁷ Webseite Pix4D Mapper: https://www.pix4d.com/product/pix4dmapper-photogrammetry-software

⁸ Webseite nFrames: https://www.nframes.com/

für jeden Punkt das scheinbare Strahlenbündel rekonstruiert werden. Jeder Bildstrahl wird dann mit der Wasseroberfläche verschnitten (ray tracing) und die Richtung des, aufgrund der Strahlbrechung beim Übergang in das optisch dichtere Medium Wasser, zum Lot hin abgelenkten Bildstrahls bestimmt. Alle wasserseitigen Bildstrahlen werden dann erneut vorwärtsgeschnitten, wobei bei mehr als 3 Bildstrahlen die 3D Position des Gewässerbodenpunktes mittels vermittelnder Ausgleichung bestimmt wird. Es sei an dieser Stelle erwähnt, dass die Textur durch Streueffekte und Signalabsorption im Wasser mit zunehmender Tiefe je nach Trübungsgrad abnimmt. Durch die beträchtliche Trübung zum Zeitpunkt der Luftbildaufnahme konnten im gegenständlichen Fall Wassertiefen nur bis zu einer Tiefe von ca. 2 m zuverlässig bestimmt werden. Die homogene Textur der bewuchsfreien hochalpinen Seen verringerte die erzielbare Eindringtiefe zusätzlich.

Alternativ können Gewässertiefen auch aus der multispektralen Bildinformation gewonnen werden (LYZENGA et al. 2006). Dabei wird für jedes Orthophotopixel eine radiometrische Maßzahl ermittelt, wobei zumeist ein logarithmiertes Verhältnis zweier Bänder zur Anwendung kommt. Anhand von Referenzdaten, die punktuell aus den Fächerecholotdaten extrahiert wurden, wird anschließend der Zusammenhang zwischen Tiefe und radiometrischer Maßzahl gebildet. Im konkreten Fall erwies sich ln(grün/rot) als optimale Maßzahl für die lineare und quadratische Regression. Die ermittelten Parameter der Regressionsfunktion wurden anschließend auf alle benetzten Orthophotopixel angewandt und somit ein flächenhaftes Rastermodell des Gewässerbodens mit einer Gitterweite von 1xGSD (10 cm) berechnet. Ergebnisse

3.4 Echolotung

Die Echolotung liefert eine sehr detaillierte Bathymetrie der Gewässersohle. Durch die erwähnte, hohe Punktdichte (bis zu 400 Punkte pro m²) sind auch kleinräumige Strukturen gut erkennbar.



Abb. 5: Blaue Lacke: über die Tiefe farbkodiertes DGM, erstellt aus der Echolotdaten mit einer Rasterauflösung von 50 × 50 cm. Die max. Tiefe lag bei ca. 10 m.



Abb. 6: Hochaufgelöstes Sidescan-Sonar Bild des südlichen Uferbereichs der Blauen Lacke. Die Auswertung der Sidescan-Daten liefert eine sehr detaillierte Morphologie. Starke Reflektoren (z.B. Steine) sind hell dargestellt. Pixelgröße 10 × 10 cm.



Abb. 7: Grünausee: über die Tiefe farbkodiertes DGM, erstellt aus der Echolotdaten mit einer Rasterauflösung von 50 × 50 cm. Die max. Tiefe lag bei ca. 12 m.

Die Begutachtung der einzelnen Streifen zeigt jedoch, dass es noch gewisse Abweichungen zwischen den Streifen gibt Abb. 7. Diese Differenz ist in den Randbereichen größer, beispielhaft hier dargestellt für zwei Streifen ergibt sich dort eine Abweichung bis zu 16 cm in der Höhe in den mittleren Bereichen gibt es Abweichungen von rund 4 cm. Diese Differenz kann mehrerlei Ursachen haben, z.B. schnelle Bewegungen des Messboots (Kurvenfahrten, Wellenschlag), Abweichungen in der Referenzierung und Ausrichtung der einzelnen Sensoren zueinander, Übertragen von Ungenauigkeiten der Positionsbestimmung mittels GNSS oder wie hier am wahrscheinlichsten eine nicht vollständige Korrektur der veränderten Schallausbreitung über die Wassertiefe. Betrachtet man die Gesamtheit der Punkte der einzelnen Streifen ergibt sich eine Standardabweichung für den Lagefehler von 11 cm. Dieser Wert reduziert sich für die gesamte Punktwolke dadurch weiter, dass die Randbereiche von mehreren Streifen abgedeckt werden und so über eine Vielzahl von Punkten interpoliert wird. Für das dargestellte Beispiel der Blauen Lacke (ca. 24.000 m²) ergeben sich für die bereinigte Punktwolke ca. 5.000.000 Bodenpunkte mit einer Standardabweichung von ca. 9 cm. Durch Feintuning der Kalibrierwerte (z.B. für die Winkelabweichung des Messsystems) ist eine weitere Verringerung dieser Abweichung möglich.



Abb. 8: Streifenabgleich der Echolotdaten der Blauen Lacke, mit Farbkodierung der Höhenabweichung [m]

3.5 Laserbathymetrie

In den folgenden Abschnitten werden die Ergebnisse der Tiefenbestimmung mittels optischer Fernerkundung präsentiert. Wegen des hohen Trübungsgrades beider Seen konnten dabei nur die Flachwasserbereiche erfasst werden. Exemplarisch werden die Ergebnisse daher anhand des südwestlichen Endes der Blauen Lacke dargestellt, welches einen flachen Uferbereich aufweist und damit eine entsprechend große Fläche mit Wassertiefen bis maximal 3.5 m.

Abb. 9 zeigt im benetzten Bereich die aus den Laserdaten ermittelte Gewässertiefenkarte soweit eine Durchdringung bis zum Grund möglich war und das digitale Orthophoto im restlichen Seebereich. An Land ist die Schummerung des Laser DGMs dargestellt. Man erkennt deutlich, dass der Gewässerboden bis zur maximalen Penetrationstiefe konsistent dargestellt ist und morphologische Details (Rinnen, Blöcke, etc.) in den Unterwasser-Laserdaten abgebildet sind. Ein quantitativer Vergleich zu den Echolotdaten als Referenz findet sich in Abschnitt 3.7.



Abb. 9: Blaue Lacke: Seebereich: Wassertiefenkarte in [m] ermittelt aus 3D Laserpunktwolke und digitales Orthophoto in Bereichen außerhalb der maximalen Penetrationstiefe; Uferbereich: Schummerung Laser DGM

3.6 Bathymetrie aus Bilddaten

Abbildung 10 zeigt das digitale Orthophoto, welches als Nebenprodukt des Dense Image Matching entstanden ist. Das Orthophoto veranschaulicht die relativ hohe Gewässertrübung der Blauen Lacke zum Zeitpunkt der Datenerfassung. Gewässerbodenstrukturen, welche die Grundvoraussetzung für photogrammetrische Punktbestimmung ist, sind lediglich im ufernahmen Flachwasserbereich zu erkennen. Dem entsprechend eingeschränkt ist die erzielbare Messtiefe, die bei den gegebenen Umständen wesentlich geringer ist als jene der Laserbathymetrie. Im südöstlichen Bildbereich sind darüber hinaus auch Schlaglichter (sun glint) zu erkennen. Zur Vermeidung von Schlaglichtern wäre eine Befliegung bei tieferem Sonnenstand (z.B. in den Morgen- oder Abendstunden) vorteilhafter gewesen. Wegen der umgebenden steilen Berghänge hätte dies aber Schatten der Felswände im See zur Folge gehabt, weswegen die Schlaglichter in Kauf genommen wurden. Abb. 11 zeigt eine Perspektivansicht der mittels Mehrmedienphotogrammetrie ermittelten 3D Punktwolke. Die Profilansicht der Abb. 11 unten ist zu entnehmen, dass der Gewässerbereich nur in der Flachwasserzone bis zu einer Tiefe von ca. 1-1.5 m zuverlässig erfasst ist. Darüber hinaus nehmen zwischen 1.5-2.5 m Wassertiefe zunächst Ausreißerpunkte mit grob falschen Höhen zu. Bei noch größerer Tiefe können mangels Textur keine Punkte mehr bestimmt werden. Dies verdeutlicht klar die Einschränkungen des Verfahrens, welches neben glatter Wasseroberfläche und klarem Wasser auch genügend Bodentextur erfordert. Verfahrensbedingt ist die maximale Eindringtiefe bei der Mehrmedienphotogrammetrie kleiner der einfachen Secchi-Tiefe.



Abb. 10: Blaue Lacke: Digitales True Orthophoto



Abb.11: Blaue Lacke: Perspektivansicht der 3D Punktwolke aus Mehrmedienphotogrammetrie (oben) und Profilschnitt (unten), Lage des Profilschnitts in der Grundrissansicht (links oben).

In Abb. 12 sind der Ablauf der radiometrischen Tiefenbestimmung (a+b), die resultierende Tiefenkarte (c) und der Vergleich mit den Fächerecholotdaten als Referenz (d) dargestellt. Abb. 12a zeigt die ausgewählten Echolotprofillinien farbkodiert über dem digitalen RGB Orthophoto. Anhand dieser Referenzinformation wurde die Regressionsfunktion gegenüber der radiometrischen Größe x=log(rot/grün) durchgeführt, deren Ergebnis Abb. 12b zeigt. Bemerkenswerterweise hat sich der rote Kanal als sehr informativ herausgestellt, obwohl in diesem Wellenlängenbereich bereits eine starke Absorption in der Wassersäule gegeben ist. Der logarithmierte Quotient mit dem Grünkanal, welcher eine relativ geringe Absorption in der Wassersäule erfährt, führt zu einem starken funktionalen Zusammenhang zwischen Radiometrie und Tiefe. Im Vergleich zum Laserbodenmodell (Abb. 9) sind in der resultierenden Tiefenkarte der Abb. 12c ähnliche Strukturen zu erkennen. Die relativ gute Übereinstimmung mit den im konkreten Fall auch flächenhaft vorliegenden Fächerecholotdaten bestätigt das Histogramm der Resthöhenfehler der Abb. 12d. Die mittlere Abweichung liegt bei 2 mm und die robust abgeschätzte Standardabweichung (GMAD) beträgt 22 cm, was umgerechnet auf die Wassertiefe einen relativen Fehler von etwa 5% bedeutet. Dies stellt für radiometrische Tiefenbestimmung einen sehr guten Wert dar. Die Eindringtiefe ist aber auch hier mit etwa 4 m begrenzt, was aber immerhin größer als die Sichttiefe ist.



Abb. 12: Tiefenbestimmung aus multispektralen Bildern; (a) Digitales RGB Orthophoto und Referenzdaten, (b) Regression: Tiefe zu radiometrischer Messgröße, (c) Tiefenkarte ermittelt aus Bildradiometrie, (d) Histogramm der Höhenabweichung zwischen Fächerecholot und radiometrisch bestimmter Tiefe.

3.7 Vergleich der Methoden

Während die Gewässertiefen in Abb. 12c verfahrensbedingt gut zu den Echolotdaten passen, da diese im Zuge der Kalibrierung als Referenz eingefügt wurden, stellt der in Abb. 13 dargestellte Soll-Ist-Vergleich zwischen den Tiefen aus der Sonar- bzw. Laservermessung einen unabhängigen Vergleich dar. Abb. 13a zeigt dabei die farbkodierten Höhendifferenzen der beiden Verfahren in einer Skala von -25 cm (rot) bis +25 cm (blau) in 5 cm Schritten. Große Teile des Überlappungsbereichs weisen dabei einen weißlichen Farbton auf, welcher geringe Differenzen kleiner 5 cm kennzeichnet. Die größten Differenzen treten dabei im Uferbereich auf, wo die Echolotvermessung Tiefen lediglich von schleifende Echos in den flachen Randbereich hinein liefert. Diese sind verfahrensbedingt weniger zuverlässig als die tieferen Bereiche in der Seemitte. Laserbathymetrie hat hier bei der Verwendung von Scannern mit sehr kurzer Pulslänge den Vorteil eines praktisch nahtlosen Übergangs zwischen Wasser und Land. Die quantitative Auswertung ist im Histogramm der Abb. 13b zusammengefasst. Die Verteilung der Höhenabweichungen ist annähernd symmetrisch um Null (Median: +3 mm, Mittelwert: -4 mm) und die Standardabweichung beträgt 9.3 cm. Die robuste Standardabweichung (σ_{MAD}) ist mit 6 cm noch günstiger und belegt eine gute generelle Übereinstimmung der unabhängig erfassten Tiefen im Überlappungsbereich.



Abb. 13: Vergleich Echolot vs. Laserbathymetrie, (a) Farbkodierte Höhendifferenzen, (b) Histogramm der Soll-Ist-Höhenabweichungen.

4 Diskussion

Die Vermessung mittels Echolot lieferte sehr gute Ergebnisse und ein detailliertes Bild der Gewässersohlen und ihrer Morphologie. Gleichzeitig war es aber die Methode, die vor Ort den größten Personal- und Zeitaufwand erforderte, da die gesamte Ausrüstung teils weite Strecken zu Fuß getragen werden musste. Der mangelnde Mobil-Funk-Empfang stellt ein zusätzliches Problem bei Verwendung eines RTK-Systems dar und macht die Verwendung von lokalen Festpunkten notwendig. Durch die flugzeugbasierte Datenerfassung konnte ein lückenloser Datensatz im Uferund Flachwasserbereich erstellt werden. Das ursprüngliche Ziel einer großflächigen Erfassung mit den Mitteln der aktiven und passiven optischen Fernerkundung konnte aber aufgrund der äußeren Bedingungen (Gewässertrübheit) nicht erzielt werden. Als Erkenntnis bleibt die Feststellung, dass die Vermessung von hoch alpinen Seen mit der klassischen Echolotvermessung zwar aufwändig, aber weniger anfällig hinsichtlich der äußeren Bedingungen ist. Laserbathymetrie und Mehrmedienphotogrammetrie stellen dagegen für einen großflächigen Einsatz höhere Anforderungen an die äußeren Bedingungen wie Flugwetter und Transparenz des Wasserkörpers.

5 Danksagung

Der Beitrag von Gottfried Mandlburger wurde durch Mittel der Deutschen Forschungsgesellschaft (DFG) im Rahmen des Forschungsprojektes "Bathymetrievermessung durch Fusion von Flugzeuglaserscanning und multispektralen Luftbildern" unterstützt.

6 Literaturverzeichnis

- GLIRA, P., PFEIFER, N. & MANDLBURGER, G., 2016: Rigorous Strip Adjustment of UAV-based Laserscanning Data Including Time-Dependent Correction of Trajectory Errors. Photogrammetric Engineering & Remote Sensing, 82(12), 945-954.
- KONGSBERG, 2018: Datenblatt GeoSwath Plus. <u>www.km.kongsberg.com/geoacoustics</u>, letzter Zugriff: 20.12.2018.
- LYZENGA, DAVID R., MALINAS, NORMAN P. & TANIS, FRED J., 2006: Multispectral Bathymetry Using a Simple Physically Based Algorithm. IEEE Transactions on Geoscience and Remote Sensing, **44**(8), 2251-59.
- MAAS, H.-G., 2015: On the Accuracy Potential in Underwater/ Multimedia Photogrammetry. Sensors, **15**(8), 18140-18152.
- MANDLBURGER, G., 2018a: A Case Study on Through-Water Dense Image Matching. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **42**(2), 659-666.
- MANDLBURGER, G., 2018b: Gewässervermessung mittels aktiver und passiver optischer Fernerkundung. Vermessung & Geoinformation, **3**, 195-206.
- METZ, B., 1977: Geomorphologische Untersuchungen zur Unterscheidung zwischen Eisrand- und Moränenstauseen. Berichte der Naturforschenden Gesellschaft zu Freiburg im Breisgau, 67, 203-215.
- PATZELT, G., 2008: Sammelbericht über die Gletschermessungen des Oesterreichischen Alpenvereins im Jahr 2007 (Gletscherbericht 2006/2007). Bergauf, 02/2008, 26-32.
- 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.
- PFENNIGBAUER, M., WOLF, C., WEINKOPF, J., & ULLRICH, A. (2014). Online waveform processing for demanding target situations. In: Proceedings SPIE Proceedings 9080, 1-10, https://doi.org/10.1117/12.2052994.
- RIEGL, 2018: Datenblatt VQ-880-G. <u>http://www.riegl.com/uploads/tx_pxpriegldownloads/Infos-heet_VQ-880-G_2016-05-23.pdf</u>, letzter Zugriff: 20.12.2018.

Encoder-Decoder network for local structure preserving stereo matching

JUNHUA KANG^{1, 2}, LIN CHEN², FEI DENG¹ & CHRISTIAN HEIPKE²

Abstract: After many years of research, stereo matching remains to be a challenging task in photogrammetry and computer vision. Recent work has shown great progress by formulating dense stereo matching as a pixel-wise learning task to be resolved with a deep convolutional neural network (CNN). In this paper we investigate a recently proposed end-to-end disparity learning network, DispNet (MAYER et al. 2015), and improve it to yield better results in some problematic areas. The improvements consist in two major contributions. First, in order to handle large disparities, we modify the correlation module to construct the matching cost volume with patch-based correlation. We also modify the basic encoder-decoder module to regress detailed disparity images with full resolution. Second, instead of using post-processing steps to impose smoothness and handle depth discontinuities, we incorporate disparity gradient information as a regularizer to preserve local structure details in large depth discontinuity areas. We evaluate our model in terms of end-point-error on several challenging stereo datasets such as Scene Flow, Sintel and KITTI. Experimental results demonstrate that our model achieves better performance than DispNet on most datasets (e.g. we obtain an improvement of 36% on Sintel) and estimates better structure-preserving disparity maps. Moreover, our proposal also achieves competitive performance compared to other methods.

1 Introduction

Stereo matching has continuously been an active research area in photogrammetry and computer vision. It is widely used in different applications, such as robotics and autonomous driving, 3D model reconstruction, object detection and recognition. The core task of stereo matching is to find pixel-wise correspondences between images, and thus to calculate the parallax (called disparity in computer vision) of corresponding pixels between images.

Recently, deep learning techniques have shown powerful capability for stereo matching. Convolutional neural networks (CNN) (LECUN et al. 1998) have first been introduced to calculate matching costs in Maching Cost CNN (MC-CNN) (ZBONTAR & LECUN 2016). Instead of using handcrafted matching cost metrics, the authors present a Siamese CNN for measuring the similarity between image patches. Most other recently suggested patch based stereo methods also focus on using CNN to generate unary terms as similarity measure (CHEN & YUAN 2016; LUO et al. 2016). Though patch based similarity measurements out-perform traditional hand-crafted ones, these algorithms require extra post-processing steps and hand-crafted regularization to produce complete disparity results. Therefore, some researchers suggested using an end-to-end network to directly estimate the disparity from stereo images. DispNet is first such end-to-end learning framework

¹ Wuhan University, School of Geodesy and Geomatics, 129 Luoyu Road, Wuhan, Hubei Province, P.R.China, 430079, E-Mail: jhkang.whu@gmail.com, fdeng@sgg.whu.edu.cn

² Leibniz Universität Hannover, Institute of Photogrammetry and GeoInformation, Nienburger Straße 1, D-30167 Hannover, E-Mail: [kang, chen, heipke]@ipi.uni-hannover.de

(MAYER et al. 2015), which was derived from FlowNet (DOSOVITSKIY et al. 2015). Both of them are restricted to rectified stereo images. The network architecture follows a coarse- to- fine fashion called auto encoder-decoder structure. It encodes the high-level semantic information at low resolution through successive convolutions and activations and then decodes the result back to the original resolution by successive deconvolutions. DispNet achieves considerable performance compared to traditional and patch based learning approaches in terms of both accuracy and speed. However, average error loss used in DispNet results in over-smoothing in output disparity, which leads to losing local structure details, especially in large disparity discontinuity areas. In addition, we find that DispNet has lower accuracy for large disparities.

In this paper, we use DispNet as the basic architecture and present a gradient regularizer for local structure preserving stereo matching. The horizontal and vertical gradients of the disparity map convey information about significant depth differences in the scene and local structure, which can be used to improve estimated disparity maps. In order to avoid over-smoothing in output disparity, especially around large disparity discontinuities, we add a gradient regularizer based on depth gradient information into our network to preserve sharp structure details. In addition, we modify the correlation layer in the cost volume construction module to deal with large disparities in large scale scenes. Finally, we also modify the structure of the encoder-decoder module to preserve more spatial information and output a full resolution disparity map.

The remainder of this paper is structured as follows: we review the related work of stereo matching based on CNNs in Section 2. Section 3 presents our methodology. Experimental results and analysis are illustrated in Section 4, followed by a set of conclusion in Section 5.

2 Related work

There is a lot of literature focusing on stereo matching research. A traditional pipeline for stereo matching includes four steps, which are matching cost computation, cost aggregation, disparity calculation and finally disparity refinement (SCHARSTEIN & SZELISKI 2002). Since current state-of-the-art studies focus on stereo matching employing deep learning techniques, we restrict our review to those CNN based methods. These approaches estimate disparities which can reflect part or all of the aforementioned four steps; they can be roughly divided into three categories: patch-based matching cost learning, post-processed regularity learning, and end-to-end disparity learning.

Patch-based matching cost learning. In this category, CNNs are introduced to compute the matching cost of image patches. MC-CNN (ZBONTAR & LECUN 2016) is a Siamese network composed of a series of stacked convolutional layers to extract descriptors of each image patch, followed by a decision module for measuring similarity. Luo et al. (Luo et al. 2016) expand on Zbontar's work and propose a notably faster Siamese network to learn a probability distribution over all possible disparities without manually pairing patch candidates. Chen and Yuan (CHEN & YUAN 2016) propose a multi-scale CNN to introduce global context by employing down-sampled images and increase the matching accuracy without enlarging the input patch. Although the patch based methods outperform most traditional stereo matching methods, which use hand-crafted features, they still require subsequent post-processing steps to produce complete results.

Post-processed regularity learning. This category learns regularization and focuses on the postprocessing of disparity maps. Scharstein and Pal (SCHARSTEIN & PAL 2007) earn the parameters of conditional random fields (CRFs) to replace heuristic priors on disparities. Li and Huttenlocher (LI & HUTTENLOCHER 2008) train a non-parametric CRF model with explicit occlusion labeling by using a structured support vector machine (CORTES & VAPNIK 1995). Guney and Geiger (GUNEY & GEIGER 2015) incorporate semantic segmentation and object recognition in a super-pixel based CRF framework to learn regularization and resolve ambiguities in reflective and textureless regions. Seki and Pollefeys (SEKI & POLLEFEYS 2017) propose a SGM-Net to learn penalty parameters for different 3D object structures. They obtain better penalties than hand-tuned SGM (HIRSCHMULLER 2008) and mitigate streaking artifacts that appear in MC-CNN.

End-to-end disparity learning. Approaches in this category incorporate matching cost computation and hand-crafted post-processing into a single learning process for joint optimization and train the whole network in an end-to-end mode. The first end-to-end stereo matching network is DispNet (MAYER et al. 2015), which has a similar structure that FlowNet (DOSOVITSKIY et al. 2015). DispNet is an encoder-decoder architecture for disparity regression. Given a pair of rectified images, it explicitly extracts features in the encoder part and then directly estimates the disparity map in the decoder part by minimizing a regression training loss based on the absolute difference between prediction and ground truth disparity. DispNet has achieved prominent performance and has become a baseline network in stereo matching. Several studies have tried to improve its performance by stacking multiple networks together based on this baseline architecture. For instance, CRL (PANG et al. 2017) is a cascade residual learning network, stacking an advanced DispNet and a residual network together for explicitly refining initial disparity. GC-Net (KENDALL et al. 2017) regresses disparity by employing 3D convolutional layers to exploit more context information. Similar to GC-Net, PSM-Net (SHAKED & WOLF 2017) uses spatial pyramid pooling and 3D convolutions to incorporate contextual information on different scales. However, high-dimensional feature based 3D convolution is computationally expensive. DenseMapNet (ATIENZA 2018) uses Dense Convolutional Networks (DenseNet) (HUANG et al. 2017) instead of the encoder-decoder structure to reduce the number of learning parameters. Although this network is fast to train, the results show limitations in terms of preserving structure details. To enforce smooth disparities, a disparity smoothness loss is introduced in an unsupervised deep neural network for single image depth estimation (GODARD et al. 2017). This smoothness loss is added with an edge-aware term using original image gradients. Inspired by this method, we apply a gradient regularizer on disparity estimation in a supervised way based on the gradients of prediction and ground truth disparity to preserve local structure details.

3 Methodology

In this study, we aim to improve performance in some problematic areas by adding some modification to DispNet. The input of our network is two rectified stereo images. First, we modify the correlation module of DispNet to deal with large disparities and modify the structure of the encoder-decoder part to obtain a disparity map with the same resolution as the input. Second, under the assumption that the desired disparity map should be locally smooth except at actual discontinuities, we present a gradient regularizer to preserve sharp structure details.

3.1 Network structure

The schematic structure of our proposed network is depicted in Fig. 1. This is a data-driven model that enables end-to-end disparity learning. From Fig. 1, it can be observed that our encoder-decoder network is composed of three main parts, namely feature extraction, cost volume construction, and disparity estimation.

Feature extraction. In this part we extract features by a Siamese structure with two sharedweights branches. In each branch, we employ two convolutional layers, conv_1, conv_2, to learn the unary features, which are followed by a rectified linear unit (ReLU), respectively. As shown in Fig. 1, the weights between the two branches of the feature extraction part are shared. It means that the network learns the same type of feature from the two input images. The output feature maps from this module are then applied to the correlation module to extract the correspondence prior between left and right images.

Cost Volume Construction. After having obtained deep unary features from the two Siamese branches, the cost volume can be constructed based on these features. Because the correlation is indeed an effective cue for finding conjugate pairs, we explicitly encode this relationship in our model, which enables our network to capture correspondences between the stereo pairs. As the input stereo images are rectified images, the y-coordinates of conjugate points are identical. Similar to DispNet, we can use a 1-D correlation layer along the x-direction (epipolar line) to construct the cost volume. First, let M_L , M_R denote the left and right feature maps with w, h, c representing their width, height and number of channels, respectively. Then, the cost volume C is created by convolving the left and right feature maps M_L , M_R up to the maximum disparity d_m . The correlation of two patches (i.e. context windows) centered at x_1 in M_L and x_2 in M_R is defined as $C(x_1, x_2) = \sum_{o \in [-k,k] \times [-k,k]} \langle M_L(x_1 + o), M_R(x_2 + o) \rangle$ (1)

where K = 2k + 1 is the patch size and $\langle \cdot \rangle$ means the convolution operation.. We restrict the search space of possible patch-pairs by setting the maximum displacement along the epipolar line. For each location x_1 in M_L , we compute the correlation $C(x_1, x_2)$ only in the interval $[x_2 = x_1, x_2 = x_1 + d_m]$, which implies a one sided search on M_R . We set d_m to 40 and increase the stride from 1 to 2 when computing the cost volume C by sliding M_L over M_R. In this way, our network can handle large correlation distances (40*4*2=320 pixel, note that from the first and second convolutional layers the feature map is downsampled by a factor of 4) without any extra computation and memory cost. After creating the resulting multi-channel maps and organizing the relative displacements in channels, we obtain a 3D cost volume of size (w × h × ($d_m + 1$)).

Encoder-Decoder module. Given the disparity cost volume, the next step is to learn a regularization function to refine our disparity estimation. We modify the deep encoder-decoder module of DispNet to output detailed disparity with the same resolution as the input. The architecture of our encoder-decoder network is presented in Fig. 1. The encoder part encodes sub-sampled features from the input and captures high-level representations by interleaving convolutional layers and pooling. It enables the network to explicitly leverage context with a wide field of view. However, it results in reduced resolution with multiple convolutions of stride 2. Therefore, unlike DispNet, which uses 4 groups of convolutional layers to downsample the features with a factor of 64, we only stack 3 groups of convolutional layers in the encoder to preserve more spatial context. Each group contains two 3×3 convolutions with strides of 2 and 1 respectively, achieving an encoded feature map with dimension ($W/32 \times H/32 \times C/32$) where W, H, C represent the width, height,

and channels, respectively. In order to obtain dense per-pixel predictions with the original input resolution, we apply 5 up-sampling blocks corresponding to six scales $(1/32, 1/16, 1/8, 1/4, 1/2, and 1 \times of$ the input size) in the decoder part to refine the coarse representation. Each block consists of a 4×4 deconvolution layer with stride of 2 to up-sample the encoded output. Similar to DispNet, skip connections are also used in the decoder part to preserve both the high-level coarse and the low-level fine information. In addition, we connect the left original image with deconvolution features, as shown in Fig. 1, to output more accurate and full resolution disparity maps, which is different from DispNet.



Fig. 1: Architecture overview of proposed method

3.2 Complementary loss

We train our model end to end with supervised learning using ground truth disparity data. In order to preserve local structure details in the output disparity, we present a gradient regularizer as an auxiliary loss. So, the loss for training contains two parts: the disparity regression term and the gradient regression term.

For the disparity regression loss \mathcal{L}_r , we use the end point error (EPE), the absolute Euclidean distance between the disparity D predicted by the model and the ground truth disparity \widehat{D} , averaged over the valid pixels. We adopt the ℓ_1 norm to regularize prediction which is widely used in previous methods. Thus, the disparity regression loss \mathcal{L}_r is formulated as:

$$\mathcal{L}_{r} = \frac{1}{N_{v}} \sum_{i,j \in v} \left\| D_{i,j} - \widehat{D}_{i,j} \right\|_{1}$$
⁽²⁾

where $\|\cdot\|_1$ denotes the ℓ_1 norm, v represents all valid disparity pixels in \widehat{D} and N_v is the number of valid pixels. As ground truth disparity maps are sometimes sparse (e.g. KITTI dataset (GEIGER et al. 2012; MENZE et al. 2015)), we average our loss over the valid pixels N_v , for which ground truth labels are available.

Besides the above disparity regression term, we use a new gradient term in our loss by considering large disparity discontinuities. We apply a new gradient regularizer on the disparity field to encourage similar change of disparities in the predicted and the ground truth disparity map and thus achieve more effective regularization. This is a key difference between our method and DispNet. As depth discontinuities are often accompanied by large disparity gradients, horizontal and vertical gradients of the disparity map convey information about significant depth differences in the scene

and local structure, which can be used to improve the quality of disparity maps. We minimize differences between gradients of the estimated disparity map and the ground truth to achieve an increase in performance. We apply the ℓ_1 norm to disparity gradients with the gradient regression loss \mathcal{L}_g defined as :

$$\mathcal{L}_{g} = \frac{1}{N_{v}} \sum_{i,j} \left[\left\| \nabla_{x} D_{i,j} - \nabla_{x} \widehat{D}_{i,j} \right\|_{1} + \left\| \nabla_{y} D_{i,j} - \nabla_{y} \widehat{D}_{i,j} \right\|_{1} \right]$$
(3)

Where $\|\cdot\|_1$ denotes the ℓ_1 norm, ∇_x and ∇_y are the horizontal and vertical gradient of the disparity maps. This gradient regression term encourages an estimated disparity map to have a similar local structure as a target disparity map and also encourages the network to find a local optimum to balance between the disparity and the gradient structure of the surfaces. The network is trained by minimizing the loss function E which is a weighted sum of these two terms, while λ_g controls the relative importance of the gradient regularizer in the optimization.

$$E = \mathcal{L}_r + \lambda_g \mathcal{L}_g \tag{4}$$

4 Experiments and Results

4.1 Dataset

In this paper, we use the synthetic Scene Flow dataset to train our model, and then evaluate it on some public competitive synthetic and real stereo datasets. An overview of these datasets is given in Tab. 1.

Dataset Name		Frames for training	Frames for testing	Ground Dis- parity	Synthetic or Real World	Resolution	
	FlyingThings3D	21818	4248	100%		960×540	
Scene Flow	Driving	8591	-	100%	Synthetic	960×540	
	Monkaa	4392	-	100%	Synthetic	960×540	
КІТТІ	KITTI2015	200	200	50% (sparse)	Bool world	1242×375	
dataset	KITTI2012	194	195	50% (sparse)		1226×370	
N	IPI Sintel	1064	564	100%	Synthetic	1024×436	
HCI		33	30	-	Real world	656×541	

Tab. 1 Overview of datasets used in our experiment

Scene Flow (MAYER et al. 2015) is a large synthetic dataset for stereo matching, first designed and used in DispNet for training CNNs to estimate disparity. This dataset is rendered by computer graphics methods and provides accurate dense ground truth, which is large enough to train a complex network. It contains three subsets and has more than 39,000 stereo frames in 960×540 pixel resolution. In this paper, we only use the FlyingThings3D subset to train our model. The Driving and Monkaa datasets are only used to evaluate our method and baselines.

The KITTI dataset was produced in 2012 (GEIGER et al. 2012) and extended in 2015 (MENZE et al. 2015, 2018). It contains stereo images of real-world complex road scenes collected from a calibrated pair of cameras mounted on a driving car. It provides 200 stereo frames with sparse ground truth obtained from a 3D laser scanner. Since the laser only provides sparse data up to a certain

distance and height and labels in some areas (e.g. sky) are hard or impossible to obtain, the ground truth in these areas are not available.

MPI Sintel (WULFF et al. 2012) is also an entirely synthetic dataset, which is created in the Blender software by rendering artificial scenes from a short open source animated 3D movie. It has 1064 training frames and provides dense ground truth disparities with large displacement, which is a very reliable test for comparison of methods. In this work, we use its final version because it contains sufficiently realistic scenes including natural image degradations.

HCI (MEISTER et al. 2012) is a challenge outdoor dataset, which contains eleven real-world scenes with a huge variety of different weather conditions, different motion, and depth layers. It has 330 frames and no ground truth. In this work, we only use it to show the visual quality of our method.

4.2 Implementation details

Training: We implemented our architecture and training phase using the Tensorflow framework (ABADI et al. 2015) and optimized our model end-to-end by choosing the Adam optimizer with default momentum parameters, $\beta 1 = 0.9$ and $\beta 2 = 0.999$. Due to hardware limitations, we trained the network on a Titan X GPU with a mini batch size of 4 image pairs. The training images were resized to 368*760 and preprocessed by normalizing them to zero mean and a standard deviation of 1. Since we used ReLu as activation functions and observed that "He initialization" (HE et al. 2015) worked better for layers with ReLu activation, we chose the "He initialization" method to initialize the weights of our network. We set the starting learning rate λ to be 1e-4 and then divided it by half every 150k-th iteration after the first 200k iterations. To avoid overfittings, we employed L2 regularization with a weight decay strength d=0.0004. The training weights of our final model were obtained at the 719k-th iteration because there was no improvement for 5 consecutive validations after this iteration. In addition, we performed online augmentation to introduce more variation in the training data, which includes geometric transformations (translation, scale) and chromatic transformations (brightness, contrast, gamma, and color).

Testing: we pre-trained our model on the FlyingThings3D dataset and evaluated it on other datasets. For evaluation of results, we used the EPE measure, which calculates the average Euclidean distance between predicted and ground truth disparity along all valid pixels. We also compared the performance of our method with other disparity estimation methods.

4.3 Results

In order to explore the effectiveness of our proposed method, we conduct two experiments on the aforementioned datasets. At the same time, we adopt DispNet as the baseline model; we present the qualitative and quantitative results.

Tab. 2 reports the corresponding experiment results in terms of EPE, where "Baseline" represents the model of DispNet, "Model_Final" represents our final model with all modifications and "Model_NoG" is the model without gradient regularizer. By comparing the results of the "Model_Final" to the "Baseline", we see that our model outperforms the baseline network in most cases and the EPE values are improved significantly (e.g. 27.9% on the Driving and 36.2% on the Sintel dataset). Since the Driving and Sintel datasets have larger disparities than the other datasets, the significant improvements on these two datasets can be directly attributed to our modification of the correlation module.

The improvement in visual quality is also distinct from the qualitative results, as show in Fig. 2. Compared with the baseline, our method performed noticeably better. We see that not only the resolutions of the disparity maps are improved, but also more detailed structures of the scenes are captured. For instance, the disparity estimates within the red boxes are improved by our method in Fig. 2. Furthermore, in large disparity discontinuity areas, our method can preserve clear edge details and produce correct disparity estimates because of using the gradient regularizer, as indicated by the red rectangular in Fig. 2.



Fig. 2: Visual results of our model and baseline model. (Colum 1: Left image; Colum 2: ground truth; Colum 3: results predicted by baseline network. Colum 4: results predicted by our final model.)

As mentioned above, we use an additional gradient regularizer in the loss function to penalize discontinuity on disparity maps. By comparing the results of the "Model_Final" to the "Model_NoG", we could verify the effectiveness of this gradient regularizer. From Tab. 2, we observe that the EPE values of the model with gradient regression loss (Model_Final) are slightly smaller than the model without gradient regression loss.

Dataset	Flying3D	Monkaa	Driving	Sintel	KITTI2015	KITTI2012
Baseline	1.68	5.78	12.46	5.66	1.59	1.55
Model_NoG	1.74	4.60	9.53	3.66	1.56	1.49
Model_Final	1.71	4.58	8.98	3.61	1.54	1.43
Improved(%)	-1.8/1.2	20.8/0.4	27.9/5.8	36.2/1.4	3.1/1.3	7.7/4.0

Tab. 2: EPE of different models on different datasets. The last row shows the improved accuracy when comparing 'Model_Final' to the baseline (first number) and to 'Model_NoG' (second number)

Since the EPE metric often favors over-smoothed solutions, it is interesting to also inspect qualitative results. Fig. 3 shows visual examples from "Model_Final" and "Model_NoS". As illustrated in the red box area, by using the gradient regularizer, our model performs well around the boundaries of objects. It is able to regularize the output effectively while learning to maintain sharpness and local structure details in the output disparity map. This is especially noticeable for the large disparity discontinuous area. These results indicate that utilizing the gradient regularizer has a positive impact on the performance.



Grount truth/left image

Model_NoS

Model_Final

Fig. 3: Comparison of results with and without the gradient regularizer. (Left: ground truth or left image; Middle: results predicted by Model_NoG without the gradient regularizer. Right: results predicted by our final model with the gradient regularizer. Note that we also give an example of a real world dataset "HCI" without ground truth in the third row.)

In addition, we investigate how well our method performs when compared with previously published methods, see Tab. 3. It is observed that, our end-to-end model achieves the best disparity estimation performance in terms of EPE in most cases, which demonstrates the strong generalization possibilities of our model. We note that SGM performs the worst. This confirms that CNNbased stereo matching algorithms have an explicit advantage for disparity estimation compared with more traditional methods. From Tab. 3, we see that the EPE values of MC-CNN-fs are larger than any other end-to-end method. That is because it needs extra post-processing steps to estimate complete disparity, which mainly decides the accuracy. The main contribution of DenseMapNet is reducing the parameter size and computation time. The results show this network has the least parameters and performs the best on the Driving dataset, with 0.3M parameters. However, the EPE values on other datasets are worse than those of our network. The above comparison demonstrates that our model achieves competitive performance through modification of correlation module and our gradient regularizer.

Dataset	Flying3D	Monkaa	Driving	Sintel	KITTI2015	KITTI2012	Parameters
SGM	8.70	20.16	40.19	19.62	7.21	10.06	
MC-CNN-fast	4.09	6.71	19.58	11.94			0.6M
Baseline	1.68	5.78	12.46	5.66	1.59	1.55	36M
DenseMapNet	5.07	4.45	6.56	4.41	2.52		0.3M
Our method	1.48	3.92	8.31	3.07	1.37	1.21	37M

Tab.3 Comparison results of our model with other methods (EPE)

5 Conclusions

In this paper, we have modified the baseline network DispNet, which we investigated experimentally. We increased the correlation range when computing the cost volume correlation module to handle large disparities. At the same time, a gradient regression loss derived from disparity gradient information is combined with the disparity regression loss to preserve sharper local structure details in large depth discontinuity area. The performance of our approach was evaluated on several challenging stereo datasets. The experiments demonstrate that our method achieves competitive performance and predicts more accurate and more detailed disparity maps in specific areas. We believe that semantic context information is crucial for stereo matching, especially in textureless regions. In the future work, we will investigate how to integrate semantic context to improve the performance in weakly textured area. At the same time, we will introduce more quality criteria, for instance, depth consistency (KocH et al. 2018), for the evaluation of our method.

6 Acknowledgements

The author Junhua Kang would like to thank the China Scholarship Council (CSC) for financially supporting her as a visiting PhD student at Leibniz University Hannover, Germany. We gratefully acknowledge the support of NVIDIA Corporation with the donation of GPUs used for this research.

7 References

- ABADI, M., AGARWAL, A., BARHAM, P., BREVDO, E., CHEN, Z., CITRO, C., CORRADO, G., DAVIS, A., DEAN, J. & DEVIN, M., 2015: Tensorflow: Large-scale machine learning on heterogeneous distributed systems.
- ATIENZA, R., 2018: Fast Disparity Estimation using Dense Networks. arXiv preprint arXiv:1805.07499.
- CHEN, J. & YUAN, C., 2016: Convolutional neural network using multi-scale information for stereo matching cost computation. IEEE International Conference on Image Processing, 3424-3428.
- CORTES, C. & VAPNIK, V., 1995: Support-vector networks. Machine Learning, 20, 273–297.
- DOSOVITSKIY, A., FISCHERY, P., ILG, E., HAUSSER, P., HAZIRBAS, C., GOLKOV, V., SMAGT, P. VAN DER, CREMERS, D. & BROX, T., 2015: FlowNet: Learning optical flow with convolutional networks. IEEE International Conference on Computer Vision, 2758-2766.
- 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.
- GODARD, C., MAC AODHA, O. & BROSTOW, G. J., 2017: Unsupervised monocular depth estimation with left-right consistency. CVPR., Vol. 2p. 7.
- GUNEY, F. & GEIGER, A., 2015: Displets: Resolving stereo ambiguities using object knowledge. IEEE Conference on Computer Vision and Pattern Recognition, 4165-4175.
- HE, K., ZHANG, X., REN, S. & SUN, J., 2015: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. IEEE International Conference on Computer Vision, 1026-1034.
- HIRSCHMULLER, H., 2008: Stereo processing by semiglobal matching and mutual information. IEEE Transactions on pattern analysis and machine intelligence, **30**, 328-341.
- HUANG, G., LIU, Z., VAN DER MAATEN, L. & WEINBERGER, K. Q., 2017: Densely connected convolutional networks. CVPR., Vol. 1p. 3.
- KENDALL, A., MARTIROSYAN, H., DASGUPTA, S., HENRY, P., KENNEDY, R., BACHRACH, A. & BRY, A., 2017: End-to-End Learning of Geometry and Context for Deep Stereo Regression. IEEE International Conference on Computer Vision, 66-75.
- KOCH, T., LIEBEL, L., FRAUNDORFER, F. & KÖRNER, M., 2018: Evaluation of CNN-based singleimage depth estimation methods. arXiv preprint arXiv:1805.01328.
- LECUN, Y., BOTTOU, L., BENGIO, Y. & HAFFNER, P., 1998: Gradient-based learning applied to document recognition. Proceedings of the IEEE., **86**, 2278-2324.
- LI, Y. & HUTTENLOCHER, D. P., 2008: Learning for stereo vision using the structured support vector machine. IEEE Conference on Computer Vision and Pattern Recognition, 1-8.
- LUO, W., SCHWING, A. G. & URTASUN, R., 2016: Efficient deep learning for stereo matching. IEEE Conference on Computer Vision and Pattern Recognition, 5695-5703.
- MAYER, N., ILG, E., HÄUSSER, P., FISCHER, P., CREMERS, D., DOSOVITSKIY, A. & BROX, T., 2015: A Large Dataset to Train Convolutional Networks for Disparity, Optical Flow, and Scene Flow Estimation. IEEE Conference on Computer Vision and Pattern Recognition, 4040-4048.

- MEISTER, S., JÄHNE, B. & KONDERMANN, D., 2012: Outdoor stereo camera system for the generation of real-world benchmark data sets. Optical Engineering, **51**, 21107.
- MENZE, M., HEIPKE, C. & GEIGER, A., 2015: Joint 3d estimation of vehicles and scene flow. ISPRS Workshop on Image Sequence Analysis (ISA), **8**.
- MENZE, M., HEIPKE, C. & GEIGER, A., 2018: Object scene flow. ISPRS Journal of Photogrammetry and Remote Sensing, **140**, 60-76.
- PANG, J., SUN, W., REN, J. S. J., YANG, C. & YAN, Q., 2017: Cascade Residual Learning: A Two-Stage Convolutional Neural Network for Stereo Matching. ICCV Workshops, 7.
- SCHARSTEIN, D. & SZELISKI, R., 2002: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision, 47, 7-42.
- SCHARSTEIN, D. & PAL, C., 2007: Learning conditional random fields for stereo. IEEE Conference on Computer Vision and Pattern Recognition, 1-8.
- SEKI, A. & POLLEFEYS, M., 2017: SGM-Nets: Semi-global matching with neural networks. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, 21-26.
- SHAKED, A. & WOLF, L., 2017: Improved stereo matching with constant highway networks and reflective confidence learning. IEEE Conference on Computer Vision and Pattern Recognition, 6901-6910.
- WULFF, J., BUTLER, D. J., STANLEY, G. B. & BLACK, M. J., 2012: Lessons and insights from creating a synthetic optical flow benchmark. European Conference on Computer Vision. Springer, 168-177.
- ZBONTAR, J. & LECUN, Y., 2016: Stereo matching by training a convolutional neural network to compare image patches. Journal of Machine Learning Research., **17**, 2.

Complementary Features Learning from RGB and Depth Information for Semantic Image Labelling

LIN CHEN¹, DAIXIN ZHAO¹, CHRISTIAN HEIPKE¹

Abstract: In this paper, we present a complementarity constraint for features computed from different sources of input data before fusion in semantic labelling. A two-branch encoderdecoder architecture with ResNet-50 is proposed and used as classification network. Our proposed complementarity constraint is added to the standard softmax cross-entropy classification loss. The impact of different weights for this constraint in multi-modal data fusion is investigated. The result of the two branch network is also compared to the one obtained with only the spectral information. The constraint is shown to improve the results consistently in our experiments. Different amounts of improvement are achieved when different weighs for the complementarity constraint are used.

1 Introduction

Semantic image labelling, called classification in remote sensing, is the process of assigning an object class label to each pixel in an input image. This labelling process for aerial and satellite images is one of the fundamental tasks in photogrammetry and remote sensing. Applications, e.g. land cover classification, land-use classification and change detection, rely heavily on semantic labelling. Therefore, semantic image labelling has been a focus of research in photogrammetry and remote sensing community for a long time.

Spectral information of images, such as RGB or IRRG, is normally the first data source for classification. However, multi-modal data such a combination of spectral and depth information is frequently available, and the height data can add additional information. For instance, the appearance of objects may change due to shadows and weather conditions (e.g. cloudy, snowy), but a digital surface model is not influenced by these effects.

In recent years, the research focus of semantic labelling has shifted from probabilistic graphical models, e.g., conditional random field (CRF), to deep neural networks. Deep neural networks are designed to extract useful features from the input automatically for the underlying task. Among many of the deep neural networks, convolutional neural networks (CNN) (LECUN et al. 1998) attracts most attention in image classification and semantic labelling. Fully Convolutional Networks (FCN), a variant of CNN without a fully connected layer, directly generates a classification map for the entire input images and is now a stand tool for semantic labelling. The encoder-decoder architecture based on FCN (NOH et al. 2015; BADRINARAYANAN et al. 2017; CHEN et al. 2018) shows state-of-the-art performance on many semantic labelling benchmarks.

One of the major concerns in using multi-source data for semantic labelling is how to properly fuse the data in a multi input branch encoder-decoder FCN. Features extracted from different

¹ Leibniz Universität Hannover, Institute of Photogrammetry and GeoInformation, Nienburger Str. 1, D-30167 Hannover, E-Mail: chen@ipi.uni-hannover.de, daixin.zhao@gmail.com, heipke@ipi.uni-hannover.de

sources of input is normally fused in the early or middle stages (i.e., before decoding) of the network. In this paper, we build a complementarity constraint to motivate the features from different source to be perpendicular to each other, thus "different" distinctive features are learned by the network. A comparison of adding this constraint with different strength for an encoder-decoder form of network proposed by us is also presented and analyzed.

2 Related Work

Semantic labelling is one of the major tasks in remote sensing image interpretation. Fully Convolutional Networks (FCN) (LONG et al. 2015) take an arbitrary size input image and then learn the features through a gradually down-sampled convolution with trainable kernels; a dense pixelwise classification map is then generated for the input image through an up-sampling stage. The final classification map can be 1/4 or 1/8 of the input image size. Several studies based on this architecture, e.g. SegNet (BADRINARAYANAN et al. 2017) and DeepLab (CHEN et al. 2015), are now standard baselines for semantic labelling. The following networks take the idea from FCN and achieve better accuracy in semantic labelling tasks. DeconvNet (NOH et al. 2015) builds a convolution-deconvolution network based on the VGG 16-layer Net (SIMONYAN et al. 2014) to upsample the convolutional feature maps to the original input image size. With the same backbone of VGGNet, SegNet presents an encoder-decoder architecture (bottleneck architecture) and stores the encoder max-pooling indices for producing sparse feature maps in the decoder stage. DeepLab v3+ (CHEN et al. 2018) takes advantages of the encoder-decoder architecture and Atrous (HOLSCHNEIDER et al. 1989) Spatial Pyramid Pooling (ASPP) from the first version of DeepLab. They utilize the dilated convolution (a.k.a. atrous convolution) to enlarge the receptive field of the convolution filter. Specifically, zero values are inserted between filter values to form a larger field-of-view without increasing the number of parameters. The rate parameter determines the number of zeros inserted between two adjacent filter matrix values, and thus controls the size of the receptive field. By using different rates of same size filters and stacking the results, a multi-scale response is imitated. These architectures have achieved notable performance improvements in semantic labelling tasks, e.g. using the PASCAL VOC 2012 (EVERINGHAM et al. 2014) and Cityscapes (CORDTS et al. 2015) datasets.

On the other hand, multi-modal data is frequently available in practical applications and is also provided in recent semantic segmentation contests, such as the ISPRS benchmark for 2D Semantic Labeling Contest (ROTTENSTEINER et al. 2014). Using multi-modal data can potentially improve the semantic segmentation performance because of the complementary information from different source. FuseNet (HAZIRBAS et al. 2016) makes use of RGB and depth (2.5D) with a two-branch-encoder and fuses the output from each branch at the end of encoder stage, and the fused information is then fed into a RGB-D decoder for generating a pixel-wise classification map. CHEN et al. (2018) investigate the impact of combinations of hand-crafted radiometric, e.g., NDVI (normalized difference vegetation index), and geometric features (e.g., nDSM, change of curvature) derived from the true orthophotos and DSM. The combination of RGB and normalized DSM (nDSM) delivers the best results in their investigation. AUDEBERT et al. (2018) explore the influence of where to fuse different source of information, e.g., IRRG, nDSM, NDVI and DSM data, in an encoder-decoder structure based variants of FuseNet. Their results show

that early fusion improves the semantic segmentation performance by learning multi-modal features jointly. Nevertheless, late fusion is helpful for hard pixels as the authors observed in experiments. In those mentioned works, two frequent fusion manners, summation and concatenation of input feature vectors for different modal, are used.

DualNet (HOU et al. 2017) focuses on learning complementary features from different subnetworks for same-source input data in object classification. Two parallel sub-networks without shared parameters extract features from the same source and then their extracted features are fused by summation (main-branch). A cross-entropy classifier based on the fused features acts as the main loss and two other cross-entropy classifiers built based on the features computed from each sub-network act as auxiliary losses. To motivate sub-networks to learn complementary features, only the loss from one sub-network and the main branch are calculated and the parameters of the other sub-network keep fixed in a single iteration. The two sub-networks are trained alternatively, which means the sub-networks exchange the role of being optimized and fixed after each iteration. This alternative learning strategy can prevent one sub-network from moving towards the same weights as the other, and hence allows both sub-networks are well trained, a joint fine tuning of all three branches is conducted, which yielded a minor performance improvement.

As far as we know, DualNet is the most closely related work to our research. However, their work concentrates on descriptor learning for object classification. In this paper, we hypothesize that if the mid-level features (those extracted at the end of the feature extraction stage in the net-work) computed from RGB and depth information are complementary to each other, then the information computed from RGB and depth will reinforce each other and thus provide more distinctive and complementary features for classification. Based on this hypothesis, we formulate a complementarity constraint for the mid-level features of RGB and depth information. The complementarity constraint leads to features which are perpendicular to each other in high dimensional feature space, thus the features span different dimensions in feature space, which means they are aimed at extracting "different" useful features for semantic segmentation.

3 Methodology

In this section, we first report the network we propose for semantic labelling. Then, the complementarity constraint is introduced, followed by details on the online augmentation we apply for training.

3.1 Network Architecture

DeepLab v3+ (CHEN et al. 2018) is selected as the basic architecture of our baseline network. Several modifications for DeepLab v3+ are introduced in our classification network. It contains an encoder and decoder stage. In the encoder stage the input image is fed into convolution blocks and downsized to a series of lower spatial resolution, e.g. 1/2, 1/4 of the input image, high dimensional mid-level feature maps are thus obtained. Then, dilated convolution is applied to the mid-level feature maps to enlarge the field-of-view. This procedure is called Atrous Spatial Pyramid Pooling (ASPP). In the decoder stage, different scales of down-sized features maps of the encoder stage are fused by concatenation to the corresponding scale feature maps in the decoder

stage. Through this fusion the spatial context and boundary details of objects can be preserved for generating finer segmentation results. After obtaining the full size up-sampled feature map, a pixel-wise classification digits map is built through a few convolution blocks. If the underlying task is pixel-wise classification, then the cross-entropy is calculated as the classification loss by using the digits map. Thus, the network output is a classification map with a same size of the input image. However, the output can also be at a lower resolution, like 1/4 of the input image, once the output quality is able to meet the requirements of application.



Fig. 1: Illustration of Atrous Spatial Pyramid Pooling (ASPP). Max-pooling, a 1×1 convolution and three atrous convolutions with rate parameters 2, 4, 6, are used to generate the ASPP feature map.

To increase the learning ability of the network, our variant of DeepLab v3+ takes advantage of the ResNet-50 network (HE et al. 2016), we modify it for our encoder stage. The ResNet-50 network has 5 stages, more specifically, 50 convolutional layers. Each stage consists of one convolutional block and 3 to 5 identity blocks. In our network, the 7×7 convolution with stride 2 in the first block is replaced by a 5×5 convolution (with unchanged stride), and the following 3×3 maxpooling is dropped. By using a smaller convolution kernel and dropping max-pooling in the first block, less smoothing is conducted and thus more explicit boundary information is preserved to differentiate objects, like buildings, that normally have sharp boundary information. Later in the encoder stage we choose dilated convolution with rate parameters 2, 4, 6 in ASPP to preserve more continuous spatial information at differing scales, as shown in Fig. 1.

To classify the aerial images with different modality, two branches for different sources of data are used in the encoder stage. IRRG and nDSM/NDVI data are fed into the two-branch-encoder separately to extract features. The extracted mid-level features are fused by summation. The feature decoder up-samples the ASPP feature map with bilinear interpolation. During up-sampling, the 1/4 and 1/8 feature maps from both branches are concatenated with corresponding scale of feature maps in the decoder. In the end, a classification map with the same size of the input image is calculated through a softmax classifier. The network architecture is described in Fig. 2.



Fig. 2: Demonstration of network architecture. 1/4, 1/8 and 1/16 feature maps from encoder stages are fused in the decoder stage with the same size up-sampled feature maps. An ASPP feature map is generated using the fused feature map up-sampled for the decoder stage. The network is trained with a softmax cross-entropy loss and a complementarity constraint.

Input of the second branch are the nDSM and the NDVI. The nDSM captures representations of height differences, while the NDVI contains the discriminative information for vegetations and other classes. Therefore, the combination of the two input data can distinguish high vegetation from low vegetation, or different classes of objects with similar height, for instance, trees and buildings. The NDVI for each individual image pixel is derived by:

$$NDVI = \frac{IR - R}{IR + R}$$
(1)

in which IR stands for the near infrared band, R represents red band.

3.2 Complementary loss

Following our hypothesis that mid-level features computed from the two branches should be complementary to each other, we formulate a complementarity constraint for learning discriminative, yet complementary features for the classification. To enforce dissimilarity of the two feature vectors computed from different source of data, they should be perpendicular to each other. The complementarity constraint is illustrated in Fig. 3.



Fig. 3: Complementarity constraint. H₁, H₂, W₁, W₂, C₁, C₂, stand for the height, width and channel numbers for feature maps from the IRRG and the nDSM + NDVI branch, respectively.

Our complementary loss is described by the cosine similarity between two non-zero vectors in high dimensional feature space. The cosine similarity lies the range [-1, 1], where a similarity of -1 is computed from two antiparallel vectors and a similarity of 1 is computed from two parallel vectors. Perpendicular vectors have a similarity of 0, which means the two features provide complementary information. The cosine similarity is defined as:

cosine similarity =
$$\frac{\mathbf{A} \cdot \mathbf{B}}{\|\|\mathbf{A}\|\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}}$$
(2)

where **A** and **B** stand for two non-zero vectors with n entries. Since the complementary loss should penalizes vector combinations providing similar information, the absolute value of cosine similarity that neatly bound the loss value between 0 and 1 is used for simplicity. When training the network with the complementary loss, we treat it as a regularization term. The final loss function is defined as:

$$\mathcal{L} = \mathcal{L}_{\text{classification}} + \lambda \cdot \mathcal{L}_{\text{complementary}}$$
(3)

The parameter λ represents the regularization strength of the complementary loss. We investigate different values of λ in our experiments, the results are discussed in Section 4.

3.3 Online augmentation

To overcome the relatively small amount of available training data and alleviate potential overfitting, online augmentation is applied in our training process. The augmentation contains random flipping (vertically or horizontally) and random rotation $(0^{\circ}, 90^{\circ}, 180^{\circ}, 270^{\circ})$ of all input data and the corresponding ground truth label. A random tag for each training sample in each mini-batch is set to switch the augmentation during training on or off. Overall, 50% of the samples used for forward propagation are augmented, and the other 50% samples are original data without augmentation.

4 Experimental results

4.1 Dataset

We use the Vaihingen image dataset from the ISPRS 2D Semantic Labelling Contest⁽¹⁾ for our experiments. The goal of this contest is to label images using multiple object categories, namely tree, building, low vegetation, impervious surfaces, car and clutter/background. This dataset contains very high resolution true orthophotos (TOP) with three bands (near infrared, red and green), manually labelled ground truth and a corresponding DSM derived from dense image matching. Instead of directly using the provided DSM, we use an nDSM released by GERKE (2014) to be independent of absolute heights. As a side effect, the noise inherent in the DSM is decreased by the filtering operation during the generation of the nDSM.

The Vaihingen dataset contains 33 TOP images with a ground sampling distance (GSD) of 9 cm, for 16 of them ground truth is provided. The test results are based on the test set containing 17 images with so called uneroded and eroded ground truth released recently. Uneroded ground truth contains the complete reference of the test data, whereas eroded ground truth ignores the object boundaries in a buffer of 3 pixels width to reduce the effects of uncertainty in boundary definition.

4.2 Parameters setup

We randomly choose 13 images for training and the other 3 images for validation. The TOP images, nDSM, NDVI and ground truth images are cropped into 256×256 pixel patches by sliding a 256×256 window with step of 64 pixels. That means adjacent patches overlap by 75%. After cropping, 4037 image patches are obtained for training.

We also utilize dropout introduced by SRIVASTAVA et al. (2014) after the 1/8 feature map concatenation in the decoder stage with the *keep_prob* parameter as 0.7, i.e. randomly dropping 30% connections during the training process to prevent overfitting.

Our model is implemented using the Tensorflow framework (ABADI et al. 2016) from scratch. We train our model for 200,000 iterations using Nesterov's accelerated gradient with momentum 0.9 and the learning rate 0.0003. Additionally, the validation is based on a sliding window with a step of 50 pixels and the classification results from each 256×256 pixel patch are assembled with equal weights to generate final classification maps. The model with the best validation overall accuracy is retained.

4.3 Evaluation Criteria

The evaluation is based on the pixel-wise confusion matrix. The correctness (precision), completeness (recall), F1 score and overall accuracy are derived from the confusion matrix:

$$Precision = \frac{t_p}{t_p + f_p}$$
(4)

⁽¹⁾ http://www2.isprs.org/commissions/comm3/wg4/semantic-labeling.html (accessed on Nov 21, 2018)

L. Chen, D. Zhao & C. Heipke

$$\text{Recall} = \frac{t_p}{t_p + f_n} \tag{5}$$

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{percision} + \text{recall}}$$
(6)

Overall Accuracy =
$$\frac{t_p + t_n}{t_p + t_n + f_p + f_n}$$
(7)

Where t_p , t_n , f_p , f_n represent the number of true positives, true negatives, false positives and false negatives, respectively.

4.4 Results

First, we compare the result of only using IRRG and using IRRG and nDSM/NDVI. Then different complementary strength λ , i.e. 0.1, 0.01, 0.001, for features computed from IRRG and nDSM/NDVI are compared. Table 1 shows the overall accuracy of different experiment setups on uneroded test dataset and eroded test dataset.

Experiment setups	OA on uneroded test dataset	OA on eroded test dataset
Only IRRG	84.4%	87.2%
$\lambda = 0$	85.3%	88.2%
$\lambda = 0.1$	86.0%	89.0%
$\lambda = 0.01$	86.2%	89.1%
$\lambda = 0.001$	86.1%	89.0%

Tab. 1: Overall Accuracy (OA) on different experiment setups

As illustrated in Table 1, the model trained only by a single IRRG branch encoder is outperformed by the two-branch-encoder network with IRRG and nDSM/NDVI as inputs, which is expected as the multi-modal data provide richer information for semantic segmentation. Notably, setting the complementary strength λ to 0.01 provides the best overall accuracy, as also shown in Table 2 and Table 3.

Table 2 and Table 3 show the experimental results with the precision, recall and F1 score calculated from the eroded and the uneroded data, respectively. Interestingly, we find that the highest F1 scores for each class are actually not given by the experiment setup with λ equals 0.01. For instance, cars are better classified by the network trained only with optical images. Our first interpretation is that the NDVI does not contain enough distinguishable features to separate cars from impervious surfaces and buildings. Also, the nDSM height difference might contain ambiguities on these relatively small objects due to problems in dense matching caused by lack of texture on cars. Therefore, the nDSM can contain some noisy information when classifying cars compared to other categories. Moreover, when the complementary strength is set to 0.1 or 0.001, we observe a relatively higher F1 score on building or trees and low vegetation.

Experiment setups	Evaluation criteria	Impervious surface	Building	Low vege- tation	Tree	Car	Clutter/ Background
	Precision	0.845	0.907	0.807	0.809	0.781	0.764
Only IRRG	Recall	0.888	0.890	0.719	0.894	0.597	0.184
	F1 score	0.866	0.898	0.760	0.850	0.677	0.296
	Precision	0.836	0.925	0.836	0.815	0.771	0.805
$\lambda = 0$	Recall	0.910	0.912	0.723	0.895	0.429	0.013
	F1 score	0.871	0.918	0.775	0.853	0.551	0.025
	Precision	0.844	0.925	0.841	0.836	0.648	0.687
$\lambda = 0.1$	Recall	0.914	0.931	0.729	0.883	0.588	0.097
	F1 score	0.878	0.928	0.781	0.859	0.617	0.170
	Precision	0.850	0.933	0.845	0.820	0.778	0.844
$\lambda = 0.01$	Recall	0.910	0.916	0.741	0.904	0.562	0.097
	F1 score	0.879	0.924	0.790	0.860	0.652	0.174
$\lambda = 0.001$	Precision	0.839	0.929	0.831	0.845	0.792	0.797
	Recall	0.910	0.900	0.773	0.888	0.514	0.136
	F1 score	0.873	0.914	0.801	0.866	0.623	0.232

Tab. 2: Experimental results based on the uneroded dataset

Tab. 3: Experimental results based on the eroded dataset

Experiment setups	Evaluation criteria	Impervious surface	Building	Low vege- tation	Tree	Car	Clutter/ Background
	Precision	0.872	0.929	0.841	0.837	0.809	0.797
Only IRRG	Recall	0.916	0.905	0.752	0.922	0.699	0.193
	F1 score	0.893	0.917	0.794	0.877	0.750	0.310
	Precision	0.867	0.947	0.867	0.844	0.791	0.827
$\lambda = 0$	Recall	0.935	0.929	0.762	0.922	0.528	0.014
	F1 score	0.899	0.938	0.811	0.881	0.633	0.027
	Precision	0.877	0.947	0.874	0.863	0.670	0.726
$\lambda = 0.1$	Recall	0.939	0.948	0.768	0.914	0.676	0.104
	F1 score	0.907	0.947	0.818	0.888	0.673	0.182
	Precision	0.881	0.952	0.877	0.852	0.806	0.867
$\lambda = 0.01$	Recall	0.934	0.934	0.781	0.930	0.669	0.104
	F1 score	0.907	0.943	0.826	0.889	0.731	0.186
$\lambda = 0.001$	Precision	0.868	0.951	0.864	0.876	0.816	0.825
	Recall	0.936	0.915	0.811	0.917	0.619	0.144
	F1 score	0.900	0.933	0.837	0.896	0.704	0.246



Fig. 4: Top: the top-left part of the test image 38, from left to the right, are shown: IRRG image (a), nDSM (b), NDVI (c), ground truth (d) and eroded ground truth (e). Bottom row from left to right: Classification results for network trained only using IRRG (f), multi-modal inputs network trained with complementary strength λ of 0 (g), 0.1 (h), 0.01 (i) and 0.001 (j).

Fig. 4 emphasizes that by using the complementary strength λ of 0.01, the network generates better classification results. The top row shows the input and ground-truth labels and the bottom row shows the classification result. In this area, low vegetation is classified as building without information from other sources. The bottom part of the input image is poorly classified except for the case of $\lambda = 0.01$. Our interpretation is that the lower vegetation is dry and shows more ambiguity to spectral features and the NDVI of buildings, and thus it cannot be differentiated well from buildings as shown in Fig. 4 (f). Involving the nDSM introduces depth information and thus contributes a lot towards a better differentiation in this case, as shown in the Fig. 4 (g, h, i, j). However, the strength of the complementary constraint does matter, as $\lambda = 0.01$ (Fig. 4(i)) delivers the best features to differentiate the lower drier vegetation and buildings.

Our experiments show that even though the complementary strength of 0.01 does not give the best results for each class, it strikes a better balance for all frequent categories and thus proves its effect.

5 Conclusion

In this work, we investigate a two-branch-encoder and decoder architecture with complementarity constraint for semantic segmentation using high resolution aerial images. Our proposed complementary constraint shows a notable performance improvement for semantic labelling with

different source of input data. In future work we will test more choices for the complementarity constraint and test our approach with different datasets.

Acknowledgement

Part of this work is done by DAIXIN ZHAO during an internship at Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover. The authors LIN CHEN and CHRISTIAN HEIP-KE are grateful to NVIDIA Corp. for the GPU donation. Also, the author LIN CHEN and DAIXIN ZHAO are grateful to CHUN YANG and JUNHUA KANG for valuable discussions.

References

- ABADI, M., BARHAM, P., CHEN, J., CHEN, Z., DAVIS, A., DEAN, J., DEVIN, M., GHEMAWAT, S., IRVING, G., ISARD, M., KUDLUR, M., LEVENBERG, J., MONGA, R., MOORE, S., MURRAY, D.G., STEINER, B., TUCKER, P., VASUDEVAN, V., WARDEN, P., WICKE, M., YU, Y. & ZHENG, X., 2016: Tensorflow: a system for large-scale machine learning. In OSDI, 16, 265-283.
- AUDEBERT, N., LE SAUX, B. & LEFÈVRE, S., 2018: Beyond RGB: Very high resolution urban remote sensing with multimodal deep networks. ISPRS Journal of Photogrammetry and Remote Sensing, 140, 20-32.
- BADRINARAYANAN, V., KENDALL, A. & CIPOLLA, R., 2017: Segnet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39, 2481-2495.
- CHEN, L.C., PAPANDREOU, G., KOKKINOS, I., MURPHY, K. & YUILLE, A.L., 2018: Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE transactions on pattern analysis and machine intelligence, **40**(4), 834-848.
- CHEN, L.C., ZHU, Y., PAPANDREOU, G., SCHROFF, F. & ADAM, H., 2018: Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. European Conference on Computer Vision (ECCV). 8-14 September 2018, Munich, Germany
- CHEN, K., WEINMANN, M., SUN, X., YAN, M., HINZ, S., JUTZI, B. & WEINMANN, M., 2018: Semantic Segmentation of Aerial Imagery via Multi-scale Shuffling Convolutional Neural Networks with Deep Supervision. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences, 4(1). 29-36
- CORDTS, M., OMRAN, M., RAMOS, S., SCHARWÄCHTER, T., ENZWEILER, M., BENENSON, R., FRANKE, U., ROTH, S. & SCHIELE, B., 2015: The cityscapes dataset. CVPR Workshop on the Future of Datasets in Vision, 1(2).
- EVERINGHAM, M., ESLAMI, S.M.A., GOOL, L.V., WILLIAMS, C.K.I., WINN, J. & ZISSERMAN, A., 2014: The Pascal Visual Object Classes Challenge: A Retrospective. International Journal of Computer Vision, 111(1), 98-136.
- GERKE, M., 2014: Use of the stair vision library within the ISPRS 2D semantic labeling benchmark (Vaihingen). Technical Report. ITC, University of Twente.
- HAZIRBAS, C., MA, L., DOMOKOS, C. & CREMERS, D., 2016: Fusenet: Incorporating depth into semantic segmentation via fusion-based cnn architecture. Asian Conference on Computer Vision, Taipei Taiwan, China, 213-228.
- HE, K., ZHANG, X., REN, S. & SUN, J., 2016: Deep residual learning for image recognition. IEEE conference on computer vision and pattern recognition, Caesars Palace in Las Vegas, Nevada, USA, 770-778.
- HOU, S., LIU, X. & WANG, Z., 2017: Dualnet: Learn complementary features for image recognition. International Conference on Computer Vision (ICCV), Oct. 22-29 2017, Venice, Italy, 502-510.
- LECUN, Y. BOTTOU, L. BENGIO, Y. & HAFFNER, P., 1998: Gradient-based learning applied to document recognition. Proceedings of the IEEE, **86**(11), 2278-2324.
- LONG, J., SHELHAMER, E. & DARRELL, T., 2015: Fully convolutional networks for semantic segmentation. IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 3431-3440.
- HOLSCHNEIDER, M., KRONLAND-MARTINET, R., MORLET, J. & TCHAMITCHIAN, P., 1989: A RE-AL-time algorithm for signal analysis with the help of the wavelet transform," in *Wavelets:* Time-Frequency Methods and Phase Space, 289-297.
- NOH, H., HONG, S. & HAN, B., 2015: Learning deconvolution network for semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 1520-1528.
- ROTTENSTEINER, F., SOHN, G., GERKE, M. & WEGNER, J.D., 2014: ISPRS semantic labeling contest. http://www2.isprs.org/semantic-labeling.html.
- SIMONYAN, K. & ZISSERMAN, A., 2014: Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv: 1409.1556.
- SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., SUTSKEVER, I. & SALAKHUTDINOV, R., 2014: Dropout: a simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, 15(1), 1929-1958.

Deep Learning zur Analyse von Bildern von Seidenstoffen für Anwendungen im Kontext der Bewahrung des kulturellen Erbes

MAREIKE DOROZYNSKI¹, DENNIS WITTICH¹ & FRANZ ROTTENSTEINER¹

Zusammenfassung: Der vorliegende Beitrag befasst sich mit der Klassifikation von Bildern mittels Deep Learning. Exemplarisch wird hier die bildbasierte Prädiktion der Entstehungsepoche von Seidenstoffen und Fertigungsskizzen betrachtet. An diesem Beispiel wird die Problemstellung von nicht eindeutig definierten Klassen sowie uneindeutigen Trainingsbeispielen aufgegriffen und ein Lösungsansatz vorgestellt. Dieser basiert auf einer erweiterten Verlustfunktion, mit welcher die uneindeutigen Trainingsbeispiele während des Trainings genutzt werden können. Hierzu werden während der Trainingsphase übergreifende Klassen definiert. Die durchgeführten Experimente zeigen, dass hiermit die Genauigkeit der Klassifikation ohne die übergreifenden Klassen verbessert werden kann. Da der exemplarische Datensatz relativ wenig Daten umfasst wird ein vortrainiertes Convolutional Neural Network verwendet und eine Augmentierung der Daten durchgeführt. Der Einfluss der Augmentierung sowie der entwickelten Verlustfunktion wird in einer Kreuzvalidierung evaluiert.

1 Einleitung

Für die Bewahrung des kulturellen Erbes sind Kunstsammlungen von hoher Bedeutung. Ebenso wichtig wie die eigentlichen Objekte sind hierbei die zugehörigen Informationen, wie beispielsweise das Datum der Entstehung, der Name des Künstlers oder die Stilrichtung. In aller Regel ist die Klassifikation von Kunstobjekten (im Sinne der Generierung oder Vervollständigung der zugehörigen Informationen) eine Aufgabe, die bislang nur von Fachleuten wie Kunsthistorikern bewältigt werden kann. Folglich erweist sich die Klassifikation mehrerer oder gar aller Objekte einer Sammlung oft als eine langwierige Aufgabe. Hieraus ergibt sich der Wunsch, die Klassifikation weitestgehend zu automatisieren. Bei ausreichender Anzahl klassifizierter Objekte ist es naheliegend, Verfahren des maschinellen Lernens heranzuziehen. Liegen beispielsweise digitale Bilder der Kunstwerke vor und ist der Hersteller von ausreichend vielen Werken bekannt, so kann ein Modell auf diesen Daten trainiert werden, das nach Abschluss des Trainings im Idealfall auch den Hersteller von weiteren Werken anhand von Bildern prädizieren kann.

Mit dieser Motivation hat sich das EU-Projekt SILKNOW unter anderem das Ziel gesetzt, eine automatische Methode für die bildbasierte Klassifikation von Seidenstoffen zu entwickeln. Die zu prädizierenden Variablen sind hierbei beispielsweise der Hersteller, die Stilrichtung oder die Herstellungsepoche des jeweiligen Seidenstoffes. In der vorliegenden Arbeit, die im Rahmen des Projektes entwickelt wurde, wird eine Methode für die bildbasierte Klassifikation vorgestellt. Als Datensatz dient exemplarisch eine Sammlung von Bildern von Seidenstoffen sowie zugehöriger Skizzen die für das Projekt zur Verfügung gestellt wurden (siehe Abbildung 1). Für alle Objekte

¹ Leibniz Universität Hannover, Institut für Photogrammetrie und GeoInformation, Nienburger Str. 1, D-30167 Hannover, E-Mail: [dorozynski, wittich, rottensteiner]@ipi.uni-hannover.de

der Sammlung liegen neben den Bildern auch gesammelte Informationen vor, z.B. über die Herstellungsepoche, also die zeitliche Einordnung der Seidenstoffe in Bezug auf die Fertigstellung. Die Herstellungsepoche wird in diesem Beitrag als zu prädizierende Variable gewählt, da sich hieraus eine Problematik ergibt, die über die ,klassische' Klassifikation von Bildern hinausgeht. In der ,klassischen' Klassifikation werden vorab disjunkte, eindeutige Klassen definiert, was jedoch bei der Prädiktion von Epochen, im Sinne von zeitlichen Abschnitten, aufgrund der Heterogenität der vorliegenden Annotationen oft nicht möglich ist. Auch in dem hier betrachteten Datensatz, in welchem die Angaben zu den Entstehungsepochen von Jahresangaben bis hin zu einer Zeitspanne von über 500 Jahren reichen und sich teilweise überlappen, ist die Definition von diskunkten Klassen nicht möglich. Hieraus ergibt sich das Problem von uneindeutigen Trainingsbeispielen bzw. Klassenlabels der Bilder. Hierunter verstehen sich Trainingsbeispiele, deren gegebene Annotation der Epoche nicht vollständig in einen Zeitraum fällt, der als Klasse definiert wurde.

Eine weitere Problematik, die in diesem Beitrag adressiert wird, ist das Training auf einem relativ kleinen Datensatz. Dies ist insbesondere bei tiefen neuronalen Netzen ein Problem, da diese oft mehrere Millionen Parameter aufweisen, deren Optimierung eine ausreichende Menge an Trainingsbeispielen erfordert.

Der hier gewählte Lösungsansatz basiert auf einem Convolutional Neural Network (CNN), das teilweise auf dem ImageNet-Datensatz (DENG et al. 2009) vortrainiert wurde. Die Verwendung eines vortrainierten Netzwerkes zur Merkmalsextraktion ist ein Ansatz, der sich in der Vergangenheit insbesondere für das anschließende Training auf kleinen Datensätze etabliert hat. Für den Umgang mit uneindeutigen Trainingsbeispielen wird eine erweiterte Verlustfunktion entwickelt, sodass auch diese für das Training des Netzwerkes genutzt werden können.

In Abschnitt 2 wird zunächst erläutert, welche Lösungsansätze für die bildbasierte Kunstklassifikation existieren und es wird ein Überblick über den bisherigen Umgang mit uneindeutigen Trainingsbeispielen gegeben. Anschließend wird in Abschnitt 3 der verwendete Datensatz vorgestellt. In Abschnitt 4 werden die Grundlagen beschrieben, auf welchen die Methodik in Abschnitt 5 aufbaut. In Abschnitt 6 wird der Ansatz experimentell evaluiert. Zuletzt wird der vorliegende Beitrag in Abschnitt 7 zusammengefasst und die wesentlichen Folgerungen werden herausgestellt.

2 Stand der Forschung

Die automatische, bildbasierte Klassifikation, Bewertung und Beurteilung von Kunstwerken ist eine Problemstellung, die bereits in zahlreichen Publikationen adressiert wurde. Frühere Arbeiten wie (BLESSING & WEN 2010) basieren oft auf manuell selektierten Merkmalen wie Histogramme der Farbwerte und Gradienten, die anschließend mit klassischen Klassifikationsmodellen, hier eine Support Vector Machine, klassifiziert werden. Die Autoren erreichen mit diesem Ansatz eine Genauigkeit von 85,1 % bei der Prädiktion des Künstlers.

Im Zuge der Weiterentwicklung der künstlichen neuronalen Netze (ANNs) und insbesondere der CNNs wurden auch diese Modelle auf kunstbezogene Klassifikationsprobleme angewendet. Beispielsweise verwenden AYUB et al. (2009) ein CNN zur Prädiktion des Preises von Gemälden. Sie erreichen hierbei jedoch nur geringe Genauigkeiten, was sie selbst unter anderem damit begründen, dass der verwendete Datensatz mit 700 Bildern zu wenig Trainingsbeispiele beinhaltet. SUR & BLAINE (2017) zeigen jedoch, dass die Klassifikation von Kunstwerken mit CNNs prinzipiell möglich ist. Sie erreichen bei der Prädiktion der Künstler (25 verschiedene) unterschiedlicher Gemälde Genauigkeiten über 80 %. Im Gegensatz zu AYUB et al. Trainieren sie ihr Modell auf einem größeren Datensatz mit 13574 Bildern. Zudem nutzen sie zur Extraktion der Merkmale ein CNN, welches zuvor auf dem ImageNet Datensatz trainiert wurde. RAZAVIAN et al. (2014) zeigen, dass der Ansatz ein vortrainiertes Netzwerk für die Merkmalsextraktion zu verwenden, insbesondere für das Training auf kleinen Datensätzen sinnvoll ist. Ein verwandter Ansatz ist die Feinanpassung der Parameter eines bestehenden Modelles (Finetuning). Hierbei werden die Parameter eines vortrainierten Netzwerkes zur Initialisierung verwendet und während des Trainings auf dem Zieldatensatz weiter optimiert. Diesen Ansatz verfolgen auch HICSONMEZ et al. (2017), die ein Modell für die Prädiktion des Illustrators von Kinderbüchern trainieren. Da der dort verwendete Datensatz nur 6468 Bilder umfasst, verwenden die Autoren eine Datenaugmentierung, mit deren Hilfe sie eine Klassifikationsgenauigkeit von 90,0 % erreichen. Neben der Prädiktion des Künstlers befassen sich einige Arbeiten mit der Klassifikation von Kunstwerken nach Stil bzw. Genre (BAR et al. 2014; TAN et al. 2016; HENTSCHEL et al. 2016). In diesen Ansätzen wird verdeutlicht, dass die Verwendung von vortrainierten Netzwerken stets geringere Klassifikationsfehler erzielt als das Trainieren eines zufällig initialisierten Netzes.

In der Recherche zu dieser Arbeit konnte die Klassifikation von Seidenstoffen in keiner Arbeit ausgemacht werden. Lediglich XIAO et al. (2018) befassen sich mit der automatischen Klassifikation von Strickmustern mit künstlichen neuronalen Netzen. Das Modell ist hierbei auf hochauflösenden Nahaufnahmen der Gewebe trainiert, wodurch sich das Problem deutlich von dem hier behandelten unterscheidet.

Obgleich sich alle genannten Arbeiten mit der bildbasierten Klassifikation von Kunst befassen, verbleiben zwei signifikante Unterschiede gegenüber der vorliegenden Arbeit. Der erste Unterschied liegt darin, dass die Datensätze, die in den vorgestellten Arbeiten herangezogen wurden, überwiegend Digitalisierungen umfassen, die durch Scanverfahren oder in Form von rektifizierten Fotografien mit frontaler Beleuchtung erzeugt wurden. In dieser Arbeit wird der Umgang mit Fotografien der Kunstwerke behandelt, die starke Schwankungen sowohl in der Beleuchtung als auch in der Platzierung des eigentlichen Objektes innerhalb der Abbildung aufweisen. Der zweite Unterschied liegt in der Wahl der zu prädizierenden Klassen. Keine der Arbeit befasst sich mit dem zuvor erläuterten Problem der uneindeutigen Klassenlabels. Hierfür existieren einige Ansätze, in denen zusätzliche, übergreifende Klassen definiert werden und somit eine Art hierarchische Klassenstruktur gebildet wird. Einen Ansatz, basierend auf Support Vector Machines, stellen VURAL et al. (2004) vor. Einen ANN basierten Ansatz liefern WU et al. (2017). Beide Ansätze unterscheiden sich von dem hier vorgestellten Ansatz, da in diesem Beitrag die übergreifenden Klassen nicht prädiziert werden sollen. Stattdessen sollen die uneindeutigen Trainingsbeispiele für die Verbesserung der Prädiktion der Basisklassen genutzt werden.

3 Datensatz

Der hier genutzte Datensatz wurde von GARÍN zur Verfügung gestellt. Er enthält 2752 Bilder von Seidenstoffen und 2168 Bilder von Entwurfsskizzen (siehe Abbildung 1). In beiden Fällen ist

meistens sowohl eine Abbildung der Vorderseite als auch der Rückseite sowie gelegentlich auch eine Detailaufnahme vorhanden. Die Bilder haben eine Größe zwischen 500 x 500 Pixeln und 4000 x 3500 Pixeln. Zudem existieren Datenblätter mit detaillierten Beschreibungen der abgebildeten Stoffe. Diese enthaltenen z.B. Angaben bezüglich des Materials, der Herstellungsepoche oder der Weberei. Letztere variiert jedoch nicht, da alle Objekte der Sammlung von der Weberei der Familie Garín stammen. Der Familienbetrieb produziert jedoch seit dem Jahr 1820, sodass Stoffe aus fast zwei Jahrhunderten vorliegen; das älteste abgebildete Objekt ist eine Skizze aus dem Jahr 1827 und das jüngste Werk stammt aus dem Jahr 2015.



Abb. 1: Beispiele für das Bild einer Vorderseite (links) und einer zugehörigen Rückseite (Mitte) eines Seidenstoffs sowie ein Beispiel für ein Bild einer Entwurfsskizze (rechts)

Um die Bilder im Rahmen einer Klassifikation der Epoche nutzen zu können, müssen die gegebenen zeitlichen Angaben zunächst einzelnen Klassen zugeordnet werden. Eine Herausforderung bei der Einordnung der Beschreibungen der Herstellungsepoche in eine Klassenstruktur ist die stark variierende Genauigkeit der Zeitangaben die sich teilweise überlappen (siehe Abbildung 2). Für manche Stoffe ist der genaue Tag der Herstellung bekannt, wohingegen für andere die Entstehungszeit lediglich auf ein Jahrhundert eingegrenzt werden kann. Überschneidungen gibt es beispielsweise bei den Zeiträumen "2. Hälfte 19. Jahrhundert" und "1889-1912". Bei einer Klasseneinteilung in einzelne Jahre oder Jahrzehnte wären beim vorliegenden Datensatz nicht ausreichend viele Beispiele pro Klasse vorhanden, wie Abbildung 2 verdeutlicht.



Abb. 2: Darstellung der vorliegenden Zeitangaben anhand eines Zeitstrahls, wobei der farblichen Codierung die Häufigkeit der jeweiligen Beschreibung zu entnehmen ist

Als Klassen werden daher im vorliegenden Beitrag halbe Jahrhunderte gewählt, sodass sich in Summe die drei Klassen "2. Hälfte 19. Jahrhundert", "1. Hälfte 20. Jahrhundert" und "2. Hälfte 20. Jahrhundert" ergeben. Alle Bilder, deren Epochenangabe vollständig in einen dieser Zeiträume fällt, werden der entsprechenden Klasse zugeordnet. Die resultierende Klassenverteilung

ist in Tabelle 1 unter der Bezeichnung "Basisklassen" dargestellt. Insgesamt lassen sich auf diese Weise 2876 der 4920 Bilder (ca. 58 %) einer der gewählten Klassen zuweisen, wobei sich diese Menge aus 2012 Seidenstoffbildern und 864 Skizzenbildern zusammensetzt. Zusätzlich werden Seidenstoffbilder, die in Bezug auf die drei Basisklassen eine uneindeutige Herstellungszeit haben sich jedoch zwei der Basisklassen umfassenden Zeiträumen von 100 Jahren zuordnen lassen, entsprechenden zusammengesetzten Klassen zugewiesen (siehe Tabelle 1).

	Basisklassen			Zusammenges	etzte Klassen
Klassenbezeichnung	2. Hälfte	1. Hälfte	2. Hälfte	1850	1900
	19. Jh.	20. Jh.	20. Jh.	-1950	-2000
Anzahl Bilder	213	2035	628	302	140

Tab. 1: Übersicht über die Verteilung der Bilder auf die Klassen

4 Grundlagen

4.1 Convolutional Neural Networks

Durch CNNs können sowohl Merkmale aus Bildern als auch die Abbildung dieser Merkmale auf eine Menge von Klassen gelernt werden (KRIZHEVSKY et al. 2012). Hierfür werden sogenannte Convolutional Layer verwendet, welche Faltungen auf den Bildern ausführen, wobei die Gewichte der Faltungsmatrizen im Training ermittelt werden. Da auf diese Weise nur lineare Abbildungen realisiert werden, werden nichtlineare Aktivierungsfunktionen in CNNs integriert. Ein Beispiel für solch eine Aktivierungsfunktion ist die ReLU-Aktivierung (Rectified Linear Unit), die negative Werte auf null abbildet und positive Werte unverändert lässt (BISHOP 2006).

Da tiefere Netze potentiell komplexere Merkmale lernen wird angenommen, dass tiefere Netze bessere oder zumindest genauso gute Ergebnisse liefern wie flache Netze (HE et al. 2016a). Um das Training solcher tiefen Netzwerke zu ermöglichen führen HE et al. (2016a) sogenannte ,Residual Blocks' ein und optimieren diese durch pre-activations (HE et al. 2016b). Der schematische Aufbau eines Residual Blocks mit pre-activations ist links in Abbildung 3 zu sehen. Die Abbildungsfunktion R, welche durch einen Residual Block gelernt wird, ist die Summe des Inputs Xt und der Residualfunktion des Inputs. Dabei geht Xt direkt in Abbildungsfunktion über eine sogenannte ,shortcut connection' ein und die Residualfunktion wird mittels Convolutional Layern modelliert. Vor jedem Gewichtslayer wird eine pre-activation durch eine Batch Normalization (BN) mit anschließender ReLU-Aktivierung realisiert. Durch eine derartige Modellierung der Abbildungsfunktion müssen im Training lediglich die Gewichte der Residualfunktion gelernt werden, was zu besseren Ergebnissen im Vergleich zu genauso tiefen trainierten Netzen ohne Residuallernen führt (HE et al. 2016b).

4.2 Training

Allgemein wird bei dem Training eines CNNs eine Verlustfunktion minimiert. Basierend auf dem Verlustwert werden durch "Backpropagation" die Gradienten der Parameter des Netzes ermittelt. Entsprechend der Gradienten werden die Gewichte iterativ aktualisiert um den Verlustwert zu minimieren. Werden nicht alle Trainingsdaten zugleich, sondern eine zufällige Teilmenge der Daten (Mini-Batch) hierfür genutzt, nennt man dies den stochastischen

Gl. 2

Gradientenabstieg (BISHOP 2006). Im Rahmen dieses Beitrags wird die Softmax Kreuzentropie als Verlustfunktion verwendet. Für die Berechnung des Funktionswertes werden zunächst alle Softmax-Aktivierungen der Klassifikationsschicht ermittelt. Bei insgesamt K zu prädizierenden Klassen gibt es K Netzwerkausgaben. Die Softmax-Aktivierung y_k der k-ten Klasse für die Netzwerkeingabe x berechnet sich zu

$$y_{k}(x, w) = \frac{\exp(w_{k}^{T}\Phi)}{\sum_{j=1}^{K} \exp(w_{j}^{T}\Phi)}$$
Gl. 1

aus den Gewichten w_k und w_j des k-ten beziehungsweise j-ten Ausgabeknotens und den Aktivierungen Φ der vorherigen Schicht im Netzwerk (BISHOP 2006). Dabei kann y_k als die Wahrscheinlichkeit interpretiert werden, dass das Sample x zu der Klasse k gehört. Anschließend berechnet sich der Verlustwert E(w) für alle n betrachteten Samples aus der Kreuzentropie der Softmax-Aktivierungen

$$E(w) = -\sum_{n}\sum_{k} t_{nk} \ln(y_{k}(\mathbf{x}_{n}, \mathbf{w}))$$

wobei $t_{nk} \in \{0, 1\}$ eine Indikatorvariable ist, welche angibt, ob das n-te Sample zur k-ten Klasse gehört (BISHOP 2006).

5 Methodik

Basierend auf den in Abschnitt 4 beschriebenen Grundlagen wird in diesem Abschnitt die Methodik erläutert, welche für die Prädiktion der Entstehungsepoche von Seidenstoffbildern angewendet wird. Diese umfasst die verwendete Netzwerkarchitektur sowie die Vorprozessierung der Daten und die Erweiterung der Softmax Kreuzentropie.

5.1 Netzwerkarchitektur

Das in diesem Beitrag genutzte CNN ist in Abbildung 3 dargestellt. Für die Merkmalsextraktion wird ein Residual Network mit 152 Convolutional Layern (HE et al. 2016b) genutzt. Als Eingabe wird ein Bild mit einer Größe von 224 x 224 Pixeln und drei Kanälen benötigt. Das Netzwerk setzt sich aus 50 Residual Blocks zusammen (siehe Abb. 3). Die Ausgabe der Merkmalsextraktion ist ein 2048-dimensionaler Vektor. Als Klassifikationsteil des Netzes werden 3 voll vernetzte Schichten mit ReLU- bzw. Softmax-Aktivierung verwendet. Für die Prädiktion der Epochen besteht der letzte Layer aus drei Knoten. Auf diese Weise werden die 2048 Merkmale auf die Menge der drei Klassen abgebildet.



Abb. 3: Schematische Darstellung eines Residual Blocks (links) sowie die vollständige Netzwerkarchitektur (rechts).

5.2 Training

Für die im vorangegangenen Abschnitt vorgestellte Architektur wird aufgrund der geringen Menge an Trainingsdaten der Merkmalsextraktionsteil mit den auf dem ILSVRC-2012-CLS ImageNet-Datensatz (RUSSAKOVSKY et al. 2015) trainierten Gewichten des ResNet-152 V2 initialisiert. Diese Gewichte bleiben während des Trainings unverändert und lediglich die Gewichte des Klassifikationsteils werden auf dem Seidenstoffdatensatz trainiert. Bei der Nutzung üblicher Verlustfunktionen im Training muss hierfür eine eindeutige Zuordnung aller Samples zu einer der definierten Klassen gegeben sein, damit sie im Training verwendet werden können. Damit im Rahmen des Trainings auch Samples mit uneindeutigen Klassenlabels verwendet werden können, wird die Verlustfunktion aus Gleichung 2 erweitert. Dafür sei zunächst eine hierarchische Klassenstruktur mit zwei Ebenen von Klassen wie links in Abbildung 4 gegeben. Die untere Ebene besteht aus K Basisklassen $\{C_1, ..., C_k, ..., C_K\}$, welche durch den zu trainierenden Klassifikator prädiziert werden können. Sei hierfür B ≔ {1, ..., k, ..., K} die Menge der Indizes der Basisklassen. Zudem gibt es eine zweite Ebene von L zusammengesetzten Klassen {C₀₁, ..., C₀₁, ..., C_{0L}}, wobei sich jede dieser Klassen aus der Vereinigung einer Teilmenge der Basisklassen zusammensetzt. Die Verlustfunktion E(w) aus Gleichung 2 kann derart erweitert werden, dass die N im Mini-Batch enthaltenen Samples einer Basisklasse k oder einer aus den Basisklassen q zusammengesetzten Klasse Q₁ zugehörig sein können. Die in die Kreuzentropie eingehenden Softmax-Aktivierungen der Basisklassen yk berechnen sich weiterhin gemäß Gleichung 1, wohingegen die Softmax-Aktivierung einer zusammengesetzten Klasse Q1 sich aus der Summe der Softmax-Aktivierungen der zugeordneten Basisklassen y_q ergibt (siehe Abbildung 4, rechts). Folglich lautet die erweiterte hierarchische Softmax Kreuzentropie

$$E(\mathbf{w}) = -\sum_{n=1}^{N} \left(\sum_{k=1}^{K} t_{nk} \ln(\mathbf{y}_{k}(\mathbf{x}_{n}, \mathbf{w})) + \sum_{l=1}^{L} t_{nQl} \ln\left(\sum_{q \in Ql} \mathbf{y}_{q}(\mathbf{x}_{n}, \mathbf{w})\right) \right),$$
GL 3

wobei die Indikatorvariablen t_{nk} und t_{nQl} eins sind, falls das n-te Sample zur k-ten Basisklasse beziehungsweise zur Q_l-ten zusammengesetzten Klasse gehört, und null anderenfalls. Auf diese Weise können die zusammengesetzten Klassen nicht vom gelernten Klassifikator prädiziert werden, jedoch während des Trainings als Positivbeispiele für die jeweils zugeordneten Basisklassen und damit implizit als Negativbeispiele für die nicht zugeordneten Basisklassen fungieren. Für den in Abschnitt 3 vorgestellten Datensatz heißt dies, dass es drei Basisklassen und zwei zusammengesetzte Klassen gibt. Die Basisklassen umfassen jeweils Zeiträume von 50 Jahren und die zusammengesetzten Klassen entsprechend 100 Jahre. Samples, die aufgrund ihrer Epochenbeschreibung einer der zusammengesetzten Klassen zugeordnet werden, können auf diese Weise ins Training einbezogen werden. Ohne die Erweiterung wäre dies nicht möglich gewesen. Ein weiteres Problem beim Trainieren des Netzes stellt neben den uneindeutigen Klassenlabels zum einen die ungleiche Verteilung der Bilder auf die einzelnen Klassen und zum anderen die verhältnismäßig geringe Menge an Trainingsdaten dar. Daher wird als Vorprozessierung eine Augmentierung des hier verwendeten Datensatzes realisiert, um eine repräsentative Grundlage für das Training zu schaffen und eine Überanpassung an die Trainingsdaten zu vermeiden.

Um eine Invarianz gegenüber der Ausrichtung des Objektes gegenüber der Kamera zu schaffen werden die Bilder zunächst um einen zufälligen kleinen Winkel rotiert und anschließend erneut mit einer 50 %-igen Wahrscheinlichkeit um 90°. Eine zusätzliche Variation erfolgt durch zufälliges horizontales und vertikales Spiegeln der Bilder. Anschließend wird ein zufälliger Bereich jedes Bildes ausgeschnitten, der 30 % - 90 % des ursprünglichen Bildes einnimmt. Dieser Ausschnitt wird dann auf die Eingangsgröße des Netzwerkes skaliert. Zusätzlich werden zufällige kleine horizontale und vertikale Scherungen durchgeführt, dies ist für den verwendeten Datensatz sinnvoll, um Deformationen des Musters durch Falten in einigen Stoffen nachzuahmen. Zusätzlich werden Kontrast, Helligkeit, Farbton und Sättigung der Bilder zufällig verändert, da angenommen wird, dass diese durch das Fotografieren der Werke mit unkontrollierten Lichtverhältnissen sowie unterschiedlichem Hintergrund variieren.



Abb. 4: Schematische Darstellung der Klassenstruktur bestehend aus Basisklassen C₁, ..., C_K und zusammengesetzten Klassen (links) sowie dem zugehörigen Ablaufschema der erweiterten Softmax Kreuzentropie (rechts)

6 Experimente

Für das Training der Klassifikationsschicht wird der Datensatz in einen 80 % der Daten umfassenden Trainingsdatensatz sowie einen Validierungs- und einen Testdatensatz, die jeweils 10 % der Daten umfassen, aufgeteilt. Der Validierungsdatensatz, mit dem die optimale Anzahl der Iterationen sowie die Größe der Mini-Batches bestimmt werden variiert im Rahmen der durchgeführten 10-fachen Kreuzvalidierung. Die Parameter der Klassifikationslayer werden mittels dem Adaptive-Momentum (KINGMA & BA 2014) unter Verwendung der Standardparameter optimiert. Dabei wird eine ausgeglichene Klassenverteilung in den Trainingsbatches erzielt, indem für jede Klasse etwa gleich viele Bilder zufällig aus dem Trainingsdatensatz gezogen werden. Hierfür werden für die unterrepräsentierten Klassen mehr synthetische Bilder generiert als für die dominante Klasse "1. Hälfte 20. Jh.", sodass anschließend eine ausgeglichene Verteilung über alle Klassen vorliegt (siehe Tabelle 2).

Klassenbezeichnung	2. Hälfte 19. Jh.	1. Hälfte 20. Jh.	2. Hälfte 20. Jh.
Anzahl originaler Bilder	213	2035	628
Anzahl synthetischer Samples je Bild	964	100	326
Resultierende Anzahl an Samples	205545	205535	205356

Tab. 2: Übersicht über die Anzahl der synthetischen Samples pro Klasse.

Abbildung 5 zeigt exemplarisch ein Originalbild und zugehörige synthetische Variationen.



Abb. 5: Beispiel für ein Bild aus dem (GARÍN) Datensatz (links) und zugehörige synthetische Daten

Die Anzahl der notwendigen Trainingsiterationen wird in jedem Experiment auf Basis des Validierungsdatensatzes ermittelt. Hierfür werden die Gesamtgenauigkeit (Anteil der korrekten Prädiktionen) und die über alle Klassen gemittelten F1-Scores (harmonisches Mittel aus Vollständigkeit und Korrektheit) betrachtet. Ist in dem Verlauf der auf dem Validierungsdatensatz basierenden Qualitätsmaße eine Sättigung zu beobachten, so wird der Testdatensatz evaluiert.

Zunächst wird das Training des Netzes ausschließlich auf Basis der Basisklassen und ohne Datenaugmentierung durchgeführt (BASIS). Somit kann die Prädiktion der Epoche durch den CNN-basierten Klassifikator im Allgemeinen bewertet werden. In den weiterführenden Experimenten wird dann der Einfluss der vorgenommenen Modifikationen evaluiert. Die erste Modifikation ist die Hinzunahme von synthetischen Daten im Training (AUG). Für den Validierungs- und den Testdatensatz werden ausschließlich die originalen Bilder verwendet. Als zweite Modifikation werden statt der synthetischen Samples die Samples der zusammengesetzten Klassen (siehe Tabelle 1) in Kombination mit der erweiterten Softmax Kreuzentropie in (siehe Gleichung 3) hinzugenommen (ERW). Das abschließende Experiment nutzt im Training sowohl synthetische Bilder als auch (ausschließlich originale) Bilder der übergreifenden Klassen (AUG+ERW). Die resultierenden Ergebnisse sind in den Tabelle 3 aufgeführt.

Experiment	F1-Score	F1-Score	F1-Score	Gesamt-	F1-Score
	2. Hälfte 19. Jh.	1. Hälfte 20. Jh.	2. Hälfte 20. Jh.	genauigkeit	(Mittel)
	[%]	[%]	[%]	[%]	[%]
BASIS	29,9	79,3	60,9	70,6	56,7
AUG	19,7	83,2	62,3	74,5	55,1
ERW	32,1	82,2	65,6	74,0	60,0
AUG+ERW	19,0	85,2	64,0	77,1	56,1

Tab. 3: Zusammenstellung der erzielten Ergebnisse in den einzelnen Experimenten.

Basierend auf einer 10-fachen Kreuzvalidierung wurde hinsichtlich der Gesamtgenauigkeit das beste Ergebnis durch die Verwendung zusätzlicher Daten sowohl durch die Datenaugmentierung als auch durch die erweiterte Verlustfunktion erzielt (AUG+ERW) (siehe Tabelle 3). Die Erweiterung des Trainingsdatensatzes mittels nur einer der beiden Methoden führt zu etwa identischen Ergebnissen auf dem Testdatensatz in Bezug auf die Gesamtgenauigkeit.

Da im verwendeten Datensatz jedoch eine unausgeglichene Verteilung der Bilder auf die drei Klassen vorliegt, ist eine Bewertung der Experimente auf Basis klassenspezifischer Indizes sinnvoll. In Tabelle 4 sind die F1-Scores der einzelnen Klassen aufgeführt. Allgemein fällt auf, dass die Klasse "2. Hälfte 19. Jh." die schlechtesten Genauigkeiten erzielt, was sich auf die geringe Menge an Trainingssamples in dieser Klasse (siehe Tabelle 1) zurückführen lässt. Entsprechend sind für die Klasse "1. Hälfte 20. Jh." mit den meisten Trainingsdaten die höchsten F1-Scores zu beobachten.

Bezüglich der Bewertung der einzelnen Experimente kann festgestellt werden, dass durch die Hinzunahme von synthetischen Trainingssamples die Genauigkeit der am stärksten unterrepräsentierten Klasse deutlich geringer wird; die Genauigkeiten der anderen Klassen können jedoch leicht verbessert werden. Als Ursache wird vermutet, dass im verwendeten Datensatz nicht ausreichend Bilder vorhanden sind, um die Charakteristika der Stoffe aus der Zeit von 1850 bis 1900 zu beschreiben.

Dies wird durch die Verbesserung des entsprechenden F1-Scores bei der Einführung von Zusatzinformationen im Training durch die Hierarchische Verlustfunktion bestätigt. Durch die Hinzunahme von 302 Bilder von Seidenstoffen, deren Herstellungsepoche teilweise in die zweite Hälfte des 19. Jahrhunderts fällt, kann der F1-Score um etwa 2 % verbessert werden. Ebenso kann die Genauigkeit der Klasse "2. Hälfte 20. Jh.", für welche ebenfalls verhältnismäßig wenig Trainingsdaten vorliegen, die Berücksichtigung von Samples mit uneindeutigem Klassenlabel im Training gesteigert werden. Somit erreicht das Experiment mit der Hierarchischen Verlustfunktion den besten mittleren F1-Score von 60 %. Werden zusätzlich zu den Bildern mit uneindeutigem Klassenlabel synthetische Samples zum Training der Klassifikationslayer genutzt, wird ein mittlerer F1-Score vergleichbar mit dem im Bezugsexperiment erreicht. Die Klassen mit verhältnismäßig vielen Trainingsbeispielen weisen verbesserte Genauigkeiten auf, wohingegen für die Klasse "2. Hälfte 19. Jh." wie auch bei dem anderen Experiment mit synthetischen Trainingsdaten einen deutlich schlechteren F1-Score erzielt.

Eine klare Aussage darüber, welches Experiment die besten Ergebnisse liefert, kann nicht getroffen werden, da dies von dem gewünschten Ziel abhängt. Um eine höchstmögliche Gesamtgenauigkeit zu erzielen ist die Kombination aus Datenaugmentierung und erweiterter Verlustfunktion die beste Wahl unter den untersuchten Ansätzen. Ist ein hoher F1-Score der stark

unterrepräsentierten Klasse erwünscht, liefert die erweiterte Verlustfunktion ohne Augmentierung die besten Ergebnisse.

7 Zusammenfassung und Ausblick

In diesem Beitrag wurde zum Umgang mit uneindeutigen Klassenlabels eine erweiterte Verlustfunktion eingeführt, welche Samples mit derartigen Labeln im Training berücksichtigen können. Dies führte in den Experimenten zu einer Verbesserung der klassenspezifischen Genauigkeiten, insbesondere der unterrepräsentierten Klasse. Zudem wurden zur Vergrößerung des Datensatzes und zum Umgang mit einer ungleichen Verteilung der Daten auf die Klassen synthetische Trainingsdaten generiert, wobei die Anzahl der Datenaugmentierungen pro Sample derart gewählt wurde, dass im Trainingsdatensatz eine ausgeglichene Klassenverteilung vorliegt. Die Ergebnisse zeigen, dass die am stärksten unterrepräsentierte Klasse dadurch schlechter prädiziert wird. Da als Ursache die geringe Menge an Daten für die Klasse "2. Hälfte 19. Jh." vermutet wird, wäre in künftigen Untersuchungen eine Erweiterung des Datensatzes durch existierende Sammlungen von Museen von Interesse. Zum einen wären die einzelnen Klassen vermutlich besser durch die Daten repräsentiert und zum anderen könnten durch die größere Menge an Daten einige vortraininerte Layer der Merkmalsextraktion für die neue Aufgabe adaptiert werden. Außerdem könnten im Rahmen von "Multi-Task Learning" Informationen über den Hersteller oder den Herstellungsort für die Prädiktion der Herstellungsepoche einen Mehrwert liefern. Interessant wäre auch die Anwendung der Hierarchischen Verlustfunktion auf andere Datensätze und deren Einfluss auf die Klassifikationsergebnisse. Denkbar wären semantische Hierarchien von Pflanzen- oder Tierbezeichnungen. Zudem wäre in weiterführenden Arbeiten die Entwicklung eines Klassifikators von Interesse, der automatisch eine zusammengesetzte Klasse prädiziert, falls die Prädiktion der Basisklassen aufgrund von fehlenden Informationen nicht möglich ist.

8 Danksagung

The research leading to these results is in the frame of the "SILKNOW. Silk heritage in the Knowledge Society: from punched cards to big data, deep learning and visual/tangible simulations" project, which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 769504.

All image data as well as the according descriptions were used in this paper's experiments with the friendly permission of Garín 1820 (http://garin1820.com/).

9 Literaturverzeichnis

- AYUB, R., ORBAN, C. & MUKUND, V., 2017: Art Appraisal Using Convolutional Neural Networks. http://cs229.stanford.edu/proj2017/final-reports/5229686.pdf
- BAR, Y., LEVY, N. & WOLF, L., 2014: Classification of artistic styles using binarized features derived from a Deep Neural Network. Workshop at the European Conference on Computer Vision, Springer, Cham, 71-84.

- BLESSING, A. & WEN, K., 2010. Using machine learning for identification of art paintings. Technical Report, Stanford University, 5 pages. https://pdfs.semanticscholar.org/1d73/0a452a5c03cc23f90d4fde71c08864f31c35.pdf
- BISHOP, C., 2006: Pattern Recognition and Machine Learning. 1st edition, Springer, New York, 235-240.
- DENG, J., DONG, W., SOCHER, R., LI, L.-J., LI, K. & FEI-FEI, L., 2009: ImageNet: A large-scale hierarchical image database. IEEE Conference on Computer Vision and Pattern Recognition, 248-255.
- GARÍN: Image data and according descriptions with friendly permission of Garín 1820. URL: http://garin1820.com/
- HENTSCHEL, C., WIRADARMA, T. & SACK, H., 2016: Fine tuning CNNS with scarce training data
 Adapting imagenet to art epoch classification. IEEE International Conference on Image Processing (ICIP), 3693-3697.
- HE, K., ZHANG, X., REN, S. & SUN, J., 2016a: Deep residual learning for image recognition. IEEE conference on computer vision and pattern recognition, 770-778.
- HE, K., ZHANG, X., REN, S. & SUN, J., 2016b: Identity mappings in deep residual networks. European conference on computer vision, Springer, Cham, 630-645.
- HICSONMEZ, S., SAMET, N., SENER, F. & DUYGULU, P., 2017: DRAW: Deep networks for Recognizing styles of Artists Who illustrate children's books. ACM on International Conference on Multimedia Retrieval, 338-346.
- KINGMA, D. P. & BA, J. L. 2014: Adam: A method for stochastic optimization. arXiv preprint, arXiv:1412.6980.
- KRIZHEVSKY, A., SUTSKEVER, I. & HINTON, G. E., 2012: ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25(1), 1097-1105.
- RAZAVIAN, A., AZIZPOUR, H., SULLIVAN, J. & CARLSSON, S., 2014: CNN Features Off-the-Shelf: An Astounding Baseline for Recognition. IEEE conference on computer vision and pattern recognition workshops, 806-813.
- RUSSAKOVSKY, O., DENG, J., SU, H., KRAUSE, J., SATHEESH, S., MA, S., HUANG, Z., KARPATHY, A., KHOSLA, A., BERNSTEIN, M., BERG, A. C. & FEI-FEI, L., 2015: ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, **115**(3), 211-252.
- SALEH, B. & ELGAMMAL, A., 2016: Large-scale classification of fine-art paintings: Learning the right metric on the right feature. International Journal for Digital Art History, 70-93.
- SIMONYAN, K. & ZISSERMAN, A., 2014: Very deep convolutional networks for large-scale image recognition. arXiv preprint, arXiv:1409.1556.
- SUR, D. & BLAINE, E., 2017: Cross-Depiction Transfer Learning for Art Classification. Technical report for CS 231A and CS 231N at Stanford.
- TAN, W. R., CHAN, C. S., AGUIRRE, H. E. & TANAKA, K., 2016: Ceci n'est pas une pipe: A deep convolutional network for fine-art paintings classification. IEEE International Conference on Image Processing (ICIP), 3703-3707.
- VURAL, V. & DY J. G., 2016: A hierarchical method for multi-class support vector machines. In: Proceedings of the 21st International Conference on Machine Learning, ACM, 105.

- WU, Z. & SAITO, S., 2017: HiNet: Hierarchical Classification with Neural Network. 5th International Conference on Learning Representations (ICLR), arXiv preprint arXiv:1705.11105.
- XIAO, Z., LIU, X., WU, J., GENG, L., SUN, Y., ZHANG, F. & TONG, J., 2018: Knitted fabric structure recognition based on deep learning. The Journal of the Textile Institute, 1-7.
- YOSINSKI, J., CLUNE, J., BENGIO, Y. & LIPSON, H., 2014: How transferable are features in deep neural networks? Advances in Neural Information Processing Systems 27, 3320-3328.

Landesweite digitale Vegetationshöhenmodelle aus historischen SW - Stereoluftbildern

CHRISTIAN GINZLER¹, MAURO MARTY¹ & LARS T. WASER¹

Zusammenfassung: Methoden der Fernerkundung erlauben die Erfassung von 3D Strukturen von Wäldern über grosse Gebiete. Landesweite Berechnungen von Oberflächenmodellen mit photogrammetrischen Verfahren aus digitalen Stereoluftbildern oder ALS Daten werden in einigen Ländern bereits operationell eingesetzt. Die Verfügbarkeit historischer Stereoluftbilder erlaubt die Berechnung digitaler Oberflächenmodelle aus der Vergangenheit mittels photogrammetrischer Methoden. Wir stellen einen Arbeitsfluss vor, mit welchem wir für die Schweiz landesweite Oberflächenmodelle für die 1980er und 1990er Jahre berechnet haben. Es wurden zwei verschiedene Bildmatching Verfahren getestet und die Ergebnisse hinsichtlich der Vollständigkeit der Bildkorrelation verglichen. Die Ergebnisse zeigen, dass beide Verfahren trotz der zum Teil geringen Kontraste in den digitalisierten schwarz-weissen Luftbildern zufriedenstellende Ergebnisse erzielten. Auf offenen Wiesen und im Gebiet der Landwirtschaft war die Vollständigkeit gering. In Waldgebieten jedoch deutlich höher.

1 Einleitung

Die Erfassung von horizontalen und vertikalen Strukturen in Wäldern erlaubt die Modellierung einer weiten Palette von Waldfunktionen. So können Ressourcen, wie zum Beispiel Holzvorrat und Biomasse, mit der Höhe von Waldbeständen geschätzt werden. Für die Schutzfunktion der Wälder gegenüber Naturgefahren, wie zum Beispiel Steinschlag und Lawinen, sind die Variabilität der Bestandeshöhen und Bestandeslücken entscheidende Parameter. Kronenschluss, Lückigkeit und Bestandesdichte, sowie die Stufigkeit von Waldrändern, sind wichtige Eingangsgrössen für die Modellierung von Biodiversität. Die Erholungsfunktion für den Menschen wird unter anderem von Lichtverfügbarkeit und Abwechslung in den Waldstrukturen beeinflusst.

Methoden aus der Fernerkundung erlauben heutzutage die Erfassung von 3D Strukturen von Wäldern über grosse Gebiete. Sowohl aktive, wie z. Bsp. Airborne Laserscanning (ALS) und Synthetic Aperture Radar (SAR), als auch passive Fernerkundungssysteme, wie multispektrale Sensoren haben sich in den letzten Jahren etabliert um 3D und 2.5D Daten der Erdoberfläche präzise zu erfassen. Landesweite Berechnungen von Oberflächenmodellen mit photogrammetrischen Verfahren aus digitalen Stereoluftbildern werden in einigen Ländern (z.B. Schweiz, Österreich, manche Bundesländer in Deutschland) bereits operationell eingesetzt.

Die Verfügbarkeit historischer Stereoluftbilder erlaubt die Berechnung digitaler Oberflächenmodelle aus der Vergangenheit mittels photogrammetrischer Methoden. In unserem Artikel stellen wir einen Arbeitsfluss vor, mit welchem wir für die Schweiz landesweite Oberflächenmodelle aus den 1980er Jahren berechnet haben. Zudem liegen aktuelle schweizweite Oberflächenmodelle aus dem Landesforstinventar (LFI) Schweiz vor.

¹ Eidg. Forschungsanstalt für Wald Schnee und Landschaft WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, E-Mail: [Christian.Ginzler, Mauro.Marty, Lars.Waser]@wsl.ch

2 Material und Methoden

Im Rahmen der Arealstatistik Schweiz wurden die analogen SW-Stereoluftbilder der landesweiten Befliegung vom Bundesamt für Landestopographie (swisstopo) der Jahre 1979-85 mit dem Leica DSW700 Scanner mit einer Pixelgrösse von 14 µm gescannt. Die Luftbilder wurden mit der Kamera RC10, mit einem mittleren Massstab von ~ 1:25'000 aufgenommen.

Die Maschenweite der Pixel entspricht somit ~0.35 m. Die Bildorientierungen wurden von swisstopo durchgeführt und für diese Studie übernommen. Der hauptsächliche Einsatz der Bilder war die Aktualisierung der Bodenbedeckung und Bodennutzung der Arealstatistik durch stereo Bildinterpretation am Bundesamt für Statistik (BFS 2018). Bei der Bildorientierung wurde speziell auf die Genauigkeit in der Lage geachtet. Die Genauigkeit in der Höhe war für den primären Einsatz der Bilder von untergeordneter Bedeutung und liegt im Bereich von +/- 1 – 5 Meter. Für den Vergleich von zwei unterschiedlichen Bildmatching Verfahren wurden sechs repräsentative Gebiete in der Schweiz mit unterschiedlicher Landnutzung und Topographie ausgewählt (Abbildung 1).



Abb. 1: Sechs Testgebiete in der Schweiz, welche die unterschiedlichen Landschaften und Topografien der Schweiz repräsentieren.

3 Verwendete Bildmatching Verfahren

In der vorliegenden Studie wurden zwei unterschiedliche Softwarepakete für die Bildkorrelation getestet. Es waren dies SocetSet 5.6.0 und Match-T von Trimble Photogrammetry 7.1.

3.1 SocetSet 5.6.0

Die Berechnung der Punktewolke im Softwarepaket SocetSet 5.6.0 von BAE Systems wurde mit dem Modul "Next Generation Automatic Terrain Extraction" (NGATE) durchgeführt. Die Bildkorrelation erfolgt dabei auf verschiedenen Pyramidenstufen bis auf die 1:1 Pixelebene und verwendet eine Kombination aus flächenbasierter und kantenbasierter Korrelation. Obwohl NGATE die Korrelation bis auf Pixelebene durchführt, können die berechneten Werte aus SocetSet nur in einem regelmässigen Raster ausgegeben werden. Die Interpolationsmethode welche die korrelierten, unregelmässigen Punkte in ein regelmässiges Raster umwandelt ist nicht dokumentiert. Beim Export der Daten kann zu jedem Punkt ein Merkmal ausgegeben werden, welches angibt, ob der Punkt korreliert oder interpoliert wurde. Die Maschenweite für das endgültige Raster wurde auf 1x1 m festgesetzt.

Die Verarbeitung der Stereomodelle erfolgte ähnlich wie in GINZLER & HOBI (2015). Für jedes Luftbild wurde die Bildabdeckung im Gelände berechnet. Aus benachbarten Luftbildern derselben Fluglinien konnten somit Stereoabdeckungen abgeleitet werden. Die Bildkorrelation wurde für jedes Stereomodell separat durchgeführt. Stereomodelle aus den Querüberlappungen der Flugstreifen wurden dabei nicht berücksichtigt. Die Bildkorrelationen wurden für jedes einzelne Stereomodell unabhängig voneinander prozessiert und konnten somit auf verteilte Prozesse ausgelagert werden. Die Steuerung der Berechnungen erfolgte dabei über eine zentrale Datenbank. Auf die zentral gespeicherten Luftbilder konnte über das Netzwerk zugegriffen werden.

Die exportierten regelmässigen Punkte wurden schliesslich in ein digitales Oberflächenmodell (DOM) umgewandelt.

3.2 Match-T

Die Punktewolken wurden mit Match-T des Softwarepaketes Trimble Photogrammetry 7.1 berechnet. Match-T verwendet dabei drei verschiedene Bildkorrelations Algorithmen, "Feature Based" (FGM), "Least Squares" (LSM) und einen "Cost Based" (CBM) Ansatz, einem Semi Global Matching Algorithmus (HIRSCHMÜLLER 2008). Diese Algorithmen werden hierarchisch durch alle Bildpyramiden angewendet, bis auf die 1:1 Pixelebene. Match-T verwendet sequentielles Muli-Matching. Dabei werden basierend auf den Blickwinkeln die am besten geeigneten Bildpaare für eine Berechnungseinheit ausgewählt. Die einzelnen Punktewolken werden anschliessend zusammengeführt und robust gefiltert. Für die Bildkorrelation der gescannten schwarz-weiss Stereobildpaare wurden die Standardeinstellungen der Software gewählt. Dabei werden für die ersten Bildpyramiden das FBM und für die letzten drei Bildpyramiden das CBM verwendet. Zur Generierung der DOMs haben sowohl gewünschte Punktedichte als auch ein sogenannter "Optimize" Parameter einen wichtigen Einfluss auf die angewandte Filtermethode für die Endpunktwolke. Die Filterung wurde vom Standard 2.5D auf den 3D-Modus umgestellt, so dass nur erfolgreich korrelierte Punkte als unregelmäßige Punktwolke exportiert werden. Die Punktdichte wurde auf den Wert 3

gesetzt, um eine Einstellung anzuwenden, die am ehesten mit der für NGATE verwendeten vergleichbar ist. Die irregulären Punktewolken wurden zu einem DOM mit der Auflösung 1 m \times 1 m linear interpoliert.

4 Qualitätsabschätzung

Ein geeignetes Mass für die Qualitätsabschätzung der berechneten Oberflächenmodelle ist die Vollständigkeit der effektiv korrelierten Punkte pro Flächeneinheit. Diese ergibt sich aus dem Verhältnis der korrelierten Punkte zur maximal möglichen Anzahl Punkte pro Flächeneinheit. Um die Vollständigkeit für die unterschiedlichen Bodenbedeckungen berechnen zu können wurde der Datensatz der Arealstatistik Schweiz verwendet. Bei der Arealstatistik Schweiz wird auf einem Stichprobennetz von 100 m × 100 m die Bodenbedeckung und Bodennutzung erhoben. Für sechs Klassen der Bodenbedeckung (geschlossener Wald, Grasland, Gebüsche, versiegelter Boden, Fels und Offenboden, Gletscher) wurden um den Stichprobenpunkt in einem Radius von 5 m die Anzahl der korrelierten Punkte für beide Korrelationsansätze gezählt. Die maximal mögliche Anzahl beträgt dabei 80 Punkte. Anders als bei Match-T, wo durch das sequentielle Multimatching mehr als 1 Punkt pro Quadratmeter entstehen können. Damit die Werte mit NGATE verglichen werden konnten wurde pro Quadratmeter aber dann nur 1 Punkt gezählt.

5 Resultate

5.1 Digitale Oberflächenmodelle

Die Generierung von digitalen Oberflächenmodellen war mit beiden Bildmatching-Verfahren erfolgreich. Durch die Subtraktion eines digitalen Geländemodells wurden beide DOMs normalisiert (nDOM). Ein Ausschnitt eines typischen Waldgebietes aus dem nördlichen Schweizer Mittelland ist in Abbildung 2 dargestellt. Die Waldstrukturen werden mit beiden Methoden gut abgebildet. Sowohl grössere Waldlücken als auch schmale Waldstreifen wurden gut modelliert. Ausschnitte aus vier Testgebieten geben einen Eindruck der erfolgreichen Modellierung der Oberflächen aus den digitalisierten schwarz-weiss Luftbildern (Abbildung 3).



Abb.2: Typisches Waldgebiet im Schweizer Mittelland. Links: Schwarz-weiss Luftbild. Mitte: Reliefdarstellung des digitalen Oberflächenmodells und des normalisierten Oberflächenmodells aus NGATE. Rechts: Reliefdarstellung des digitalen Oberflächenmodells und des normalisierten Oberflächenmodells aus Match-T.



Abb. 3: Orthofotos und die Reliefdarstellungen der entsprechenden DOMs mit eingefärbten nDOMs in verschiedenen Testregionen mit unterschiedlichen Waldstrukturen (von oben nach unten Ausschnitten aus den Testgebieten 1,3,5,6)

5.2 Vollständigkeit

Die Vollständigkeit wurde für beide Korrelationsresultate berechnet und für die sechs unterschiedlichen Bodenbedeckungsklassen analysiert (Tabelle 1). NGATE erreichte in allen Bodenbedeckungsklassen höhere Vollständigkeiten. Der grösste Unterschied ergibt sich im geschlossenen Wald mit 83 % Vollständigkeit bei NGATE und 40 % Vollständigkeit bei Match-T. Bei allen anderen Bodenbedeckungsklassen war die Vollständigkeit bei NGATE um 10 % - 20 % höher.

Tab. 1: Mittlere Vollständigkeit und Anzahl der Referenzpunkte für sechs verschiedene Bodenbedeckungsklassen in den Untersuchungsgebieten.

Bodenbedeckung	Mittlere Vollständigkeit [%]		Anzahl Referenzpunkte
-	NGATE	Match-T	
Geschlossener Wald	83	43	34'805
Grasland	68	51	36'992
Gebüsche	92	68	6'915
Versiegelter Boden	94	66	3'000
Fels, Offenboden	91	70	18'973
Gletscher	66	54	6'093

Die geringste Vollständigkeit wurde für die Bodenbedeckung Grasland mit 68 % (NGATE) und 51 % (Match-T) erzielt. Unterschiede fanden sich bei dieser Bodenbedeckung in den verschiedenen Testgebieten. Im Flachland konnten deutlich weniger Punkte korreliert werden als in den Bergregionen (Abbildung 4).



Abb. 4: Vollständigkeit der Bildkorrelation für die Bodenbedeckungsklasse "Grasland" im Flachland (links) und in Berggebieten (rechts).

6 Fazit & Ausblick

Die Verfügbarkeit von orientierten historischen Luftbildern ist eine sehr wertvolle Datenquelle zur Berechnung von digitalen Oberflächenmodellen aus der Vergangenheit. Obwohl die Flugparameter nicht für eine dichte Bildkorrelation ausgelegt waren, können durch paarweises Image-Matching sehr grosse Gebiete nahezu lückenlos korreliert werden. Die zum Teil sehr geringe Textur in den digitalisierten schwarz-weissen Luftbildern führt speziell im Offenland, auf Wiesen und Feldern zu einer geringen Vollständigkeit. Oft werden nur an Nutzungsgrenzen und Strassenrändern Punkte erfolgreich korreliert. Auf Waldflächen werden jedoch deutlich höhere, zufriedenstellende Vollständigkeiten erzielt. Die Wahl der Methode für die Bildkorrelation ist nicht entscheidend. Auch wenn in dieser Studie mit NGATE eine höhere Vollständigkeit der Korrelationspunkte erreicht wurde, so waren die abgeleiteten Oberflächenmodelle von NGATE und Match-T durchaus vergleichbar.

Neben dem vorliegenden landesweiten Datensatz aus der Luftbildbefliegung 1979-85 wird zudem ein analoger Datensatz mit den Luftbildern 1990-98 gerechnet. Somit stehen nebst den spektralen Zeitreihen seit den 1980er Jahren zusätzlich auch Zeitreihen zur Oberfläche und Struktur von Wäldern zur Verfügung. Diese werden in kommenden Arbeiten hinsichtlich der Dynamik und Störungen in Wäldern über die gesamte Schweiz analysiert.

7 Literaturverzeichnis

BFS - BUNDESAMT FÜR STATISTIK, 2018: Arealstatistik: http://www.landuse-stat.admin.ch

- GINZLER, C. & HOBI, M., 2015: Countrywide stereo-image matching for updating digital surface models in the framework of the Swiss National Forest Inventory. Remote Sensing, 7(4), 4343-4370.
- HIRSCHMÜLLER, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. IEEE Transactions on Pattern Analysis and Machine Intelligence, **30**, 328-341.

Detektion und Extraktion von Gewässersohlenpunkten aus Laserbathymetriedaten unter Nutzung von Full-Waveform-Stacking

DAVID MADER¹, KATJA RICHTER¹, PATRICK WESTFELD², ROBERT WEIß³ & HANS-GERD MAAS¹

Zusammenfassung: Laserbathymetrie erlaubt eine effiziente und flächenhafte Erfassung einer Gewässersohle. Die Messmethode wird jedoch durch die Gewässertrübung stark limitiert, sodass eine zuverlässige Detektion von Gewässersohlenpunkten bei stärkerer Trübung oder in tieferen Bereichen nicht mehr möglich ist. Dies führt zu einer unvollständigen Erfassung der Gewässersohle. In diesem Beitrag werden erweiterte Auswertemethoden vorgestellt, welche im Vergleich zu Ergebnissen einer Standardprozessierung die Eindringtiefe erhöhen und die zuverlässige Extraktion und Detektion von Gewässersohlenpunkten in tieferen Gewässerbereichen ermöglicht. Die Methodik basiert auf der Analyse von korrelierten Nachbarschaftsinformationen unter Annahme einer stetigen Gewässersohle. Die Ergebnisse zeigen eine signifikante Erhöhung der Eindringtiefe bei einer hohen Zuverlässigkeit der zusätzlich extrahierten Gewässersohlenpunkte.

1 Einleitung

Zu den Aufgaben der Wasserstraßen- und Schifffahrtsverwaltung gehört ein aktives Wasserstraßenmanagement, das eine optimale Wasserbewirtschaftung sicherstellt. Dabei müssen die Auswirkungen bautechnischer Veränderungen wie Buhnenabsenkungen, Buhnenfeldberäumungen und Streichlinienaufweitung auf das Gewässerbett dokumentiert werden. Moderne Laserbathymetriesysteme (Airborne LiDAR Bathymetry, ALB) sind in der Lage, die Geometrie von Wasserwechselzonen und anderen Flachwasserbereichen effizient zu erfassen. Im Vergleich zu konventionellen hydroakustischen Messverfahren wie Echolot- und Fächerlotmessungen kann dabei die Wirtschaftlichkeit deutlich erhöht werden. Das Potential des Messverfahrens wird allerdings durch die Gewässertrübung stark limitiert. Bei stärkerer Gewässertrübung oder in tieferen Bereichen kann der Gewässerboden nicht mehr zuverlässig detektiert werden.

Die Abbildungen 1 und 2 verdeutlichen die Problematik am Beispiel der Gewässerschle der Bundeswasserstraße Elbe, welche zur Erfüllung der oben genannten Aufgaben möglichst lückenlos zu erfassen ist. Abbildung 1 zeigt den Bereich der Fahrrinne, der mit konventionellen Echolotsensoren vermessen wurde. Die Echolot-Messungen sind dabei Messprinzip-bedingt auf die Fahrrinne begrenzt. In Abbildung 2 sind die Ergebnisse der Datenerfassung mit einem Laserbathymetriesystem visualisiert. Zur Auswertung der Full-Waveform Daten wurden Standardprozessierungsmethoden (online waveform processing) genutzt. Da die Elbe eine hohe Gewässertrübung aufweist,

¹ Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Helmholtzstraße 10, D-01069 Dresden, E-Mail: [david.mader, katja.richter1, hans-gerd.maas]@tu-dresden.de

² Bundesamt für Seeschifffahrt und Hydrographie, Neptunallee 5, D-18057 Rostock, E-Mail: patrick.westfeld@bsh.de

³ Bundesanstalt für Gewässerkunde, Am Mainzer Tor 1, D-56068 Koblenz

konnte der Gewässerboden im Flachwasserbereich auf diese Weise nicht vollständig detektiert werden.

Ziel des Beitrags ist die Entwicklung von erweiterten Auswertealgorithmen, welche im Kern auf einer Analyse von korrelierten Nachbarschaftsinformationen und der Annahme einer stetigen Gewässersohle basieren. Dadurch soll eine höhere Eindringtiefe, eine höhere Datendichte und damit einhergehend eine größere Gewässersohlenabdeckung mit einer hohen Zuverlässigkeit erreicht werden. In diesem Beitrag werden die Methodik und die Ergebnisse vorgestellt, diskutiert und abschließend Ziele für weiterführende Arbeiten genannt.





Abb. 1: Abschnitt der Elbe bei Elster (Elbe). Die aus Echolotmessungen des Vermessungsschiffes ROSSLAU extrahierten Gewässersohlenpunkte der Fahrrinne sind farbcodiert dargestellt. (Quelle: Bundesanstalt für Gewässerkunde)



Abb. 2: Abschnitt der Elbe bei Elster (Elbe). Gewässersohlenpunkte aus der Online-Prozessierung des RIEGL VQ880-G sind farbcodiert dargestellt. (Quelle: Bundesanstalt für Gewässerkunde)

2 Untersuchungsgebiet und Datengrundlage

Das Untersuchungsgebiet umfasst die Elbe zwischen Klöden und Elster. Zur Evaluierung der entwickelten Methoden wurden drei Teilgebiete ausgewählt, die in Abbildung 3a dargestellt sind. Die Befliegung des Gebietes fand im Frühjahr 2015 mit einem RIEGL VQ-880-G statt. Während der Messungen wurde das zurückgestreute Signal am Empfänger in Form einer Full-Waveform digitalisiert und aufgezeichnet (RIEGL 2017). Die aufgenommenen Full-Waveforms bestehen aus 60 bis 200 Samples (Intensitätswerte) mit einem konstanten Abtastintervall von 5,75·10⁻¹⁰ Sekunden. Es wurde dabei eine Datendichte für Gewässersohlenpunkte von bis zu 35 Punkten/m² erreicht. Die zu prozessierenden Gebiete umfassen das Gewässer und den angrenzenden trockenen Uferstreifen mit einer Breite von bis zu 20 m.

Für die Entwicklung, Ausführung und Auswertung der erweiterten Methodik kamen die Full-Waveform-Daten (Abb. 3b), die Trajektorie (Abb. 3c) und die aus Standardprozessierung resultierende (klassifizierte) Punktwolke (Abb. 3d) zur Anwendung. Die Evaluierung der Ergebnisse erfolgte auf Grundlage der mit Standardmethoden extrahierten Gewässersohlenpunkte (Abb. 3e) sowie der Echolotdaten der Fahrrinne (Abb. 3f), welche vom Vermessungsschiff ROSSLAU im Frühjahr 2014 aufgenommen wurden.



Abb. 3: Lage der Evaluationsgebiete Elbe Nord (a1), Elbe Süd (a2) und Elbe Ost (a3). Für die erweiterte Prozessierung sind Full-Wavefrom-Daten (b), Position des Laserbathymetriesystems (c) und mit Standardverfahren extrahierte Wasseroberflächen- sowie Landpunkte nötig (d). Für die Evaluierung der Ergebnisse werden bereits detektierte Gewässersohlenpunkte (e) sowie Echolotdaten (f) verwendet.

3 Methodik

3.1 Allgemeiner Ablauf der Prozessierung

Das Potential der Auswerteverfahren für die ALB-Daten ist stark von der Trübung des Gewässers abhängig (GUENTHER et al., 1978). D.h., je trüber ein Gewässer ist, desto schwächer wird das zurückkommende Bodensignal bei größer werdenden Gewässertiefe sein – mit der Folge, dass Bodenpunkte nicht mehr zuverlässig von Rauschsignalen separierbar sind. Durch die Nutzung von korrelierten Nachbarschaftsinformationen sollen Lage und Weite des Bereichs im Full-Waveform-Signal bestimmt werden, wo das Bodenecho erwartet wird. Damit stellt eine geeignete Nachbarschaftsanalyse die Grundlage für eine zuverlässige Detektion und Extraktion von Gewässerbodenpunkte in der Prozessierung dar.

Die Bearbeitung der Daten erfolgt in vier Prozessierungsschritte (Abb. 4). In einem ersten Schritt wird das Untersuchungsgebiet in Rasterzellen unterteilt (Kap. 3.2), welche im weiteren Verlauf sequentiell abgearbeitet werden. Anschließend werden die Informationen zur Gewässertiefe von benachbarten Rasterzellen erfasst, um eine grobe Näherung der Gewässertiefe für die zu bearbeitende Rasterzelle zu ermitteln (Kap. 3.3). Durch Anwendung eines Full-Waveform-Stacking-Ansatzes (Zusammenfassung von mehreren Full-Waveforms zu einer Gesamtwaveform) über alle Full-Waveforms innerhalb einer Rasterzelle wird der Suchbereich für das (schwache) Bodensignal weiter eingegrenzt (Kap. 3.4). Die durchgeführte Nachbarschaftsanalyse bildet die Grundlage für die abschließende Bestimmung der Gewässersohlenpunkte aus den jeweiligen Full-Waveforms (Kap. 3.5).



Abb. 4: Workflow für die Detektion und Extraktion von Gewässersohlenpunkten auf Basis redundanter Nachbarschaftsinformationen.

3.2 Rasterung des der Messdaten

Die Grundidee des Auswerteverfahrens beruht auf der Analyse und Nutzung korrelierter Nachbarschaftsinformationen. Ausgehend davon, dass benachbarte Full-Waveforms ähnliche Eigenschaften (z.B. Gewässerbodenpunkte) besitzen, können durch eine gemeinsame Auswertung die Gewässertiefe zuverlässiger bestimmt und Fehler im Signalverlauf besser identifiziert werden. Dazu ist es notwendig, eine geeignete Definition von Nachbarschaft bzgl. Anzahl, räumlicher Entfernung, und Verteilung von potentiell benachbarten Messdaten (bestehend aus Full-Waveform, Gewässeroberflächenpunkt und Ursprung der Pulsemission) festzulegen.

Die vorliegende Messkampagne weist im Allgemeinen eine gleichmäßige Punktverteilung und eine hohe Punktdichte mit bis zu 35 Punkten pro Quadratmeter auf. Für die erste Entwicklung der erweiterten Auswerteverfahren kam daher zunächst die leicht zu implementierende Methode der Rasterung zur Anwendung (Abb. 5). Für die Bestimmung der Rasterzellengröße spielt die Punktdichte und die Gewässerbodentopographie eine wichtige Rolle. Mit zunehmender Rasterzellengröße steigt die Anzahl an Full-Waveforms und damit die Zuverlässigkeit des Ergebnisses. Andererseits sinkt die Korrelation zwischen den Messdaten mit zunehmendem Abstand zueinander. Dies kann ähnlich wie bei einer Tiefpassfilterung zu einer Art Verschmierungseffekt führen, wodurch Gewässerbodensignale beim Full-Waveform-Stacking nicht korrekt oder gar nicht mehr identifiziert werden können. Für den hier vorgestellten Fall ergaben empirische Untersuchungen eine geeignete Rasterzellengröße zwischen 2,5 m × 2,5 m bis 3 m × 3 m und daraus resultierend 218 bis 315 Messdaten pro Rasterzelle.



Abb. 5: Aufteilung der Messdaten in Rasterzellen.

Nach der Definition des Rasters erfolgt die auf der räumlichen Lage basierende Zuordnung der Messdaten zu den Rasterzellen. Folglich enthalten die Rasterzellen Oberflächenpunkte aus der Standardprozessierung (Abb. 3d) sowie die dazugehörigen Full-Waveforms (Abb. 3b) und Laserscannerpositionen zum Zeitpunkt der Pulsemission (Abb. 3c). Zusätzlich erfolgt auf Grundlage der Oberflächenpunkte eine Klassifizierung der Rasterzellen in Land- und Wasser-Rasterzellen. Eine grobe Näherung der Gewässertiefe basierend auf den Informationen der Nachbarrasterzellen (Kap 3.3) sowie eine stacked Full-Waveform (Kap 3.4) werden im Verlauf der Prozessierung ermittelt, der jeweiligen Rasterzelle zugeordnet und für die Detektion des Bodensignals verwendet. Die verwendete Methode zur Rasterung der Daten erwies sich als zweckmäßig für die hier vorgestellte Messkampagne. Dennoch ist die empirische Bestimmung der Rasterzellengröße nicht optimal und bei der Auswertung von Gewässersohlen unterschiedlicher Charakteristik auch nicht praktikabel. Eine automatische Definition der Rasterzellengröße basierend auf den Eigenschaften der Messdaten ist aufgrund des großen Einflusses der unbekannten Gewässerbodentopographie nicht trivial, sodass weitere Möglichkeiten einer räumlichen Aufteilung der Daten untersucht werden müssen. Dabei können folgende Varianten in Betracht gezogen werden:

- eine flexible Anpassung der Rasterzellengröße während der Prozessierung, in Abhängigkeit der Gewässertiefe und gegebenenfalls des Trübungsgrades; geringe Trübung und geringe Gewässertiefen führen zu kleineren Rasterzellen und umgekehrt
- die Ermittlung einer vorgegebenen Anzahl der räumlich nächstgelegenen Messdaten für jede zu prozessierende Full-Waveform
- die Einbeziehung aller Messdaten, die eine maximale Lage-Distanz aufweisen

3.3 Nachbarschaftsinformation zwischen Rasterzellen

Die Prozessierung der Messdaten einer Rasterzelle (Kap. 3.4) kann zu mehreren möglichen Gewässertiefen als Ergebnis führen, wodurch eine fehlerhafte Bestimmung der Gewässertiefe nicht ausgeschlossen ist. Um dem entgegenzuwirken, wird vor der Prozessierung der einzelnen Rasterzellen näherungsweise die wahrscheinliche Gewässertiefe der Rasterzelle unter Einbeziehung von Informationen über die Nachbarschaft der umliegenden Rasterzellen bestimmt.

Dazu wird zuerst für jede Wasser-Rasterzelle geprüft, wie viele direkt benachbarte Rasterzellen bereits über Gewässertiefeninformationen verfügen. Weist eine Wasser-Rasterzelle zwei oder mehr Nachbarn mit bekannten bzw. prozessierten Gewässertiefeninformationen auf, erfolgt eine Eingrenzung der möglichen Gewässertiefe für diese Rasterzelle, und es wird anschließend ein Full-Waveform-Stacking mit allen inkludierten Full-Waveform-Daten durchgeführt (Kap. 3.4). Gefälle oder Anstiege der Gewässersohle werden durch einen definierbaren Toleranzbereich bzgl. der zulässigen Gewässertiefe berücksichtigt.

Anfänglich werden keine Wasser-Rasterzellen mit Informationen über Gewässertiefen vorliegen, sodass die Rasterzellen am Ufer als erstes prozessiert werden. Die Gewässertiefe der Land-Rasterzellen wird für die Berechnungen mit 0,0m festgesetzt. Aufbauend darauf erfolgt iterativ die Bearbeitung der verbleibenden Rasterzellen zur Gewässermitte hin (Abb. 6), bis alle Rasterzellen prozessiert wurden. Als Ergebnis erhält jede Rasterzelle eine genäherte Gewässertiefe, welche für die Prozessierung der Rasterzellen selber (Kap. 3.4) als Kontrolle dient.



Abb. 6: Iterative Prozessierung der Rasterzellen unter Hinzunahme von Nachbarschaftsinformationen, beginnend am Uferbereich hin zur Gewässermitte. Braun=Land-Rasterzellen mit einer Gewässertiefe gleich 0,0 m. Gelb = zu prozessierende Rasterzellen. Rot = noch nicht prozessierte Rasterzellen. Grün = bereits prozessierte Rasterzellen.

3.4 Prozessierung der Rasterzellen – Full-Waveform-Stacking

Zur Bestimmung der Gewässertiefe der einzelnen Rasterzellen werden redundante Informationen aller in der Rasterzelle liegenden Full-Waveform-Daten in Form eines Full-Waveform-Stacking Ansatzes, ähnlich wie in STILLA et al. (2007), RONCAT et al. (2016) und PLENKERS et al. (2013), verarbeitet. Hierfür werden alle Full-Waveforms der jeweiligen Rasterzelle zueinander ausgerichtet (z.B. am Wasseroberflächen-Peak) und aufaddiert (Abb. 7). Durch das Zusammenfassen von Full-Waveforms wird der zufällige Rauschanteil des zurückkommenden Signals und erratische Rückstreueffekte innerhalb der Wassersäule stark reduziert bzw. eliminiert. Wiederkehrende Merkmale (u.a. Gewässersohlenpunkte) mit einer schwachen Signalintensität werden verstärkt und sind somit besser und zuverlässiger detektierbar.

Als Ergebnis wird eine mittlere Sampleanzahl zwischen Wasseroberfläche und Gewässersohle bestimmt, wodurch die Detektion des eigentlichen Bodensignals in den einzelnen Full-Waveforms gestützt wird. Diese Vorgehensweise ist insofern vorteilhaft, da – insbesondere bei unebenem Boden – Verschmierungseffekte durch das Full-Waveform-Stacking vermieden werden.



Abb. 7: Prinzipskizze zum Full-Waveform-Stacking. "WF1-WF5" repräsentieren einzeln aufgenommene Full-Waveforms, während die "stacked WF" die Summe der "WF1-WF5" darstellt. Der grüne Bereich zeigt den resultierenden Suchbereich innerhalb jeder einzelnen Full-Waveform. Rote Kreise = "falsche" Bodensignale. Schwarze Kreise = korrekte Bodensignale.

3.5 Detektion und Extraktion der Gewässersohlenpunkte

Basierend auf den Ergebnissen des Full-Waveform-Stackings kann der wahrscheinlichste Gewässersohlenpunkt extrahiert werden. Dazu wird im ermittelten Suchbereich die Full-Waveform analysiert und, sofern vorhanden, das lokale Maximum detektiert. Anhand der bekannten Strahlengeometrie und der Sampleanzahl bzw. der Laufzeit des Laserstrahls zwischen Wasseroberfläche und Gewässersohle können die Punktkoordinaten berechnet werden.

$$\mathbf{P}_{\rm GS} = \mathbf{P}_{\rm WO} + t_2 * c_W * \mathbf{X}_2 \tag{1}$$

Abbildung 8 und Gleichung 1 beschreiben den prinzipiellen Weg vom Wasseroberflächenpunkt P_{WO} zum Gewässersohlenpunkt P_{GS} . Die Koordinaten des Wasseroberflächenpunktes werden aus der Standardprozessierung übernommen. Der Richtungsvektor X_1 wird aus dem Ursprung der Pulsemission P_{LS} und dem Wasseroberflächenpunkt P_{WO} ermittelt. Auf Grundlage des Richtungsvektors X_1 sowie der aus Nachbarpunkten der Wasseroberfläche berechneten lokalen Wasseroberflächennormalen (RICHTER et al. 2018) erfolgt die Refraktionskorrektur nach dem Gesetz von Snellius. Folglich kann die neue Ausbreitungsrichtung X_2 für den Laserstrahl bestimmt werden. Mit der ermittelten Laufzeit *t2*, entsprechend der Ausbreitungsgeschwindigkeit in Wasser *c*w angepasst, wird der Gewässersohlenpunkt berechnet.



Abb. 8: Schematische Visualisierung zur Berechnung eines Gewässersohlenpunktes.

4 Ergebnisse und Diskussion

In diesem Abschnitt erfolgt die Präsentation und Auswertung der Ergebnisse der drei gezeigten Teilgebiete, um Aussagen bzgl. Genauigkeit und Zuverlässigkeit treffen zu können. Grundlage hierfür bilden Referenzdaten mit einer übergeordneten Genauigkeit, welche für die vorliegenden Echolotdaten und die Bodenpunkte aus der Standardprozessierung jedoch nicht gewährleistet werden kann. In Konsequenz werden sie somit als Vergleichsdaten gegenübergestellt.

Bei der Auswertung der Daten kam eine Rasterzellengröße von $2,5 \text{ m} \times 2,5 \text{ m}$ mit bis zu 218 Punkten pro Rasterzelle zur Anwendung. Eine erste visuelle Analyse der Ergebnisse anhand eines für das Untersuchungsgebiet repräsentativen Profils (Abb. 9) zeigt, dass die neudetektierten Bodenpunkte große Teile der Gewässersohle mit Eindringtiefen von bis ca. 2,0 m sehr gut und zuverlässig abbilden (im Vergleich zu ca. 1,65 m für die Standardprozessierung).

Für eine Genauigkeits- und Zuverlässigkeitsanalyse der Resultate dienen die Root-Mean-Square (RMS) Abweichung der Höhenkomponenten zwischen Ergebnis- und Vergleichsdaten (Tab. 1, Zeile 1). Für eine Einschätzung der Güte der neudetektierten Gewässersohlenpunkte wurde als Genauigkeitsanforderung die in VOSSELMANN & MAAS (2010) angegebene mittlere Höhengenauigkeit von $\pm 0,25$ m bei ebener Wasseroberfläche für Airborne LiDAR Bathymetrie verwendet. Dies wird durch die International Hydrographic Organization bekräftigt, welche eine zulässige Unsicherheit von 0,25 m als Minimum Standard for Hydrographic Surveys angibt (IHO 2011). Zudem erfolgt die Prüfung der Zuverlässigkeit der neudetektierten Gewässersohlenpunkte, indem der prozentuale Anteil jener Punkte ermittelt wird, welche in der Höhekomponente 0,25 m und weniger von den Vergleichsdaten abweichen (Tab. 1, Zeile 2).

Die Ergebnisanalyse der Untersuchungsgebiete bestätigt eine gute Repräsentation der Gewässersohle durch die neudetektierten Bodenpunkte. Die RMS-Abweichungen bzgl. der Höhenkomponente zu den Vergleichsdaten liegen zwischen 0,14 m bis 0,18 m (Abb. 10 und Tab. 1). Der Punktanteil, welcher eine Höhenabweichung von 0,25 m nicht überschreitet, beträgt 83,48 % bis 91,08 %. Die gute Höhengenauigkeit und die hohe Zuverlässigkeit sowie eine erreichte Eindringtiefe von bis zu 2,0 m führen zu einer weitläufigeren Erfassung der Gewässerbodentopographie (Abb. 11a-c). Für eine komplette Erfassung der Gewässersohle durch neuprozessierte ALB- und Echolot-Daten wären höhere Eindringtiefen nötig gewesen (Abb. 11d-f).



Abb. 9: In der Breite gestauchtes Profil durch die Elbe. Blau= Wasseroberfläche, Rot = Echolotdaten, Schwarz = Gewässersohle aus Standardprozessierung, Grün = Gewässersohle aus Full-Waveform-Stacking Ansatz.



Abb. 10: Farbcodierte Darstellung der RMS-Höhenabweichungen für (a) Elbe Nord, (b) Elbe Süd und (c) Elbe Ost. Braun = Landpunkte.



Abb. 11: Abdeckung durch die neudetektierten Punkte im Gebiet Elbe Nord (a, d), Elbe Süd (b, e) und Elbe Ost (c, f). Braun = Landpunkte, Schwarz = Punkte der Standardprozessierung (a - c); Echolot-Daten (d - f) und Grün = Punkte der Full-Waveform-Stacking Prozessierung.

Tab. 1: Zeile 1: RMS-Höhenabweichungen zwischen neudetektierte Bodenpunkten und Vergleichsdaten. Zeile 2: Anzahl der Punkte, die eine Höhenabweichung von max. 25 cm zu den Vergleichsdaten aufweisen, sowie prozentualer Anteil zur Gesamtpunktezahl.

	Elbe Nord	Elbe Süd	Elbe Ost
RMS	0,14 m	0,15 m	0,18 m
max. ±0,25 m	54642 / 60457 Punkte 90,38 %	77053 / 84600 Punkte 91,08 %	108870 / 130414 Punkte 83,48 %

5 Fazit & Ausblick

Die Zielstellung einer erhöhten Eindringtiefe und einer höheren Datendichte sowie damit einhergehend einer vollständigen Erfassung der Gewässersohle bei einer hohen Zuverlässigkeit wurde zu einem Großteil erreicht. Die Ergebnisse demonstrieren die sehr gute Anwendbarkeit der entwickelten Methoden auf die gegebenen Daten. Die Erhöhung der Eindringtiefe von ca. 1,65 m auf bis zu 2,0 m entspricht einen Zuwachs um 21 %. Dies bewirkt gleichzeitig eine flächenmäßig deutlich ausgedehntere Repräsentation des Gewässerbodens durch die neudetektierten Punkte. Anhand der Untersuchungen mit den Vergleichsdaten der Gewässersohle konnte eine hohe Zuverlässigkeit bei der Verwendung von Full-Waveforms als Datenbasis nachgewiesen werden. Etwa 87,33 % der detektierten Punkte haben eine Höhenabweichung von 25 cm zu den Vergleichsdaten nicht überschritten. Die RMS Abweichungen in der Höhenkomponente für alle neudetektierten Punkte lag bei 0,16 m und zeigt eine gute Höhengenauigkeit.

Die Lücke zu den Echolot-Messungen konnte dennoch nicht immer geschlossen werden, da dafür Eindringtiefen größer 2,0m nötig gewesen wären. Dennoch stellen die prozessierten Daten eine signifikante Verbesserung bzw. Erweiterung zu den bereits vorhandenen Daten dar.

Die entwickelten Methoden bieten Potential für weitere Anpassungen und Verbesserung, sodass in zukünftigen Arbeiten nachstehende Ziele verfolgt werden:

- Adaptive Definition von Nachbarschaften bei der Auswahl von Full-Waveforms für das Full-Waveform-Stacking
- Test des Potentials einer alternativen räumlichen Aufteilung der Prozessierungsgebiete in Verbindung mit dem Full-Waveform-Stacking (siehe Abschnitt 3.2)
- Verbesserung der Erfassung von dynamischeren Gewässerbodentopographien, um u. a. die Probleme im Uferbereich zu beheben
- Entwicklung weiterer Strategien zur Steigerung der Zuverlässigkeit
- Weiterführende Untersuchungen bzgl. der Verwendung von signaldämpfungskorrigierten Rückstreuquerschnitten für eine zusätzliche Erhöhung der Eindringtiefe

Zudem soll das Potential der hier vorgestellten erweiterten Auswertemethoden bei Anwendung auf maritime Gewässer untersucht werden. In maritimen Gewässern kann im Allgemeinem von einer geringeren Trübung und somit von einer höheren Eindringtiefe ausgegangen werden. Derzeit erfolgt eine Erfassung der Gewässerbodentopographie meist mit hydroakustischen Verfahren wie Vertikal- und Fächerecholot. Dennoch kommen in küstennahen Regionen vermehrt Methoden der Laserbathymetrie zur Anwendung (CHRISTIANSEN 2016; ELLMER et al. 2014). Es ist zu erwarten, dass die Übertragung der vorgestellten Auswertemethoden auf maritime ALB-Daten zu einer signifikanten Verbesserung der Erfassung von Meeresbodentopographien führt.

6 Danksagung

Die Arbeiten zur Laserbathymetrie werden von der Deutschen Forschungsgemeinschaft (DFG) gefördert. Wir danken auch der Bundesanstalt für Gewässerkunde (BfG) und dem Wasserstraßenund Schifffahrtsamt Dresden (WSA Dresden) für die Unterstützung unserer Forschungsarbeiten sowie der Firma Airborne Hydro Mapping GmbH (AHM) für die gute Zusammenarbeit bei der Aufbereitung der Befliegungsdaten.

7 Literaturverzeichnis

- CHRISTIANSEN, L., 2016: New techniques in capturing and modelling of morphological data. Hydrographische Nachrichten, HN 105, 20-23.
- ELLMER, W., ANDERSON, R. C., FLATMAN, A., MONONEN, J., OLSON, U. & ÖlÅS, H., 2014: Feasibility of laser bathymetry for hydrographic surveys on the Baltic Sea. The International Hydrographic Review, (12), 33-50.
- GUENTHER, G. C. & GOODMAN, L. R., 1978: Laser applications for near-shore nautical charting. Proc. SPIE Ocean Optics V, 160, 174-184.
- INTERNATIONAL HYDROGRAPHIC ORGANIZATION, 2011: Manual on Hydrography, Publication C-13. International Hydrographic Bureau Monaco, Feb. 2011.
- PLENKERS, K., RITTER, J. R. & SCHINDLER, M., 2013: Low signal-to-noise event detection based on waveform stacking and cross-correlation: application to a stimulation experiment. Journal of Seismology, **17**(1), 27-49.
- RICHTER, K., MADER, D., WESTFELD, P. & MAAS, H.-G., 2018: Numerical simulation and experimental validation of wave pattern induced coordinate errors in airborne LiDAR bathymetry. The International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 42(2), 961-967.
- RIEGL, 2017: Waveform Extraction Library. RIEGL Laser Measurement Systems GmbH. Printed: 31.03.2017.
- RONCAT, A. & MANDLBURGER, G., 2016: Enhanced detection of water and ground surface in airborne laser bathymetry data using waveform stacking. EGU General Assembly Conference Abstracts, 18, 17016.
- STILLA, U., YAO, W. & JUTZI, B., 2007: Detection of weak laser pulses by full waveform stacking. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(3-W49A).
- VOSSELMAN, G. & MAAS, H.-G., 2010: Airborne and terrestrial laser scanning. CRC, 2010.

Die Möglichkeiten der Full Waveform für bathymetrische Airborne LiDAR Daten

WOLFGANG DOBLER^{1,2}, FRANK STEINBACHER¹, RAMONA BARAN¹, WERNER BENGER¹, MARCEL RITTER^{1,3} & WOLFGANG LEIMER¹

Zusammenfassung: Heutige bathymetrische Laserscanner sind in der Lage, die gesamten Full-Waveform (FWF) Daten neben der Online-Prozessierung (OP) abzuspeichern. Für einen Laserscanflug liefern die FWF Daten zusätzliche, intrinsische Informationen, die einen vertieften Einblick in die Geometrie der Sohle erlauben. Der besondere Fokus liegt in jenem Sohlbereich für das Projekt Elbe mit einer Wassertiefe größer 0.5 m, die mit der OP teilweise nicht mehr erfasst werden können. Für die Entfaltung der FWF wird der Richardson-Lucy Algorithmus verwendet. Der Wasserkörper (ab einer Tiefe > 0.5 m) in den FWFs wird mit Hilfe eines e-Funktion Fittings bestimmt. Anschließend wird ein Waveform-Stacking durchgeführt, um das Signal-Rausch-Verhältnis zu verbessern. Abschließend wurden dann mit dem Gradienten der FWF und einem variablen Schwellenwert Peaks in der FWF gesucht und in 3D Punktkoordinaten überführt. Die Datengrundlage für die FWF Auswertungen ist ein Teilabschnitt der Elbe in Deutschland, für den mit Hilfe von Echolotdaten auch die Höhengenauigkeit der zusätzlich aus den FWFs extrahierten Sohlpunkte überprüft wurde.

1 Einleitung

Während eines bathymetrischen Laserscanflugs (STEINBACHER & AUFLEGER 2013) verarbeitet die Online-Prozessierung (OP) des Riegl-Lasers VQ880-G in Echtzeit die FWF zu geometrischen Koordinaten. Da die Aufnahmefrequenz 550 kHz beträgt, kann die OP keine rechenintensiven Auswertealgorithmen durchführen, um möglichst viele Objekte im Zielgebiet zu erfassen und als Punktkoordinate auszugeben. Hier bietet sich eine nachträgliche - wie seit Jahren schon vielfach angewendet - Offline-Prozessierung an (GUENTHER et al. 2000), um eine genauere Auswertung durchzuführen. Der Fokus der Auswertung liegt hier vor allem auf größeren Wassertiefen und einer höheren Punktdichte in tieferen Bereichen, in der die OP nur geringe oder keine Sohlpunkte ausgibt. Die einzelnen Schritte für eine zusätzliche Auswertung werden im Folgenden dargestellt und sind Bestandteil einer fertigen Massenprozessierungskette innerhalb der Software HydroVISH geworden.

2 Konvertierung vom Riegl-Format nach HDF5/F5

Nach erfolgter Streifenausgleichung und Georeferenzierung müssen derzeit die OP-Punktdaten und FWFs aus drei verschiedenen Dateien zusammengeführt werden. Anhand des GPS-Zeitstempels (in Wochensekunden) des emittierten Laserpulses (aus las-Datei) und der Entfernung

¹ Airborne Hydro Mapping GmbH, Feldstraße 1b, A-6020 Innsbruck

² Universität Innsbruck, Arbeitsbereich für Wasserbau, Technikerstr. 13, A-6020 Innsbruck

³ Universität Innsbruck, Institut für Informatik, Technikerstr. 21a, A-6020 Innsbruck

zwischen dem detektierten Punkt und der Trajektorie (aus sdc-Datei) kann der Zeitpunkt der Detektion innerhalb einer Waveform (wfm-Datei) errechnet werden. Die relevanten Datenfelder der Punkte und zugehörigen FWFs werden dann im F5-Layout (BENGER 2004; RITTER 2009), basierend auf dem HDF5-Format (THE HDF5 GROUP 2018), zur Weiterverarbeitung abgespeichert. Das aus dem High-Performance-Computing Bereich stammende Open-Source Format HDF5 erlaubt schnelle Zugriffe auf große Daten und flexible Datenstrukturen. Das F5-Layout bietet hierzu eine konkrete Organisation für raumzeitliche Daten und ermöglicht erst die lückenlose Extraktion und Verknüpfung aller Sensordaten.

3 Voruntersuchung

Die Offline-Prozessierung macht nur dann Sinn, wenn in den FWFs optisch ein Bodensignal zu erkennen ist. Hierfür werden mit dem Softwarepaket HydroVish der AHM GmbH (Derivat von Vish (BENGER et al. 2007)) alle FWFs entlang ihrer Richtungsvektoren gerendert. In Abb. 1 ist dabei klar zu erkennen, dass die OP nicht alle Sohlpunkte als Koordinaten ausgespielt hat und eine vertiefende Analyse der FWFs sinnvoll ist, da das physische Signal mit den Detektoreigenschaften - der sogenannten Systemwelle - verschmiert wird.



Abb. 1 oben: Querschnitt entlang der Elbe aus der Online-Prozessierung. Unten: Gerenderte FWFs entlang der Strahlrichtung. Sohlpeaks sind sichtbar, die wesentlich mehr umfassen als in der OP. In der Detailansicht ist der Doppelpeak aufgrund der Systemwelle sichtbar (siehe auch Abb. 6, Systemwelle).

4 Grobe Vorklassifizierung des Wasserkörpers mit Hilfe einer e-Funktion

Das grüne (490–575 nm) Laserlicht hat die Eigenschaft Wasser zu durchdringen (MOBLEY 1994) und kann daher verwendet werden, um die Sohle von Gewässern zu vermessen. Der vom Flugzeug aus emittierte Laserpuls wird als erstes an der Luft-Wasser Grenze teilweise reflektiert. Diese Reflexion ist meistens ein markanter Peak in der FWF und hängt vom Einfallswinkel und der Rauigkeit der Wasseroberfläche ab. Das restliche Laserlicht, das in den Wasserkörper eindringt, wird exponentiell gedämpft (PETZOLD 1972). Wenn genügend Laserlicht bis zur Sohle vordringt, ist ein zweiter - vor allem bei tieferen Gewässern - wesentlich kleinerer Peak in der FWF zu finden.

Der Amplitudenverlauf im Wasserkörper folgt dem Lambert-Beer'schen Gesetz:

$$A_s = A_0 e^{-\kappa s} \quad (1)$$

wobei A_s die Amplitude, s der Weg der FWF, A_0 die Anfangsamplitude und κ der Dämpfungskoeffizient ist.

Durch umformen erhält man κ:

$$-ln\left(\frac{A_s}{A_0}\right)\frac{1}{s} =: \kappa(s) \qquad (2)$$

Mit Gleichung (2) wird für jede Amplitude der FWF ein κ berechnet. Anschließend wird aus der Funktion $\kappa(s)$ der Median κ als repräsentativer Wert für die diskret gegebenen Werte s verwendet. Beispielhaft wurde in Abb. 2 das Einpassen der e-Funktion in eine FWF (hier Wasserkörper) dargestellt. Das Kriterium für ein erfolgreiches Einpassen der e-Funktion ist die integrierte absolute Differenz zwischen Amplitude A_s und dem dazugehörigen Wert aus der e-Funktion. Mit einem Schwellenwert für die Differenz wird dann festgelegt, ob es sich hier um einen Wasserkörper handelt, wie in Abb. 3 gezeigt.



Abb. 2: Eingefügte e-Funktion am maximalen Peak



Abb. 3 : Klassifizierung des Wasserkörpers aufgrund eines Schwellenwertes für die Abweichung vom exponentiellen Verlauf der Amplitude: rot kein Wasserkörper; grün (letzter Peak) bzw. orange (erster Peak) Wasserkörper

5 Stapelung von benachbarten Wellenformen

Die Sohlpeaks einer FWF sind in vielen Fällen meistens sehr gering ausgeprägt (z.B. bei größeren Wassertiefen oder starke Trübung) und deren Amplitudenhöhe ist nur geringfügig größer als das Rauschen. Um das Signal-Rausch-Verhältnis zu verbessern, wird eine Stapelung von Wellenformen durchgeführt (STILLA et al. 2007). Bei diesem Ansatz werden benachbarte Wellenformen (Max. Entfernung: 25 cm, ähnlicher Richtungswinkel) addiert und anschließend ein Mittelwert gebildet. Zusätzlich werden die FWFs zum Hauptpeak verschoben. Es werden nur solche FWFs für die Stapelung herangezogen, die auch als Wasserkörper klassifiziert worden sind (Abb. 4 links). In Abb. 4 (rechts) ist dann eine gemittelte FWF dargestellt: Die orange Linie demonstriert einen gegenüber der blauen Referenz-FWF rauschverminderten Verlauf sowie zwei prominente Peaks (Übergang Luft nach Wasser und Sohlpeak).



Abb. 4: Links: Mit der blauen Referenz-Waveform (Nr. 4189) werden benachbarte FWFs gesucht (strichpunktierte Linien) und entlang des Hauptpeaks ausgerichtet. Rechts: Die gemittelte FWF in orange weist einen ruhigeren Verlauf auf als die blaue Referenz-Waveform.
6 Entfaltung der FWF

Die vom Sensor gemessene FWF ist eine Faltung der Systemwelle mit dem differentiellen Rückstreuquerschnitt (dRSQ oder Target Funktion). Dadurch kommt es typischerweise zu einer Verschmierung der Wellenform und zu einer Verringerung der Signalauflösung. Um nun die Objektform zu erhalten, findet man in der Literatur zahlreiche Möglichkeiten für die Entfaltung, wie z.B. Wiener Filter, B-Splines (RONCAT et al. 2010) oder auch mit Hilfe der Dekomposition (SCHWARZ et al. 2017). Für die Entfaltung der FWF-Daten des Elbe-Abschnittes haben wir uns hier für den Algorithmus nach Richardson-Lucy (RL) entschieden, der in anderen Projekten gute Ergebnisse lieferte (WU et al. 2011).

Da die Systemwelle der Firma Riegl in analytischer Form beschreibbar ist, wurde ein Upsampling der FWF um den Faktor 10 durchgeführt. Dadurch erhält man eine feinere Auflösung der FWF von $\Delta s = 17$ cm auf $\Delta s = 1,7$ cm. In Abb. 5 ist eine Entfaltung beispielhaft durchgeführt worden: Die Target-Funktion besteht aus einer e-Funktion, die mit zwei nahe beieinander liegenden Peaks am Ende versehen ist. Die Target-Funktion wird nun mit der Systemwelle gefaltet und man erhält eine FWF mit einem verschmierten Doppelpeak am Ende. Anschließend wird ein normalverteiltes Rauschen (Größe der Schwankung wie in Abb. 2 vor dem Hauptpeak) aufgebracht und eine Entfaltung mit dem RL-Verfahren durchgeführt. In Abb. 5 ist erkennbar, dass durch das Upsampling das rückgefaltete Signal besser mit der Target-Funktion übereinstimmt (FWF zur Anschaulichkeit verschoben). Das Ergebnis für eine gemessene FWF ist in Abb. 6 dargestellt; die Peaks im rückgefalteten Signal sind gut zu erkennen.



Abb. 5: Analytische Testfunktion: Target Objekt entspricht e-Funktion mit zwei Peaks an der Sohle: Rückentfaltung der Systemwelle mit FWF via Richardson-Lucy. Links: Originalgröße der FWF (Bin-Länge 17.2cm). Rechts: Gleiche Entfaltung mit verdichteter Punktanzahl (Bin-Länge 1.72cm).



Abb. 6: Entfaltung an einer gemessenen FWF. Zwei Peaks in Sohlnähe kommen prominent zum Vorschein durch die Rückentfaltung. Die Größenordung des Versatzes aufgrund der Faltung FWF mit Systemwelle ist zu erkennen.

Da die Entfaltung nach RL ein iteratives Verfahren ist und ein Upsampling um den Faktor 10 durchgeführt worden ist, ist die Berechnung für alle FWFs sehr zeitaufwändig. Den Ergebnissen vorausgreifend, ist auch untersucht worden, ob alternativ zur mathematischen exakten Entfaltung auch einfach nur der Versatz der Systemwelle berücksichtigt werden kann (siehe Abb. 6, Versatz). Das Ergebnis für die Elbe zeigt, dass eine Entfaltung in diesem Fall nicht notwendig ist und daher darauf verzichtet worden ist.

7 Peak-Detektion

Aus den FWFs werden Punktkoordinaten abgeleitet. Der verwendete Algorithmus basiert auf einem Gradientenfeld, in dem abhängig von einem Schwellenwert Peaks detektiert werden. Zusätzlich wird der Schwellenwert in tieferen Gewässern geringer angesetzt, um auch die schwach ausgeprägten Sohlpeaks zu finden (siehe Abb. 7).



Abb. 7: Mit einem variablen Schwellenwert werden im Gradientenfeld Peaks detektiert.

In Abb. 8 ist der Vergleich mit der Online-Prozessierung (OP) dargestellt: Die OP zeigt den Beginn der Sohle an, weiter Richtung Flussmitte werden keine Sohlpunkte mehr detektiert. Mit der Gradientenauswertung werden mehr Sohlpunkte gefunden, auch wenn noch ein deutliches Rauschen um die Sohle vorhanden ist. Wesentlich besser ist die Auswertung mit der Stapelung der FWFs, bei der ein geringeres Rauschen zu erkennen ist (Suchradius 25 cm). Für die Auswertung mit der Entfaltung nach RL ist soweit keine wesentliche Verbesserung zu erkennen, auch wenn tendenziell das Rauschen etwas geringer ist.



Abb. 8: Vergleich der verschiedenen Auswertemethoden mit der Online-Prozessierung.

8 Ergebnisse Projekt Elbe

Mit der gezeigten Vorgehensweise wurden die FWF-Daten eines 20km langen Abschnittes der Elbe bei Klöden inklusive einiger Nebengewässer ausgewertet. Die Befliegung erfolgte am 07.

September 2015 unter Verwendung des VQ880-G. Die Daten der Online-Prozessierung wurden bereits im Rahmen eines Pilotprojektes zur DGM-W-Modellierung verwendet (für WSA, BFG und BAW). Es wurden insgesamt FWFs für 18 Streifenpaare (Vor- & Rückstreifen) ausgewertet. An Hand dieser Auswertung wird der Mehrgewinn durch die nachträgliche Offline-Prozessierung der FWF-Daten deutlich (Abb. 9). In einem Profilschnitt wurden zudem die klassifizierten Offline-Sohlpunkte den Airborne Laser Bathymetry (ALB)-online-Sohlpunkten und den Echolotpunkten aus 2015 gegenübergestellt (Abb. 10). Es wird deutlich, dass Lücken zwischen den ALB-Online-Geländepunkten und Echolot 2015 geschlossen bzw. deren Größe deutlich reduziert werden können unter Berücksichtigung der Offline-klassifizierten Sohlpunkte aus der FWF. Zudem zeigt sich bei visueller Betrachtung eine gute Übereinstimmung zwischen FWF-Offline-Sohle und Echolot 2015 eine Standardabweichung von 9 cm aufweist (Abb. 11).



Abb. 9: Draufsicht Elbe. Blaue Linie ist Wasser-Land-Grenze. Links: Sohl- und Bodenpunkte aus der OP, rechts Sohlpunkte aus Offline-Prozessierung in rot



Abb. 10: Elbe-Querschnitt mit ALB-Online-Punkten (Gelände & Vegetation) eingefärbt nach RGB sowie Wasserspiegel in blau und Echolot 2015 in grün. In rot sind die Offline-klassifizierten Sohlpunkte zum Vergleich aus der FWF eingeblendet.



Abb. 11: Oben: Differenzen der gerasterten Echolot-Daten mit den FWFs: Die Standardabweichung beträgt 9 cm. Unten: Wahrscheinlichkeitsverteilung der Differenzen gegen die theoretischen Quantilwerte aus einer Normalverteilung.

9 Zusammenfassung

Für das Projekt Elbe wurden die FWF-Daten als Offline-Prozessierung ausgewertet. Mit der Prozessierungs- und Visualisierungssoftware HydroVish können die massiv größeren Rohdaten (Faktor 100 im Vergleich zu ausschließlich Koordinaten) der FWF vollständig und flächendeckend dargestellt werden. Dies erlaubt als erster Schritt eine Abschätzung ob eine Offline-Prozessierung überhaupt notwendig ist. Anschließend wurde mit dem beschriebenen Exponentialansatz der Wasserkörper aus den Wellenformen klassifiziert. Mit den daraus resultierenden FWFs wurde dann eine Stapelung durchgeführt. Damit konnte das Signal-Rausch Verhältnis verbessert werden. Eine Entfaltung nach Richardson-Lucy (RL) wurde anhand eines Testquerschnitts durchgeführt. Dabei stellte sich heraus, dass eine Entfaltung nach RL zu keiner signifikanten Verbesserung der Ergebnisse in diesem Projektdatensatz führte. Daher wurde hierauf verzichtet. Die Peaks in der FWF wurden stattdessen über den Gradienten mit Hilfe eines variablen Schwellenwertes bestimmt. Das Ergebnis der Punktkoordinaten aus der FWF wurde mit den Echolot-Daten der Elbe verglichen: Die Standardabweichung beträgt hier 9 cm und liegt in der Größenordnung der Genauigkeit der LiDAR-Daten (10 cm).

Aktuelle Untersuchungen befassen sich im Zuge des integrativen Aufbaus einer topobathymetrischen Prozesskette für LiDAR Daten (vom Sensor zum Datenprodukt in einer Prozessumgebung) mit der Einsatzfähigkeit von modernen Methoden aus dem Bereich der Artificial-Intelligence, wie z.B. DeepLearning (DAVIS & MARCUS 2016) für Klassifizierung oder Peak-Detektion.

10 Literaturverzeichnis

- BENGER, W., 2004: Visualization of General Relativistic Tensor Fields via a Fiber Bundle Data Model. Berlin: Lehmanns Media-LOB.
- BENGER, W., RITTER, G. & HEINZL, R., 2007: The Concepts of VISH. In 4. High-End Visualization Workshop, Obergurgl, Tyrol, Austria, June 18-21, 2007, 26-39). Berlin, Lehmanns Media-LOB.de.
- DAVIS, E. & MARCUS, G., 2016: The scope and limits of simulation in automated reasoning. *Artificial Intelligence*. https://doi.org/10.1016/j.artint.2015.12.003
- GUENTHER, G. C., CUNNINGHAM, A. G., LAROCQUE, P. E. & REID, D. J., 2000: Meeting the Accuracy Challenge in Airborne Lidar Bathymetry. EARSeL EProceedings.
- MOBLEY, C. D., 1994: Light and Water : Radiative Transfer in Natural Waters. Light and Water : Radiative Transfer in Natural Waters. https://doi.org/10.1002/9781118622179
- PETZOLD, T., 1972: Volume Scattering Functions for Selected Ocean Waters. Scripps Institution of Oceanography. https://doi.org/10.5811/westjem.2013.7.18472
- RITTER, M., 2009: Introduction to HDF5 and F5. Baton Rouge. Retrieved from https://www.cct.lsu.edu/cct-trs/download.php?file=CCT-TR-2009-13/CCT-TR-2009-13.pdf
- RONCAT, A., BERGAUER, G. & PFEIFER, N., 2010: Retrieval of the backscatter cross-section in fullwaveform LIDAR data using B-splines. The Int. Archives Photogramm. Remote Sens. Spatial Inf. Sci., 38(3B), 137-142.
- SCHWARZ, R., PFEIFER, N., PFENNIGBAUER, M. & ULLRICH, A., 2017: Exponential decomposition with implicit deconvolution of lidar backscatter from the water column. PFG - Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 85(3), 159-167. https://doi.org/10.1007/s41064-017-0018-z
- STEINBACHER, F. & AUFLEGER, M., 2013: Airborne hydromapping: A new approach to manage hydraulics, morphology and ecology. International Journal on Hydropower and Dams, 20, 74-77.
- STILLA, U., YAO, W. & JUTZI, B., 2007: Detection of weak laser pulses by full waveform stacking. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(3/W49), 25-30.
- THE HDF5 GROUP, 2018: Hierarchical data format version 5.
- WU, J., VAN AARDT, J. A. N. & ASNER, G. P., 2011: A comparison of signal deconvolution algorithms based on small-footprint LiDAR waveform simulation. IEEE Transactions on Geoscience and Remote Sensing, 49(6), 2402-2414. https://doi.org/10.1109/TGRS.2010.2103080

Detaillierte Analyse komplexer Wellenformen eines UAV Laser Bathymetrie Profilmessgerätes

ROLAND SCHWARZ¹ & MARTIN PFENNIGBAUER¹

Zusammenfassung: Lineare Erfassung von Profilen in der Hydrologie und im Wasserbau hat trotz der Verfügbarkeit von mehrdimensionaler Modellierung nach wie vor seine Berechtigung. Der von RIEGL entwickelte Bathycopter, ein unbemannter Oktokopter ausgestattet mit dem Bathymetric Depth Finder RIEGL BDF-1, ist ein für die Profilerfassung speziell für die kleinräumige, unbemannt-luftgestützte Anwendung geeignetes UAV. Das primär erwartbare Ergebnis aus einer Befliegung mit diesem Vermessungs-UAV ist ein topografisches Profil eines Wasserkörpers bestehend aus den Linien des Bodens und der Wasseroberfläche. Bei der verwendeten Aufnahmetechnik werden dabei zunächst die kompletten Kurvenformen der rückgestreuten Laserechos aufgezeichnet, um in einer Nachverarbeitung die Punkte der Oberfläche und des Bodens zu extrahieren. Die aufgezeichneten Wellenformen zeigen dabei hohen Detailreichtum, der derzeit ungenutzt und noch nicht zur Gänze verstanden ist. Während das Studium des Verlaufs der Rückstreuung in der Atmosphäre zum Standardrepertoire in der Laser-Messtechnik zählt, ist die Beschäftigung mit den Ausbreitungsdetails in der Wassersäule noch nicht vollständig im technischen Alltag angekommen. Eine Unterscheidung von Artefakten, die durch das Verfahren selbst verursacht sind, und Strukturen, die für das Verständnis der Wassersäule nützlich sind, könnte ein erster Schritt zu einer Nutzbarmachung dieser durch das Verfahren ohnehin anfallenden und derzeit brachliegenden Daten sein.

1 Einleitung

Die genaue Kenntnis der Morphologie von Gewässern ist für eine Vielzahl von Anwendungen von erheblicher Bedeutung. Für die Schifffahrt, Fischerei, Bauprojekte, Umweltschutz, Abschätzung von Risiken durch Fluten und Stürme oder die Erfassung der entstandenen Schäden ist es unerlässlich, wiederholt hochauflösende und präzise Vermessungen von Ufer, Wasserspiegel, Gewässergrund und der umliegenden Landschaft durchzuführen. Bei schiffbaren Gewässern kann die Erfassung des Gewässergrunds mittels Echolot erfolgen. Die gewonnenen Daten können dann mit jenen von Landvermessung kombiniert werden, wobei der unmittelbare Uferbereich meist gar nicht oder nur unzulänglich erfasst werden kann. Dieses Problem wird durch Airborne Laserscanning gelöst, wo von der Luft aus eine nahezu lückenlose Erfassung von Gewässern und deren Umland erfolgen kann. Für kleine, seichte Gewässer oder kleinflächige Vermessungsprojekte kommt nach wie vor die traditionelle Methode der Vermessung mittels Messlanze, die direkt mit einem GPS Empfänger ausgestattet ist oder deren Position vom Land aus mittels Totalstation vermessen wird, zum Einsatz. Speziell bei Fließgewässern werden auf diese Art und Weise Profile in vorgegebenen, regelmäßigen Abständen erfasst.

¹ Riegl Research Forschungsgesellschaft mbH, A-3580 Horn, E-Mail: [rschwarz, mpfennigbauer]@riegl.com

Inspiriert durch diese klassische Methode der Vermessung von Fließgewässern und unter Berücksichtigung der Einschränkungen hinsichtlich Gewicht, Leistungsaufnahme und Größe entwickelte RIEGL einen Laser-Entfernungsmesser zur Erfassung von Gewässerprofilen von unbemannten fliegenden Plattformen (UAVs) aus. Der mit einem bei 532 nm, also im sichtbaren grünen Wellenlängenbereich, emittierenden Laser ausgestattete Entfernungsmesser wird dabei von dem UAV über das Gewässer bewegt und erfasst dabei Ufer, Wasseroberfläche und Grund, soweit es die Wasserqualität und -tiefe zulassen. Durch Verzicht auf einen Scanmechanismus muss zwar auf die flächige Erfassung des Geländes verzichtet werden, dafür kann aber Größe und Gewicht des Messinstruments klein genug gehalten werden, um auch von small UAVs der Gewichtsklasse < 25 kg getragen werden zu können. Die geringe Flughöhe beim Einsatz von UAVs von < 50 m ermöglicht die Aufzeichnung der gesamten Echowellenform zur späteren Nachbearbeitung mit geeigneter Software. Darüber hinaus ergibt sich der Vorteil, dass aufgrund der geringen Fortbewegungsgeschwindigkeit von typisch wenigen m/s die Echosignale aufeinanderfolgender Laserschüsse zu einer Messung kombiniert werden können. Daraus ergeben sich Echo-Wellenformen mit gegenüber der Einzelmessung verbessertem Signal-Rauschverhältnis. Die detaillierte Analyse dieser Wellenformen ergibt nicht nur Lage von Oberfläche und ggf. Grund sondern kann auch Informationen über die Wassersäule, wie Absorptionskoeffzient aber

Grund sondern kann auch Informationen über die Wassersäule, wie Absorptionskoeffzient aber auch etwaige Schichtungen liefern. Dies ist Gegenstand dieses Beitrags. Wir verwenden dazu einen Datensatz der auf Hawaii im Hafen von Hilo rund um einen Wellenbrecher aufgenommen wurde.

2 LIDAR Wellenformen in der Bathymetrie

2.1 Full-Waveform LIDAR

Seit mehr als einem Jahrzehnt sind so genannte "Full-Waveform" LIDAR (Light Detection and Ranging) Systeme im kommerziellen Einsatz (MALLET & BRETAR 2008) die in der Lage sind, den kompletten vom Ziel zurückgestreuten Wellenzug eines ausgesendeten Laserimpulses im zeitlichen Verlauf zu digitalisieren und teilweise oder vollständig aufzuzeichnen. In der Nachverarbeitung können die so erfassten Daten mit Hilfe von Algorithmen unterschiedlichen Komplexitätsgrades und Rechenzeitaufwandes analysiert werden. Ein Standardverfahren ist die Zerlegung in Gauss-Glockenkurven (WAGNER et. al 2006), bei dem Positionen von Einzelzielen sowie deren Amplituden ermittelt werden.

Die Gestalt der empfangenen Wellenform hängt sowohl von der Zielsituation (Einzelziel, Mehrfachziel) als auch von der Form des ausgesandten Lasersignals und der Filtercharakteristik des Empfängers ab.

2.2 Superpositionsprinzip

Bei inkohärenter Detektion des empfangenen Laserimpulses sowie bei Betrieb des Detektors und seiner Empfangselektronik im linearen Bereich kann die Gültigkeit des Superpositionsprinzips angenommen werden. Dieses Prinzip besagt, dass sich das Gesamtsignal aus den an einzelnen Zielen zurückgestreuten Teilwellen durch Summation zusammensetzt. Das Empfangssignal

$$p_r(t) = B \sum_i \beta_i p_s \left(t - 2 \frac{n}{c_0} z_i \right) \tag{1}$$

einer Anzahl diskreter Ziele an den Entfernungen z_i vom Sensor mit den Rückstreukoeffizienten β_i und einer Skalierungskonstanten *B* wird aus um die doppelten Laufzeiten $\frac{n}{c_0}z_i$ verschobenen Versionen des effektiven Sendepulses $p_s(t)$ gebildet. Die Konstante *B* fasst die Eigenschaften des Sensors, wie Sende- und Empfangsapertur, Verstärkungsfaktor des Empfängers und die Grunddämpfung zufolge der Distanz zwischen Sensor und Ziel zusammen. Für *n* ist der Wert des im jeweiligen Medium gültige Wert des Brechungsindex zu einzusetzen, c_0 ist die Lichtgeschwindigkeit. Gleichung (1) gilt mit guter Näherung für Ziele deren Ausdehnung klein gegen die Entfernung zum Sensor ist. Der effektive Sendepuls setzt sich dabei aus der Form des ausgehenden Laserpulses und der Impulsantwort des Empfangsfilters zusammen und wird im Folgenden als Systemwellenform bezeichnet (WAGNER et al. 2006).

2.3 Verteilte Rückstreuung aus der Wassersäule

Der Laserpuls erfährt bei der Ausbreitung im Wasser eine Dämpfung und wird an Inhomogenitäten und Partikeln gestreut. Die in Richtung zum Sensor zurückgestreute Lichtmenge steht im Zusammenhang mit den inhärenten optischen Eigenschaften von Wasser, dem Absorptionskoeffizient *a* und dem Streukoeffizient *b*, als auch den Strahleigenschaften und dem Sichtfeld der Empfangsoptik. Ein vereinfachtes Modell, streng nur gültig für ein infinitesimal kleines Sichtfeld und homogene unendliche ausgedehnte Wassersäule (GORDON 1982) lautet

$$p_r(t) = B \int_{z_0}^{\infty} p_s \left(t - 2 \frac{n(z)}{c_0} z \right) \beta e^{-2k(z-z_0)} dz$$
⁽²⁾

wobei z_0 die Entfernung bis zur Wasseroberfläche ist, β der Rückstreukoeffizient und k = a + b, der Dämpfungskoeffizient, der die Summe aus Absorbptions- und Streukoeffizient ist. Der exponentielle Zusammenhang nach Gleichung (2) ist die so genannte "Single Scattering" Lösung, die auch im "Multiple Scattering" Fall Bedeutung hat, obwohl dann der Dämpfungskoeffizient durch einen effektiven Dämpfungskoeffizienten ersetzt werden muss. Unter Beachtung, dass n(z) für $z > z_0$ durch den als konstant angenommenen Brechungsindex n_w von Wasser ersetzt werden kann, kann man die Gleichung 2 durch Variablensubstitution umschreiben zu

$$p_r(t) = B \int_{\tau_0}^{\infty} p_s(t-\tau) \frac{c_0}{2n_w} \beta e^{-k \frac{c_0}{n_w}(\tau-\tau_0)} d\tau.$$
 (3)

Das Modell setzt sich demnach aus der vom Sensor abhängigen System Wellenform $p_s(t)$ und aus dem Exponentialmodell für die Wassersäule über eine Faltungsbeziehung zusammen. Die Systemwellenform, ein gerätespezifischer Parameter, gewinnt man indem man mit dem Sensor auf eine planare, im rechten Winkel zur Laserstrahlachse angeordnete, Fläche mit einer Ausdehnung größer als der Laserfootprint misst und die dabei gewonnene Wellenform aufzeichnet.



Abb. 1: Prototypische Bathymetrie Wellenform bestehend aus zwei Interface Echos, Oberfläche und Boden, sowie dem Echo der Wassersäule.

2.4 Oberfläche – Wassersäule – Boden

Wasseroberfläche und Boden sind im Unterschied zur Wassersäule diskrete Ziele und deshalb in erster Annäherung als Dirac Funktionen modellierbar. Mit

$$\sigma(z) = \beta_s \delta(z - z_0) + \left(u(z - z_0) - u(z - z_1) \right) \beta_v e^{-k(z - z_0)} + \beta_b \delta(z - z_1)$$
(4)

erhält man für die prototypische Bathymetrie-Wellenform den Ausdruck

$$p_r(t) = \int_{z_0}^{\infty} p_s\left(t - 2\frac{n_w}{c_0}z\right)\sigma(z)dz$$
(5)

wobei u(z) die Heaviside`sche Sprungfunktion ist, z_0 die Distanz bis zur Wasseroberfläche, z_1 die Distanz bis zum Boden ist und $\beta_s, \beta_v, \beta_b$ die zugehörigen Rückstreukoeffizienten sind.

Abb. 1 zeigt eine prototypische Bathymetrie-Wellenform nach Gleichung (4) und Gleichung (5).

2.5 Feinstruktur

Sowohl Absorptionskoeffizient als auch Streukoeffizient unterliegen räumlichen Schwankungen. Das Modell nach Abschnitt 2.4 muss deshalb gegebenenfalls um zusätzliche Komponenten erweitert werden.



3 Datensätze – Hafen von Hilo

Abb. 2: Übersicht: Wellenbrecher und Messprofile. Weißes Band: VUX-SYS; graue und gelbe Linie: BDF-1; grüne Linien: Echolot; Hintergrund Bild und Karten: Imagery @2018 Google, Map data @2018 Google.

Für die Untersuchungen stand ein Datensatz zur Verfügung, der im Rahmen einer Evaluierung der Anwendbarkeit von Unmanned Laser Scanning (ULS) für die regelmäßige Prüfung der strukturellen Integrität von Hafenmauern im Juni 2018 aufgenommen wurde (LEWINTER et al. 2018). Neben der Erfassung der Hafenmauer in 3D wurden auch zwei Befliegungen mit einem bathymetrischen Laser-Entfernungsmesser durchgeführt. Details über die Vorverarbeitung der Trajektorie und die Datenerfassung können bei LEWINTER et al. (2018) nachgelesen werden.

3.1 3D - Punktwolke

Insgesamt wurden drei Befliegungen mit dem topographischen Scanner RIEGL VUX-1LR durchgeführt. Für diesen Bericht dienen diese Datensätze allerdings nur zu Illustrationszwecken. In Abb. 2 ist der Flug mit der Nummer 1 als weißes Band im rechten unteren Bildteil zu erkennen. Die Flughöhe betrug 30 m bei einer Dauer von 30 Minuten. Die Daten standen als RiPROCESS (eine Software der Firma RIEGL) Projekt zur Verfügung.

3.2 2D – Profilpunkte

Mit dem BDF-1 von RIEGL wurden zwei Flüge durchgeführt, beide in einer Flughöhe von 20 m und einer Dauer von ca. 20 Minuten. Der BDF-1 ist ein speziell für die Bathymetrie entwickelter Sensor mit einem Laser der Wellenlänge 532 nm, um unter die Wasseroberfläche eindringen zu können. Die nominale Messrate des BDF-1 beträgt 4 kHz und die Fluggeschwindigkeit lag bei 2 ms⁻¹. Daraus ergibt sich ein Punktabstand von 0.5 mm und nach einer Mittelung von je hundert Messungen ein Punktabstand von 5 cm. Vermessen wurden Wasserkörper auf der Hafenseite und im offenen Meer, die Ergebnisse wurden entsprechend auf zwei Datensätze aufgeteilt. In Abb. 2 sind die Flüge über dem Hafenbecken in Grau und die Flüge über dem offenen Meer in Gelb zu erkennen.

3.3 Echolot Messungen

Als Referenz standen Messungen mit dem Echolot vom April 2016 zur Verfügung, die nach Baggerarbeiten im Hafenbecken durchgeführt worden waren (PORTLAND 2016). Die Lotungen lagen in einer maschinenlesbaren Form vor. Die vertikale Komponente wurde durch Vorzeichenumkehr in eine Höhenkoordinate umgewandelt und um einen Offset von 69,511 Surveying Foot verschoben (Siehe Abschnitt 5.1).

4 Methodik

4.1 Vorverarbeitung

Um die Performance von Bathymetrie-Sensoren zu steigern, ist eine hohe Empfindlichkeit wünschenswert. Eine hohe Empfindlichkeit erfordert ein hohes Signal-Rausch-Verhältnis (SNR), das durch Mittelung vor der Analyse erhöht werden kann. Das SNR nimmt bei der Mittelung von *M* aufeinander folgenden Messungen gemäß $SNR_{avg} = SNR\sqrt{M}$ zu. Die hohe Empfindlichkeit des Sensors geht auch mit einer erhöhten Anzahl von Schein-Echos einher, die durch spontane Stromimpulse im Photodetektor ausgelöst werden. Um eine Verfälschung der gemittelten Wellenformen durch die Ausreißer zu verhindern, werden diese vor der Mittelung entfernt, indem nur das 95 Perzentil der Abtastwerte jedes Abtastzeitpunktes verwendet wird.

4.2 Zerlegung und Klassifikation der Wellenform

Die Basis für ein Modell, das zusätzliche Freiheitsgrade zur Verfügung stellt mit denen es an Wellenformen angepasst werden kann, die durch Unterwasserstrukturen verursacht werden, wurde von SCHWARZ et. al (2017) vorgestellt. Dieses Modell, das ausschließlich aus Exponentialtermen besteht, wird um zwei Elemente erweitert und besteht damit aus den folgenden drei Elementen:

$$\sigma_{E}(z) = (u(z - z_{0}) - u(z - z_{0} - T))\beta_{E}e^{-K(z - z_{0})}$$

$$\sigma_{B}(z) = (u(z - z_{0}) - u(z - z_{0} - T))\beta_{B}$$

$$\sigma_{D}(z) = \delta(z - z_{0})\beta_{D}$$
(6)

 σ_E für Exponentialsegmente, σ_B für Rechtecksegmente und σ_D für impulsförmige Dirac Segmente. Das Model für die empfangene Wellenform lautet damit

$$p_r(t,\varphi) = \int_0^\infty p_s\left(t - 2\frac{n_w}{c_0}\right) \sum_i \sigma_{Mi}(z) \, dz. \tag{7}$$

Der Index *M* kann dabei die Werte *E*, *B*, *D* für die verschiedenen Teilmodelle annehmen und φ steht für ein Tupel dessen Elemente die Parameter aller Teilkomponenten sind.

Die Abtastwerte der empfangenen Wellenform an den diskreten Zeitpunkten t_n werden mit $y(t_n) = y_n$ bezeichnet. Die Zerlegung der Wellenform wird nun schrittweise durchgeführt, wobei in jedem Schritt das Modell um eine weitere Komponente ergänzt wird. Angenommen das Modell des m - ten Teilschrittes lautet $p_{r,m}(t, \varphi_m)$ dann wird zunächst eine um diesen reduzierte Wellenform

$$y_{m+1}(t_n) = y_m(t_n) - p_{r,m}(t_n, \varphi_m)$$
(8)

ermittelt, anhand der ein einzelnes Elementmodel nach Gleichung (6) ausgewählt und benutzt wird, um Parameterwerte zu schätzen. Diese Parameter werden dem Tupel φ_m angehängt, so dass daraus das neue Tupel φ_{m+1} entsteht. Mittels einer nichtlinearen "Least Squares"-Methode, wie zum Beispiel Levenberg-Marquardt, wird nun das gesamte neue Parameter Tupel angepasst

$$\varphi_{m+1,opt} = argmin_{\varphi_{m+1}} \sum_{n} [y(t_n) - p_{r,m+1}(t_n, \varphi_{m+1})]^2.$$
(9)

Die Iteration wird abgebrochen, wenn die Wurzel aus dem mittleren quadratischen Restfehler mit den Messfehlern in einer vergleichbaren Größenordnung steht oder wenn keine Verbesserung zwischen den Schritten erreicht wird. Im Unterschied zu SCHWARZ et. al (2017) erlauben wir in dem hier verwendeten Modell überlappende Positionierungen der Teilmodelle.

Die Parameter des resultierenden Tupels werden für die Erzeugung einer Punktwolke herangezogen. Die Parameter, die den z_0 entsprechen werden als Zielorte interpretiert und die Flächen der Teilmodelle werden als Amplituden verwendet. Die zusätzlichen Parameter *T*, *K* fließen als zusätzliche Attribute in die Klassifikation für die Oberfläche und den Boden ein. Der empirisch ermittelte Klassifikationsalgorithmus berücksichtigt zunächst nur Punkte, deren Attribute ausgedehnte Ziele (T > 0) mit großer Amplitude qualifizieren. Wenn es nur einen solchen Punkt gibt, gilt er als Wassersäule. Bei Vorliegen mehrerer Punkte wird der Punkt mit der zweithöchsten Amplitude für die Wassersäule verwendet, aber nur wenn er vor dem Punkt mit der höchsten Amplitude liegt. Die Wasseroberfläche wird dann als vordere Flanke der so bestimmten Wassersäule bestimmt. Für den Boden wird dann der Punkt mit der größten Amplitude aus den restlichen Punkten gewählt. Alle anderen Punkte bleiben ohne Klassifikation.

4.3 Wellenform als Pseudo-Punktwolke

Eine Betrachtung der Nachbarschaftsbeziehungen einzelner Wellenformprofile legt eine dreidimensionale Darstellung nahe, auf deren horizontaler Achse die Laufzeit und auf deren vertikaler Achse die laufende Nummer des Wellenformprofils aufgetragen ist. Die dritte Dimension wird durch Zuweisung eines Grauwertes gebildet, welcher der Amplitude des Abtastwertes der Wellenform entspricht. In Abb. 3 ist eine beispielhafte Darstellung von 100 Wellenformprofilen als Wellenformstapel zu erkennen.



Abb. 3: Dreidimensionale Wellenformdarstellung als Wellenformstapel. Oben: Abtastwerte eines Wellenformprofils. Unten: Aufeinanderfolgende Wellenformprofile mit Amplituden als Grauwerte kodiert. Dunkle Stellen entsprechen großer Amplitude.

Fügt man dieser Darstellung eine weitere, vierte, Dimension hinzu indem man die mit Graustufen eingefärbten Wellenformprofile aus dem Sensor Koordinatensystem (Abb. 3) in das Projektkoordinatensystem bringt, erhält man eine Darstellung gemäß Abb. 4 die Ähnlichkeiten mit einen Vorhang aufweist.



Abb. 4: Vierdimensionale (Vorhang-) Darstellung der Wellenformen als Pseudo Punktwolke. Dunkle Einfärbung bedeutet hohe Amplitude, helle Stellen niedrige Amplitude. Die Oberkante des Vorhangs entspricht der Wasseroberfläche.

5 Ergebnisse und Diskussion

5.1 Hafenboden

Hauptanwendung des BDF-1 ist die Vermessung des unter Wasser liegenden Geländes. Die vorliegenden Daten wurden aber nicht zum Zweck einer umfassenden Vermessung des Hafengrundes erfasst, sondern nur um das Potential des Sensors für diesen Einsatzzweck abzuschätzen.



Abb. 5: BDF - Boden Abweichung. Der Abstand ist nur an wenigen Stellen größer als 0.5 m. Tiefen bis ca. 5 m bzw. 6 m wurden erfasst.

Die komplexen Wellenformen, siehe Abb. 6, mit einer hohen Anzahl von lokalen Maxima lassen die zuverlässige Detektion des Bodenechos zu einem nicht trivialen Problem werden. Mit einem auf dem im Abschnitt 4.2 beschriebenen Klassifikationsalgorithmus basierenden Algorithmus wurden die Oberflächen- und Bodenechos ermittelt.

Als Referenz für die Bodenprofile des BDF-1 stand eine Echolotvermessung aus dem Jahr 2016 zur Verfügung. Leider war es nicht möglich, diese Messungen für einen absoluten Vergleich zu verwenden, da trotz Referenzierung auf MLLW (Mean Lower-Low Water), in beiden Fällen, ein ungeklärter vertikaler Versatz von mehr als 3m bestehen blieb, wie uns vom Operator mitgeteilt wurde. Wir haben deshalb den mittleren Abstand zwischen Echolotung und (refraktions-korrigierten) BDF-1 Bodenpunkten bestimmt und vor dem Vergleich in Abzug gebracht. Die Echolotungen wurden durch Resampling einer Triangulation mit der Dichte von 1Pkt./m² interpoliert, neuerlich trianguliert und mit einem Laplace Kernel geglättet. Die Dreiecke wurden mit einer der Tiefe unter dem Projektwasserspiegel entsprechenden Skala eingefärbt. Der Projektwasserspiegel wurde so festgelegt, dass er mit der Wasseroberfläche der BDF-1 Messung zusammenfällt. In Abb. 5 ist die Differenz der BDF-1 Bodenpunkte und der interpolierten Echolot Fläche als hauptsächlich weiße Linie zu sehen. Obwohl die Struktur des Bodens nicht sehr ausgeprägt ist, kann man insbesondere in der Nähe der Anlegestelle erkennen, dass die Verläufe

innerhalb einer Abweichung von 0,5 m einander folgen. Im Vergleich mit Abb.2 ist zu sehen, dass Tiefen größer als 5 m bis 6 m nicht mehr erfasst wurden.

5.2 Dämpfungskoeffizient

Bei Zerlegung der Wellenformen wie in Abschnitt 4.2 beschrieben, werden Dämpfungskoeffizienten *K* gewonnen. Abb. 6 zeigt ein Beispiel einer Zerlegung. Die Faltung der in Rot dargestellten Kurve mit der Systemwellenform (nicht dargestellt) ergibt die schwarze, durchgezogene Linie, in der Abbildung als Modell bezeichnet. Die Abtastwerte der Messung sind durch Kreise markiert, das Modell fällt in der Darstellung praktisch zusammen mit den Abtastwerten. Es ist zu erkennen, dass einem im Wesentlichen exponentiellen Abfall kleinere Abweichungen überlagert sind. Die vordere Flanke des exponentiellen Korpus definiert zugleich auch die Wasseroberfläche.



Abb. 6: Wellenformzerlegung. Das Modell entsteht aus der Faltung der Systemwellenform mit der Zerlegung.

Die Dämpfungskoeffizienten K haben die Dimension m^{-1} und ihre Kehrwerte Z = 1/K können als optische Tiefe interpretiert werden. Dieser Zusammenhang macht K bzw. Z interessant für die Messung der Schwebstoffe (siehe z.B. ZHAO et. al (2018)). Ermittelt man ein Histogramm aller im Datensatz enthaltenen optischen Tiefen, so erhält man zunächst eine Verteilung die zwei Maxima aufweist. Eine Unterteilung der optischen Tiefen in zwei Klassen, eine Klasse, die alle Z enthält, die von den Exponentialtermen stammen welche die Wasseroberfläche definieren, und eine Klasse, die alle anderen enthält, führt auf Abb. 7 (a) und (b).



Abb. 7: Histogramme Optischer Tiefen, (a) Wasseroberfläche, (b) Unterwasser

Für den vorliegenden Datensatz existieren keine Referenzmessungen der Wassergüte oder eine Bestimmung der Secchi-Tiefe. Beachtet man den Zusammenhang (GUENTHER 1985)

$$Z_s K = \eta \tag{10}$$

worin Z_s die Secci Tiefe ist und η ein Wert zwischen 1,1 und 1,7, der in komplexer Weise von den inherenten optischen Eigenschaften Dämpfung und Streuung des Wasserkörpers abhängt, so erhält man

$$Z_s = \eta Z. \tag{11}$$

Für Z_s erhält man so unter Berücksichtigung von Abb. 7 (a) Werte von 2,5 m bis 3 m. Aus den Beobachtungen aus Abschnitt 5.1 folgt, dass die erreichbare Tiefe ca. 5 m bis 6 m ist, entsprechend der doppelten Secchi-Tiefe, ein Wert, der mit den Erwartungen über die mögliche Performance des Sensors zusammenstimmt.

Die Verteilung der optischen Tiefen in Abb. 7 (b) weist im Unterschied zu (a) eine wesentlich höhere Asymmetrie zu größeren Werten auf, aber auch (a) zeigt eine nicht Gaußförmige Form. Während im Fall (a) die erreichte Tiefenperformance eine plausible Erklärung für den Median bzw. den Mittelwert liefert, haben wir für den Fall (b) keine Erklärung. Wir halten deshalb eine genauere systematische Untersuchung im Besonderen im Zusammenhang mit der im nächsten Abschnitt 5.3 behandelten Schichtstruktur für lohnenswert.

5.3 Schichtstruktur

Der untersuchte Datensatz zeigt Abweichungen, die der prototypischen Wellenform (Abb. 1) überlagert sind. Diese Abweichungen sind zunächst an der isoliert betrachteten Wellenform nicht von zufälligen Variationen zu unterscheiden. Ein Indiz dafür, dass es sich nicht um rein statistische, beispielsweise durch die Sensorelektronik verursachte, Störungen handelt, ist die Tatsache, dass die Abweichungen auch nach einer Summation von 100 aufeinanderfolgenden Wellenformen deutlich zu sehen sind.



Abb. 8: Vorhangdarstellung der ungemittelten Wellenformen. Der Zeitabstand zwischen den beiden Profilen ist 74 Minuten. Der Vorhangdarstellung überlagert sind die ermittelte Oberfläche (blau) und der Boden (braun)

Die Grunddaten in Abb. 6 zeigen beispielsweise so eine über 100 Datensätze gemittelte Wellenform, entsprechend einer Zeitspanne von 25 ms bzw. 5 cm. Fluktuationen des Mediums die schneller ablaufen oder kürzere Ausdehnung aufweisen, werden durch den Mittelungsprozess erheblich gedämpft. Das Vorhandensein der Abweichungen lässt deshalb vermuten, dass diese Abweichungen Strukturen des Mediums, also der Wassersäule, offenbaren. Abb. 8 zeigt zwei, wie im Abschnitt 4.3 beschrieben, erzeugte Wellenformvorhänge der ungemittelten Wellenformen, denen die ermittelte Oberfläche (blau) und der Boden (braun) überlagert wurden. Es ist eine deutlich ausgeprägte horizontale Schichtstruktur zu erkennen die sogar bereits vor der Mittelung zutage tritt und nicht durch den Messprozess verursacht scheint. Darüber hinaus sind auch sehr langsame Veränderungen zu beobachten. Die beiden Profile in Abb. 8 weisen einen Unterschied in der Aufnahmezeit von 74 Minuten auf und wurden deshalb ausgewählt, da am Kreuzungspunkt zu erkennen ist, dass sich die Schichtung während dieser Zeit verändert hat. An anderen Kreuzungspunkten ist diese Schichtung unverändert geblieben.

Im Unterschied zu den horizontalen Strukturen sind die vertikalen Strukturen offensichtlich mit dem Messprozess korreliert, sie treten entlang der Laserstrahlrichtung auf. Benachbarte Wellenformen sind dabei aber einander nicht einfach nur über eine Skalierung ähnlich, was auf eine Schwankung der Empfindlichkeit des Sensors hindeuten würde, sondern zeigen auch strukturelle Unterschiede.

Bei der Zerlegung der Wellenform, wie in Abschnitt 4.2 beschrieben, fallen neben der Klassifikation der Oberflächen und Bodenpunkte auch weitere diskrete Punkte und ihnen zugeordnete Amplituden an. Stellt man diese Punkte als mit der Amplitude eingefärbte Punktwolke dar, so erhält man Abb. 9, die denselben Bildausschnitt wie Abb. 8 zeigt. Zweifellos erhöht diese Darstellungsart als reine Punktwolke nicht unbedingt die Übersichtlichkeit, dennoch hat sie aber den Vorteil, dass der Einfluss der System-Wellenform kompensiert wurde und die sich abzeichnende Linienstruktur mit konkreten geometrischen Orten unter der Wasseroberfläche identifiziert werden kann.



Abb. 9: Punktwolkendarstellung der Wellenformzerlegung.

Eine interessante Erweiterung der Darstellung über die Punktwolke hinaus wäre die Darstellung der in der Zerlegung gefundenen Exponential- und Puls-Segmente als Lineare, ausgedehnte, Elemente. Wir erwarten dadurch eine deutlicher hervortretende Schichtstruktur, vergleichbar der in Abb. 8.

Die zutage tretenden Strukturen legen es nahe, quantitative Zusammenhänge mit der Schwebstoffkonzentration zu untersuchen. Es ist auch die Frage zu stellen ob es einen Zusammenhang mit dem Temperaturprofil gibt und ob die Messdaten für eine Bestimmung der Wassergüte verwendet werden können.

5.4 Offene See

Die Gültigkeit der Wellenformzerlegung nach Abschnitt 4.2 beruht unter anderem darauf, dass vorausgesetzt werden kann, dass das Superpositionsgesetz für die einzelnen Komponenten gilt. Bei sehr starker Aussteuerung des Empfängers, die bei großer optischer Tiefe oder Messung auf Land auftreten kann, ist die Voraussetzung für das Superpositionsgesetz aber nicht mehr erfüllt. Das obere Teilbild in Abb. 10 zeigt, dass das Bodenecho auf offener See so stark ist, dass der empfindliche Kanal, der für die Wellenformzerlegung herangezogen wird, so übersteuert wird, dass für die Messung ein zweiter im Gerät vorhandener, unempfindlicher Kanal, im Bild rot dargestellt, herangezogen wird. Trotz Übersteuerung werden die Wellenformen auch im empfindlichen Kanal weiter aufgezeichnet. Das Auftreten einer Übersteuerung kann nur Signalwerte verfälschen, die zeitlich gesehen nach dem Übersteuerungszeitpunkt liegen. Da das Signal von der Wasseroberfläche vor dem Übersteuerungszeitpunkt zufolge des Bodenechos empfangen wird, kann die Wasseroberfläche mit Hilfe der Wellenformzerlegung im empfindlichen Kanal bestimmt werden. Das untere Teilbild von Abb. 10 zeigt die solcherart bestimmte Wasseroberfläche, ohne die eine Brechungskorrektur der Bodenechos nicht möglich ist.



Abb. 10: Messprofile über offener See. In Rot das sehr starke Bodenecho. Oberes Teilbild: Wasseroberfläche wegen Über-steuerung nicht bestimmt. Unteres Teilbild: Durch teil-weise Aus-wertung des übersteuerten Kanals wird eine Bestimm-ung der Wasseroberfläche möglich.

6 Fazit & Ausblick

Mit dem von der Firma RIEGL vor zwei Jahren vorgestellten BDF-1, (Bathymetric Depth Finder) ist es möglich georeferenzierte Profile im Wasserkörper aufzunehmen. Die Komplexität der Wellenformen erschwert einerseits zwar die Bestimmung der Bodenechos, zeigt andererseits aber einen Detailreichtum der räumlichen Strukturen. Wir haben gezeigt, dass eine Schichtstruktur in der Wassersäule zu Tage tritt von der angenommen werden muss, dass sie in einem Zusammenhang mit tatsächlichen stationären Inhomogenitäten des Mediums steht und nicht durch Fluktuationen und Signalrauschen verursacht wird. Wir erwarten, dass quantitative Zusammenhänge mit der Konzentration von gelösten Sedimenten oder Schwebstoffe unter der Oberfläche hergestellt werden können und regen Experimente an, mit denen diese Gesetzmäßigkeiten aufgedeckt werden können.

7 Dank

Die vorgestellte Arbeit beruht auf Datenmaterial, das uns von "US Army Corps of Engineers Cold Regions Research and Engineering Laboratory, Remote Sensing/GIS Center of Expertise" (Adam L. LeWinter, Peter J. Gadomski und David C. Finnegan) zur Verfügung gestellt wurde. Wir danken unserer Kollegin Andrea Spitzer und den Kollegen Alexander Haring und Christian Sevcik für ihre Unterstützung.

8 Literaturverzeichnis

- GORDON, H. R., 1982: Biodiversity: Interpretation of airborne oceanic lidar: effects of multiple scattering. Applied Optics, **21**, 2996.
- GUENTHER, G. C., 1985: Airborne laser hydrography system design and performance factors. National Oceanic and Atmospheric Administration Rockville MD, USA. https://apps.dtic.mil/dtic/tr/fulltext/u2/a488936.pdf.
- LEWINTER, A., PFENNIGBAUER, M., GADOMSK, P. J., FINNEGAN, D. C., SCHWARZ, R., TRUONG, M.-L. & PODOSKI, J. H., 2018: Unmanned aircraft system-based lidar survey of structures above and below the water surface. Hilo Deep Draft Harbor Breakwater, Hawaii. Lidar Remote Sensing for Environmental Monitoring XVI, International Society for Optics and Photonics, 10779, 107790Q.
- MALLET, C. & BRETAR, F., 2008: Full-waveform topographic lidar: State-of-the-art. ISPRS Journal of Photogrammetry and Remote Sensing, **64**, 1-16.
- PORTLAND, D. C. O. E., 2016: Condition Survey, Hawaii Hilo Harbour.
- SCHWARZ, R., PFEIFER, N., PFENNIGBAUER, M. & ULLRICH, A., 2017: Exponential Decomposition with Implicit Deconvolution of Lidar Backscatter from the Water Column. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 85, 159-167.
- WAGNER, W., ULLRICH, A., DUCIC, V., MELZER, T. & STUDNICKA, N., 2006: Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. ISPRS Journal of Photogrammetry and Remote Sensing, **60**, 100-112.
- ZHAO, X., ZHAO, J., ZHANG, H. & ZHOU, F., 2018: Remote Sensing of Suspended Sediment Concentrations Based on the Waveform Decomposition of Airborne LiDAR Bathymetry. Remote Sensing, 10, 247.

Single Photon LiDAR – Grundlagen und erste Evaluierungsergebnisse

GOTTFRIED MANDLBURGER^{1,2} & HUBERT LEHNER³

Zusammenfassung: In diesem Beitrag werden die Grundlagen von Single Photon LiDAR (SPL) erläutert und die technologischen Unterschiede zu konventionellem Laserscanning aufgezeigt. Anhand der 3D Punktwolken zweier Messkampagnen, die im Juli bzw. September 2018 in Wien unabhängig mit beiden Sensorsystemen erfasst wurden, werden erste Evaluierungsergebnisse präsentiert. Es zeigt sich, dass die Höhengenauigkeit bei SPL etwa um einen Faktor 2 geringer ist als bei konventionellem Laserscanning. Die planimetrische Genauigkeit, welche wesentlich von der Größe des Laserabtastflecks abhängt, ist beim SPL wegen der geringen Strahldivergenz (0.08 mrad) selbst bei höherer Flughöhe (4.000 m) nur unwesentlich größer als bei typischen Erfassungen mit konventionellen ALS Systemen. Die große Messdistanz erfordert allerdings eine hoch-genaue IMU. Eine Überlegenheit von SPL gegenüber konventionellen Systemen besteht bei der Flächenleistung in Folge der simultanen Messung von 100 Teilstrahlen, der daraus resultierenden hohen Messrate (6 MHz) und der großen Flugstreifenbreite (>2.000m).

1 Einleitung

In den vergangenen Jahren sind im Bereich des Flugzeuglaserscannings neben etablierten Systemen auch Geräte auf den Markt gekommen, die eine Entfernungsmessung beim Eintreffen eines oder weniger Photonen erlauben (single photon sensitivity). Hierbei haben mittlerweile zwei Systeme auch in der kommerziellen Nutzung Bedeutung erlangt: Beim sogenannten Geiger-mode Li-DAR (GmLiDAR, LiDAR=Light Detection and Ranging) wird mit der Reflexion eines stärker divergenten Laserpulses eine Matrix von Einzelphoton-sensitiven Elementen (Geiger-mode Avalanche Photo Diode Array, GmAPD) beleuchtet (KIM et al. 2013; STOKER et al. 2016). Bei der als Single Photon LiDAR (SPL) bezeichneten Technologie wird ein hoch kollimierter Laserimpuls durch ein diffraktives optisches Element in 10x10 Teilstrahlen, sogenannte beamlets, aufgespaltet (DEGNAN 2016). Da für jedes beamlet ein Einzelphoton-sensitives Detektor-Array zum Einsatz kommt, das sich in gewissen Aussteuerbereichen ähnlich einer APD im Linearbetrieb verhält, ist die SPL Technologie mehrzielfähig, was vor allem für die Durchdringung von Vegetation von großer Bedeutung ist. Durch die hohe Sensorempfindlichkeit können beide Systeme aus großer Flughöhe (4.000-10.000 m) betrieben werden und ermöglichen dadurch gegenüber konventionellen Laserscannern bei vergleichbarer Punktdichte eine höhere Flächenleistung (800-1.000 km²/h bei 8 Punkten/m²) zum Preis einer geringeren Höhengenauigkeit (~10 cm) und einer höheren Rate an Ausreißer-Punkten (PFENNIGBAUER & ULLRICH 2018; ULLRICH & PFENNIGBAUER 2016).

¹ Universität Stuttgart, Institut für Photogrammetrie, Geschwister-Scholl-Str. 24D, D-70174 Stuttgart, E-Mail: gottfried.mandlburger@ifp.uni-stuttgart.de

² TU Wien, Department für Geodäsie und Geoinformation, Gußhausstraße 27-29/E120, A-1040 Wien

³ Stadtvermessung Wien, Magistratsabteilung 41, Muthgasse 62, A-1190 Wien

Die höhere Flächenleistung Einzelphoton-basierter Systeme ist vor allem für großflächige Topographieerfassung in der Landesvermessung von Interesse. So kommen sowohl GmLiDAR als auch SPL bereits beim 3D Elevation Program (SUGARBAKER et al. 2014), einer Initiative zur flächendeckenden Höhendatenerfassung in den USA zum Einsatz und werden dort hinsichtlich der Erfüllung der geforderten Genauigkeitsstandards evaluiert (STOKER et al. 2016). Während das Geigermode System von Harris Corporation als Dienstleistung vornehmlich in Amerika betrieben wird, ist das Single Photon LiDAR System seit der Übernahme des Sensorherstellers SigmaSpace Corporation durch Hexagon/Leica Geosystems unter dem Namen SPL100 kommerziell verfügbar. Bei Letzterem handelt sich um ein hybrides Messsystem, bestehend aus einem Navigationssystem (GNSS/IMU), dem Laserscanner und einer 80 Megapixel RGBI Kamera (RCD30). Der Sensor ist auf effiziente großflächige Erfassung von Topographie optimiert, eignet sich aber wegen der eingesetzten grünen Wellenlänge (λ =532 nm), dem kreisförmigen Scanmuster mit einer Neigung von wahlweise 10°, 15°, 20° oder 30°4 gegenüber der Nadirrichtung und der hohen Detektorempfindlichkeit grundsätzlich auch für bathymetrische Anwendungen bei sehr klaren und seichten Gewässern (DEGNAN 2016)⁵. In Europa ist gegenwärtig ein Sensor stationiert. Dieser wird in einigen Ländern (z.B. Spanien, Finnland, Niederlande, Österreich) im Rahmen von Pilotprojekten erprobt. So ist beispielsweise die gesamte Fläche der nordspanischen Provinz Navarra mit dem SPL100 erfasst worden. Diese Daten wurden kürzlich öffentlich zugänglich gemacht⁶. In Österreich wurden unter anderem im Auftrag der Stadt Wien Testflächen aufgenommen.

In diesem Beitrag werden zunächst die Grundlagen der Einzelphoton-basierten Techniken erläutert und die technologischen Unterschiede zu konventionellem Laserscanning aufgezeigt (Abschnitt 2). Daran anschließend werden die Messkampagnen vom August/September 2018 vorgestellt, bei welchen ein Teil des Stadtgebiets von Wien mit einem Leica SPL100 Sensor und einem RIEGL VQ-1560i erfasst wurden (Abschnitt 3). Im Abschnitt 4 erfolgt die Beschreibung der Auswertemethoden für die Datenevaluierung, und die entsprechenden Ergebnisse werden in Abschnitt 5 präsentiert und diskutiert. Der Beitrag endet mit einer Zusammenfassung der wesentlichen Erkenntnisse und einem Ausblick (Abschnitt 6).

2 Funktionsprinzipien

In diesem Abschnitt werden die prinzipielle Funktionsweise von konventionellem Flugzeuglaserscanning (Abschnitte 2.1 und 2.2) sowie von Einzelphoton-basiertem LiDAR (Abschnitte 2.4 und 2.4) erläutert.

2.1 Grundlagen

Beim Flugzeuglaserscanning handelt es sich um ein dynamisches, polares, aktives Multisensor-Messsystem bestehend aus der Navigationseinheit (Global Navigation Satellite System, Inertial

⁴ <u>https://leica-geosystems.com/products/airborne-systems/topographic-lidar-sensors/leica-spl100</u>

⁵ Trotz der grundsätzlichen bathymetrischen Fähigkeiten sei an dieser Stelle darauf hingewiesen, dass der Hersteller das System ausschließlich zur effizienten 3D-Kartierung der Überwasser-Topographie propagiert.

⁶ Siehe: <u>ftp://ftp.cartografia.navarra.es/5_LIDAR/5_4_2017_NAV_ca_EPSG25830/</u>

Measurement Unit) zur Bestimmung der Raumlage und –stellung der Messplattform und dem Laserscanner, welcher Auslenkwinkel des Laserstrahls und Entfernungen zu den die Laserstrahlung reflektierenden Zielen misst (PFEIFER et al. 2015). Zur Entfernungsmessung kommt das Impuls-Laufzeit-Verfahren zur Anwendung, bei dem die Zeitdifferenz zwischen emittiertem Laserpuls und Eintreffen des reflektierten Echos am Empfänger gemessen wird. Bei bekannter Ausbreitungsgeschwindigkeit kann damit auf die Distanz zwischen Sensor und Ziel geschlossen werden. Eine Laserstrahl-Ablenkeinrichung (schwingender, rotierender oder nutierender Spiegel) sorgt dabei zusammen mit der Vorwärtsbewegung der Plattform für die systematische Abtastung der Erdoberfläche.

Diese grundsätzlichen Eigenschaften gelten sowohl für konventionelles wie auch für Einzelphoton-basiertes Laserscanning. Der wesentliche Unterschied besteht in der Detektortechnologie, die in den folgenden Unterabschnitten für die einzelnen Systeme näher erläutert ist. Einen grundsätzlichen Überblick über die Situation der eintreffenden Laserstrahlung aus der Sicht des Detektors gibt die Abb. 1.



Abb. 1: Prinzipskizze: Eingehende Laserstrahlen aus Sicht des Empfängers für konventionelles, Geigermode und Single Photon LiDAR

2.2 Konventionelles ALS – Linear-mode LiDAR

Bei der klassischen Impuls-Laufzeit-Laserentfernungsmessung, die gegenwärtig als Abgrenzung zum Geiger-mode auch als "linear-mode LiDAR" bezeichnet wird, trifft die reflektierte Strahlung eines hoch kollimierten Laserpulses auf genau einen Detektor. Das Gesichtsfeld (FOV) des Empfängers ist dabei wesentlich größer als der Durchmesser des reflektierten Laserstrahls (siehe Abb. 1 links). Die Umwandlung der optischen Strahlungsleistung in digitale Form erfolgt dabei zweistufig (ULLRICH & PFENNIGBAUER 2016): Zunächst wird die empfangene Laserstrahlung mittels einer Photodiode (Avalanche Photo Diode, APD) in ein analoges Signal umgewandelt, welches dann über einen Analog-Digital-Wandler (Analog Digital Converter, ADC) in digitale Information konvertiert wird. Die Bezeichnung "linear mode" geht dabei auf den Betriebsbereich der APD zurück, in welchem ein linearer Zusammenhang zwischen optischer Leistung und ausgehendem Signal besteht. Es sei der Vollständigkeit darauf hingewiesen, dass bei modernen Sensoren zur Erhöhung des erfassbaren Dynamikbereichs mehrere APDs/ADCs mit unterschiedlicher Empfindlichkeit zum Einsatz kommen (WAGNER et al., 2004). Jedenfalls liefern linear-mode APDs Informationen über die empfangene Signalstärke und ermöglichen durch radiometrische Kalibrierung Rückschlüsse auf die Reflektivität bzw. Materialeigenschaften der beleuchteten Objekte (LEHNER & BRIESE 2010; BRIESE et al. 2012; PFEIFER et al. 2015; WAGNER 2010).

Die eigentliche Zielerfassung erfolgt entweder durch entsprechende Bauteile im Laserscanner selbst (Diskret-Echo-Systeme) oder durch zeitlich hochauflösende Abtastung der gesamten rückgestreuten Signalenergie (Full Waveform, FWF). Die erfassten Wellenformen werden entweder direkt in der Firmware des Sensors verarbeitet (PFENNIGBAUER et al. 2014) oder zur detaillierten Analyse im Postprocessing gespeichert (PFEIFER et al. 2015). FWF Erfassung und Aufzeichnung ist derzeit jedenfalls ausschließlich für konventionelle Sensoren möglich, mit entsprechenden Vorteilen für die Genauigkeit der Entfernungsmessung, der Trennbarkeit von einzelnen Objekten innerhalb des Strahlkegels und der Beschreibung der Objekteigenschaften (Ausdehnung, Neigung, Reflektivität). Für die zuverlässige Detektion eines einzelnen Objekts sind jedenfalls einige 100 Photonen erforderlich, wobei für topographische Anwendungen vornehmlich Laserwellenlängen im nahen Infrarot (λ =1064/1550 nm) eingesetzt werden. Selbst sehr empfindliche Einzeldetektorsysteme, wie sie etwa in der Laserbathymetrie bei einer Wellenlänge im sichtbaren Bereich des Spektrums (λ =532 nm) zur Anwendung kommen, benötigen ca. 250 Photonen für die Erfassung eines Objekts (MANDLBURGER & JUTZI 2018). Die erzielbare Punktdichte hängt dabei neben Fluggeschwindigkeit und Reflektivität der beleuchteten Objekte ausschließlich von der Pulswiederholrate und dem Scanmechanismus ab.

2.3 Geiger-mode LiDAR

Zur Steigerung der Flächenleistung kommt beim GmLiDAR eine Matrix von Einzelphoton-sensitiven Detektoren, ein sogenanntes Geiger-mode Avalanche Photo Diode (GmAPD) Array, zum Einsatz (STOKER et al. 2016). Die APD jeder einzelnen Detektorzelle ist dabei im Geiger-mode, d.h. durch Anlegung einer Vorspannung über der Durchbruchsspannung in einen Modus versetzt, bei dem bereits das Eintreffen eines oder weniger Photonen den Lawineneffekt auslöst, der zu einem sprunghaften Ansteigen der Spannung am Ausgang des Detektors führt (ULLRICH & PFENNIGBAUER 2016). Dadurch kann jede Zelle als binärer Detektor angesehen werden, wobei das Eintretens des Lawineneffekts (=Durchbrechen der Photodiode) über einen Time-To-Digital (TDC) Konverter den Stopimpuls für die Laufzeitmessung auslöst. Nach dem Durchbrechen einer Zelle ist diese infolge der resultierenden Ladungsflut für längere Zeit inaktiv. Insbesondere kann nach der Detektion eines ersten Echos für diese Zelle kein weiteres Echo vom selben Laserimpuls detektiert werden. Erst nach einem Reset des gesamten Detektorarrays im Zuge der Emission des nächsten Laserpulses sind die ggf. durchgebrochenen Detektorzellen wieder aktiv.

Durch die Einzelphoton-Empfindlichkeit der Detektorzellen kann die Flughöhe wesentlich gesteigert werden. Die damit verbundene potenzielle Reduktion der Punktdichte wird durch das Detektorarray kompensiert, die als range camera aufgefasst werden kann. Der relativ große Laserabtastfleck am Boden wird also von der Detektormatrix räumlich aufgelöst erfasst. Das Gesichtsfeld des Detektors ist dabei so gestaltet, dass alle 32×128=1024 Detektorzellen beleuchtet werden (siehe Abb. 1, Mitte). Für die planimetrische Auflösung ist demnach (in Abhängigkeit der Flughöhe) die Größe der Detektorzelle maßgeblich und nicht wie bei konventionellen Systemen die Größe des

Laserabtastflecks am Boden. Bei dem von Harris Corporation betriebenen System beispielsweise beträgt der Punktabstand 35 cm bei einer nominellen Flughöhe von 10 km und einer Größe einer Detektorzelle von 35 µm. Dies setzt aber voraus, dass es zu keinen Spontanauslösungen in der Atmosphäre kommt und jede Zelle tatsächlich auslöst. Beides ist in der Praxis nicht der Fall. Allerdings kommt es zu einer weiteren Steigerung der Punktdichte, da sich die relativ großen Footprints benachbarter Laserimpulse in der Regel überlappen. Im offenen Bereich kann die hohe Punktdichte zur Glättung der verfahrensbedingt größeren Entfernungsmessfehler herangezogen werden. Vegetationsdurchdringung ist wegen der mangelnden Mehrzielfähigkeit nur eingeschränkt möglich. Abschließend sei erwähnt, dass das Harris GmLiDAR System für effiziente Topographieerfassung optimiert ist und mit einer Wellenlänge im nahen Infrarot operiert (λ =1064 nm). Radiometrische Information kann durch den binären Charakter des Detektors nicht gemessen, sondern lediglich aus der lokalen Punktdichte geschätzt werden.

2.4 Single Photon LiDAR

Bei der als Single Photon LiDAR bezeichneten Technologie kommen ebenfalls Einzelphoton-sensitive Detektoren zum Einsatz. Anders als beim GmLiDAR, wird Laserstrahlung im sichtbaren grünen Bereich des Spektrums (λ =532 nm) verwendet, womit SPL inhärente bathymetrische Fähigkeiten aufweist. Zur Erhöhung der Punktdichte wird der Laserstrahl durch ein diffraktives optisches Element in 10x10 Teilstrahlen, sogenannte beamlets, aufgespalten (DEGNAN, 2016). Die 100 simultan emittierten, einander aber nicht überlappenden Teilstrahlen interagieren mit der Oberfläche, und jeder Teilstrahl wird von einem eigenen, auf die jeweilige Raumrichtung des Laserstrahls ausgerichteten Detektor, empfangen (siehe Abb. 1, rechts). Jeder Einzeldetektor besteht dabei wiederum aus einer Matrix von Einzelphoton-sensitiven Zellen, wobei jede dieser Zellen durch entsprechende Vorspannung in den Geiger-mode versetzt wird. Die genaue technische Realisierung ist dabei vom Sensorhersteller nicht offengelegt. Mögliche Implementierungen sind Micro Channel Plate PhotoMultiplier Tubes (MCP-PMT) oder Silicon PhotoMultipliers (SiPM) (AGISHEV et al. 2013; DEGNAN 2002)

Während die Raumrichtung für jeden einzelnen Teilstrahl bekannt ist, steht am Ausgang des jeweiligen Teilstrahl-Detektors ein Summensignal zur Verfügung. Mit anderen Worten, eine weitere Verfeinerung der räumlichen Auflösung durch die einzelnen Detektorzellen ist nicht möglich. Eine Abtastung des Summensignals am Ausgang des Detektors, und damit die Aufzeichnung der vollen Wellenform, ist derzeit noch nicht möglich, die Signalamplitude kann aber im Gegensatz zum GmLiDAR gemessen werden. Der radiometrische Informationsgehalt ist allerdings nicht vergleichbar ist zu jenem konventioneller (multi-photon) APDs, die im linearen Modus betrieben werden.

Die Einzelphoton-Empfindlichkeit erlaubt aber auch für SPL größere Flughöhen und damit potenziell größere Flächenleistungen. Dies ist, wie bereits eingangs erwähnt, für großflächige Datenerfassung etwa in der Landesvermessung von Interesse. Wegen der Verfügbarkeit eines entsprechenden Sensors in Europa und der daraus resultierenden höheren Relevanz für den deutschsprachigen Raum, wird der weitere Fokus auf den Vergleich von SPL und konventionellem FWF Laserscanning (forthin: Waveform LiDAR) gelegt.

3 Datensätze

Für die Erprobung der Single Photon LiDAR Technologie beauftragte die Stadt Wien (Magistratsabteilung 41 – Stadtvermessung) die Firma COWI im Rahmen eines Pilotprojektes mit der Erfassung eines Testgebiets. Der aufgenommene Flugblock ist etwa 8 km breit und 20 km lang und reicht von der Lobau östlich der Donau über den Bereich der Wiener Innenstadt bis in den westlich angrenzenden Wiener Wald (siehe Abb. 2a). Die geforderte Punktdichte war 20 Punkte/m² im Einzelstreifen (letzte Echos). Die Befliegung wurde am 29.07.2018 mit dem Hexagon/Leica Geosystems SPL100 Sensor aus einer Flughöhe von ca. 4.000 m durchgeführt. Für die Datenerfassung wurde ein Palmer Scanner (kreisrundes Punktmuster) mit einem Öffnungswinkel des Scankegels von 15° eingesetzt, woraus sich eine Streifenbreite von 2.000 m ergibt. Bei einer theoretischen Querüberlappung von 20% sind daher 5 Flugstreifen erforderlich, um das Gebiet abzudecken. Um Scanschatten im Stadtgebiet zu reduzieren wurde das Gebiet mit einer Querüberlappung von >50% also mit 10 Flugstreifen beflogen.



Abb. 2: Übersicht Flugblock Wien. (a) Flugstreifen Single Photon LiDAR (SPL100), rote Rechtecke: Detailuntersuchungsgebiete der Abb. 3, (b) Flugstreifen Waveform LiDAR (VQ-1560i) als Teil eines größeren Flugblocks

Als Vergleichsdatensatz steht eine Aufnahme desselben Gebiets mit einem konventionellen Full Waveform Laserscanner zur Verfügung, welche am 20.09.2018 bei vergleichbarem Belaubungszustand mit dem RIEGL VQ-1560i durchgeführt wurde. Der gesamte Flugblock ist in Abb. 2b dargestellt. Die Zielpunktdichte im Einzelstreifen lag dabei ebenso bei 20 Punkten/m². Auch hier betrug die Querüberlappung der Flugstreifen 50%. Zur Erfassung des Gebietes wurden 18 Längsstreifen mit einer Geschwindigkeit von 120 Knoten (60 m/s) aus einer Flughöhe von 750 m über Grund beflogen. Der Scanner arbeitet mit zwei Laserquellen, deren Pulse über ein gemeinsames rotierendes Polygonrad abgelenkt werden und dabei zwei vertikale Scanebenen mit einem Scanwinkelbereich von jeweils $\pm 30^{\circ}$ aufspannen⁷. Die x-förmig angeordneten Scanlinien sind jeweils um $\pm 14^{\circ}$ gegenüber der Querachse gedreht, sodass der gesamte Scanbereich quer zur Flugrichtung

⁷ http://www.riegl.com/uploads/tx_pxpriegldownloads/RIEGL_VQ-1560i_Datasheet_2018-10-16.pdf

58° beträgt. Die Streifenbreite ergibt sich damit zu 840 m. Die einzelnen Ziele (Echos) werden vom Scanner über online Verarbeitung der vollen Echowellenform ermittelt. Zu jedem Echo stehen neben der Geometrie auch zusätzliche Attribute wie Signalamplitude, Reflektivität und ein Qualitätsmaß (pulse shape deviation) zur Verfügung.

Für die weiteren Analysen wurden aus dem Gesamtgebiet zwei 400×400 m² große repräsentative Bereiche extrahiert und detailliert untersucht (siehe Abb. 3). Dabei handelt es sich um die in Abb. 2a eingezeichneten, rot markierten Bereiche. Die östliche Testfläche ist ein typisches Stadtgebiet mit großen Wohnhausblöcken, Straßenfluchten und einer Kirche mit einem 70 m hohen Turm (Abb. 3b). Anhand dieses Datensatzes kann einerseits die Erfassung der engen Straßenschluchten, aber auch die innere Messgenauigkeit an glatten Flächen (Straßen, Dächer, etc.) untersucht werden. Das westlich gelegene Gebiet ist eine Waldfläche mit unterschiedlich hohen Bäumen, Forststraßen, einer größeren offenen Fläche und einem alleinstehenden Gebäude (Abb. 3a). Dieser Datensatz eignet sich vor allem zur Analyse der Penetrationsfähigkeit von SPL. Die beiden Gebiete sind im Folgenden mit "Stadt" und "Wald" bezeichnet.



Abb. 3: Schummerungen des Digitalen Höhenmodells der Detailuntersuchungsgebiete Wald (a) und Stadt (b). Die zugrunde liegenden Daten stammen aus der Waveform LiDAR Befliegung

4 Auswertemethoden

Für die im nächsten Abschnitt dargestellten Ergebnisse wurden jeweils die folgenden Verarbeitungsschritte für beide Testgebiete (Stadt, Wald) und für beide Datensätze (SPL100, VQ-1560i) ausgeführt:

- Visuelle Kontrolle im 3D Punktviewer: Für den SPL100 Datensatz wurden dabei die ungefilterten Punkte inklusive der Streupunkte oberhalb und unterhalb des Geländes herangezogen.
- Durchdringungsfähigkeit: Die Evaluierung erfolgt in diesem Fall qualitativ durch visuelle Begutachtung von Profilschnitten im bewaldeten Bereich.

- Punktdichteabschätzung: Dazu wurden die streifenweisen Punktwolken jeweils in ein Raster mit einer Kantenlänge von 2,5 m einsortiert und pro Zelle die Punktdichte berechnet. Visualisierungen und quantitative Auswertungen wurden im Anschluss für den Einzelstreifen und den Gesamtblock erstellt.
- Streifenhöhendifferenzen: Die Überprüfung der Passgenauigkeit erfolgte durch Berechnung von Streifen-Höhenmodellen (DHM), deren Klassifizierung in raue und glatte Bereiche anhand der Standardabweichung der Gitterpunktinterpolation (gleitende Schrägebene / moving planes) und die anschließende Berechnung von Differenzmodellen für glatte Bereich der Streifenüberlappung.
- Untersuchung der Nachbarschaftsgenauigkeit bzw. des Messrauschens: In den glatten Bereichen gibt die Standardabweichung der Gitterpunktinterpolation Aufschluss über das Messrauschen. Die Residuen zur ausgleichenden Schrägebene kennzeichnen hier die relative Genauigkeit der Punktmessung. Die Auswertung erfolgt dabei durch statistische Analyse (Histogramm, statistische Kennzahlen).

5 Ergebnisse und Diskussion

In diesem Abschnitt werden die Ergebnisse der Datenevaluierung und des Vergleichs zwischen dem Single Photon LiDAR Datensatz (SPL100) und dem Waveform LiDAR Datensatz (VQ-1560i) anhand von Abbildungen vorgestellt und diskutiert.

Durch die Einzelphoton-Sensitivität der SPL Technologie können Entfernungen zu Objekten selbst bei sehr schwachen Signalen gemessen werden. Dies ermöglicht größere Flughöhen und in weiterer Folge eine größere Streifenbreite als dies bei konventionellem Laserscanning der Fall ist. Die Flugblockübersicht der Abb. 2 verdeutlicht dies. Während für die Erfassung des Projektgebietes mit dem Waveform LiDAR Sensor bei 50%-iger Streifenüberlappung 18 Fluglinien erforderlich waren, reichen beim SPL System selbst bei effektiver Überlappung von 65% 10 Flugstreifen aus. Die SPL Flugstreifen sind dabei mehr als doppelt so breit wie die Waveform LiDAR Streifen (2.000 m vs. 840 m). Für beide Befliegungen wurde die Flugplanung auf eine Zielpunktdichte von 20 Punkten/m² ausgelegt. Diese wird beim SPL System trotz größerer Flughöhe (4.000 m vs. 750 m) durch eine höhere Scanrate von 5 MHz erreicht (Pulsrepetitionsrate: 50 kHz, 100 Teilstrahlen pro Puls).



Abb. 4: Ungefilterte SPL100 3D-Punktwolke des Untersuchungsgebietes Stadt eingefärbt nach Intensität [DN]; (a) Perspektivansicht des gesamten Bereichs (b) Profilschnitt Kirche

Um aus solch großen Flughöhen Entfernungen zu Geländepunkten messen zu können, kommen die in Abschnitt 2.4 beschriebenen Einzelphoton-sensitiven Detektoren zum Einsatz. Durch die hohe Empfindlichkeit werden im praktischen Einsatz allerdings nicht nur Nutzpunkte (Gelände, Vegetation, Gebäude, etc.) sondern auch Streupunkte, die von Reflexionen aus der Atmosphäre stammen, erfasst. Diese wesentliche Charakteristik von 3D SPL Punktwolken zeigt die Abb. 4 für das städtische Untersuchungsgebiet in einer Perspektivansicht des gesamten Ausschnitts (a) sowie einem Profilschnitt im Bereich der Kirche. Die Punkte sind dabei nach Signalintensität eingefärbt, wobei zu erkennen ist, dass die meisten Streupunkte niedrige Intensität aufweisen und ihre Dichte geringer ist als die der Nutzpunkte. Diese beiden Eigenschaften stellen die Grundlage für die erforderliche Datenfilterung dar (DEGNAN 2016). Erheblich schwieriger ist die Bereinigung all jener Punkte mit hoher Intensität, die knapp unter der Oberfläche liegen. Gängige DGM Filterverfahren versagen in diesem Fall, da tief liegende Punkte prinzipiell bevorzugt werden.

Alle weiteren Untersuchungen beruhen auf der gefilterten Punktwolke, welche auch das Lieferprodukt der Befliegungsfirma darstellte. Die Abb. 5 zeigt die bereinigte SPL sowie die Waveform LiDAR Punktwolke des städtischen Untersuchungsgebietes wiederum im Gesamten und für den Detailbereich der Kirche. Dabei ist die größere Prägnanz (Abbildungsschärfe) der Waveform Li-DAR Daten klar ersichtlich. Besonders ausgeprägt ist dies am Beispiel des Kirchturms, bei welchem architektonische Details im Waveform LiDAR Datensatz (Abb. 5d) abgebildet sind, in der SPL Punktwolke (Abb. 5c) aber fehlen. Deutlich zu erkennen ist dies auch am Profilschnitt in Abb. 5e, in welchem die SPL Punkte blau und die Waveform LiDAR Punkte braun dargestellt sind.



Abb. 5: Gefilterte 3D-Punktwolken eingefärbt nach Intensität [DN], Untersuchungsgebiet Stadt mit Detailausschnitt Kirche; (a+c) Single Photon LiDAR, (b+d) Waveform LiDAR, (e) Profilschnitt: SPL (blau), Waveform LiDAR (braun)



Abb. 6: 3D-Punktwolke, Ausschnitt aus Untersuchungsgebiet Wald; (a) SPL, (b) Waveform LiDAR, (c) Profilschnitt: SPL (blau), Waveform LiDAR (hellbraun)

Abb. 6 zeigt die 3D Punktwolke eines Ausschnittes aus dem Untersuchungsgebiet Wald und hat die Evaluierung der Durchdringungsfähigkeit von SPL zum Thema. Prinzipiell ist die Technologie mehrzielfähig, im konkreten Vergleichsdatensatz ist die Abdeckung mit Bodenpunkten allerdings deutlich geringer als beim Waveform LiDAR Datensatz. Dies ist vor allem im Profilschnitt (Abb. 6c) zu erkennen, in welchem über weite Bereiche lediglich die braunen Punkte des konventionellen Systems vorhanden sind. Die Belaubungssituation kann dabei bei beiden Erfassungszeitpunkten (SPL: Juli, Waveform LiDAR: September) als äquivalent angesehen werden. Die bessere Fähigkeit zur Vegetationsdurchdringung konnte auch durch ein höheres Maß an Mehrfachechos von max. 5 Echos/Puls beim konventionellen System gegenüber max. 2 Echos/Puls beim SPL System quantitativ erfasst werden.



Abb. 7: Punktdichte [Punkte/m2] Untersuchungsgebiet Stadt; (a) SPL, (b) Waveform LiDAR

Abb. 7 zeigt die Gesamtpunktdichtekarte für das städtische Untersuchungsgebiet für beide Befliegungen. Die angestrebte Punktdichte von 20 Punkten/m² im Einzelstreifen wurde beim SPL Datensatz nicht erreicht, was eher auf die Flugplanung (Fluggeschwindigkeit, Scannerdrehrate) zurückzuführen ist als auf die Technologie. Die Punktdichtekarte erlaubt aber dennoch interessante Einblicke. So ist aus Abb. 7a für den SPL Datensatz eine deutlich erhöhte Dichte an den Hausfassaden zu erkennen, welche auf die konsequent schräge Scanrichtung zurückzuführen ist. Der kreisförmige Scanmechanismus (Palmer Scanner) ermöglicht dabei neben Seitblicken auch Vor- und Rückblick in der Mitte des Scanstreifens. Der Waveform LiDAR Datensatz weist hier immer wieder Stellen mit hellgrünem Farbton (10-20 Punkte/m²) auf. Andererseits sind in diesem Datensatz die Straßenfluchten mit konstant hoher Punktdichte (>40 Punkte/m² im Flugblock) erfasst.

Als weiteres Element einer Standard ALS Qualitätskontrolle wurden auch die relativen Höhendifferenzen in den Streifenüberlappungsbereichen berechnet und analysiert. Abb. 8 zeigt die Ergebnisse für beide Detailuntersuchungsgebiete und beide Datensätze. Während die Streifendifferenzen im Waldgebiet vollflächig dargestellt sind (Abb. 8a/b), sind diese in den Abb. 8c/d für das Stadtgebiet lediglich in den glatten Bereichen ausgewiesen. Nur in den glatten Bereichen können zuverlässige Aussagen über die Passgenauigkeit getroffen werden. Die statistischen Histogrammauswertungen der Abb. 8e/f beziehen sich daher auf die maskierten Höhendifferenzen der Abb. 8c/d. Sowohl die farbkodierten Differenzkarten als auch die Histogramme weisen für den SPL Datensatz höhere Abweichungen auf. Die Standardabweichung der Höhendifferenzen liegt für SPL bei 6 cm im Vergleich zu 2 cm für den Waveform LiDAR Datensatz. In den SPL Daten zeigt sich im städtischen Untersuchungsgebiet auch ein geringer systematischer Offset von ca. 1 cm (Mittelwert: 1,2 cm, Median: 1,3 cm). Ein Faktor 3 im Genauigkeitspotenzial zwischen konventionellem und Single Photon LiDAR wurde auch bereits von MANDLBURGER & JUTZI (2018) anhand eines Datensatzes im spanischen Navarra berichtet.



Abb. 8: Streifenhöhendifferenzen [m]: obere Reihe: Single Photon LiDAR, untere Reihe: Waveform LiDAR; (a+b) Untersuchungsgebiet Wald, ungefiltert, (c+d) Untersuchungsgebiet Stadt, raue Bereiche gefiltert, (e+f) Histogramme der der Höhendifferenzen zu (c+d)

Abschließend zeigt die Abb. 9 die Ergebnisse der Untersuchung der Nachbarschaftsgenauigkeit in den glatten Bereichen des städtischen Testgebiets. In Abb. 9a und b sind die Standardabweichung der Gitterpunkthöhen, bestimmt mittels gleitender Schrägebeneninterpolation aus den jeweils 12 nächsten Nachbarpunkten, farbkodiert von blau (0-5 mm) über grün (5-15 mm) bis rot (50 mm) dargestellt. Der SPL Datensatz weist hierbei in den Straßenbereichen eine bessere innere Genauigkeit auf als der Waveform LiDAR Datensatz. In beiden Fällen liegen die Nachbarschaftsgenauigkeiten unter den von den Herstellern angegebenen Werten. Die Standardabweichung der Gitterpunktshöhen ist im Waveform LiDAR über die gesamte Szene relativ konstant. Demgegenüber fällt in den SPL Daten ein größerer Unterschied zwischen Straßenflächen und Dächern auf. Eine

genauere Analyse der Hintergründe dieses Phänomens ist Gegenstand von künftigen Untersuchungen. Jedenfalls kann für den vorliegenden Datensatz keine globale Aussage hinsichtlich der Überlegenheit des einen oder andern Systems mit Hinblick auf die Relativgenauigkeit gemacht werden. Vielmehr weisen die SPL Punkte gegenüber den Waveform LiDAR Daten an manchen Stellen eine deutlich höhere Streuung auf (z.B. Dachfläche in Abb. 9e), während dies an anderen Stellen nicht der Fall ist (z.B. Abb. 9f). Die quantitative Auswertung ergibt für beide Datensätze im Mittel ähnliche Werte (Mittelwert SPL: 12 mm vs. Waveform LiDAR: 9 mm).



Abb. 9: Nachbarschaftsgenauigkeit; (a+b) Single Photon LiDAR, (c+d) Waveform LiDAR, (e+f) Profilschnitte an unterschiedlich geneigten Dachflächen, SPL (blau) und Waveform Li-DAR (braun)

6 Zusammenfassung und Ausblick

In diesem Beitrag wurden die Funktionsprinzipien von konventionellem Full Waveform Laserscanning und die Einzelphoton-sensitiven Technologien Geiger-mode LiDAR und Single Photon LiDAR beschrieben. Speziell lag der Fokus dabei auf der SPL Technologie, da mit dem Hexagon/Leica SPL100 nunmehr ein Sensor auch in Europa verfügbar ist, der gegenüber herkömmlichen Laserscannern eine höhere Flächenleistung für die Topographieerfassung bei gleicher Punktdichte verspricht.

Mittlerweile liegen zwei Befliegungen eines Teilgebietes der Stadt Wien mit dem SPL100 (Juli 2018) und dem RIEGL VQ-1560i (September 2018) vor. Erste Ergebnisse der Datenevaluierung wurden in dem Beitrag vorgestellt. Dabei kann festgehalten werden, dass (i) mit der SPL Technologie aufgrund der höheren Flughöhe und der daraus resultierenden größeren Streifenbreite weniger Flugstreifen zur Erfassung der selben Fläche bei vorgegebener Punktdichte erforderlich sind,

(ii) die Waveform LiDAR 3D-Punktwolken die Topographie sowie Gebäude prägnanter wiedergeben, (iii) Vegetationsdurchdringung bei der SPL Technologie gegeben ist, diese im Untersuchungsgebiet im Vergleich zu Waveform LiDAR aber moderat war, (iv) das kreisförmige Scanmuster des SPL100 günstig für die Erfassung von Fassaden im Stadtbereich ist, die Straßenfluchten aber aus demselben Grund oftmals abgeschattet sind, und (v) die relative Nachbarschaftsgenauigkeit bei beiden Systemen besser als 5 cm ist, wobei eine höhere Schwankungsbreite beim SPL Datensatz vor allem bei geneigten und schwach reflektierenden Flächen festgestellt werden konnte.

Der vorliegende Beitrag hat Vor- und Nachteile der jeweiligen Technologie aufgezeigt, wobei die Datenanalyse auch Fragen aufgeworfen hat, deren Beantwortung Gegenstand künftiger Untersuchungen sein wird. Inwieweit Single Photon LiDAR als Ersatz für konventionelles Laserscanning, etwa für einen großflächigen Einsatz in der Landesvermessung ist, hängt in erster Linie von den Nutzeransprüchen ab und kann daher aus wissenschaftlicher Sicht nicht global beantwortet werden.

7 Danksagung

Die Arbeiten von Gottfried Mandlburger zu diesem Beitrag wurden durch Mitteln der Deutschen Forschungsgesellschaft (DFG) im Rahmen des Forschungsprojektes "Bathymetrievermessung durch Fusion von Flugzeuglaserscanning und multispektralen Luftbildern" unterstützt.

8 Referenzen

- AGISHEV, R., COMERÓN, A., BACH, J., RODRIGUEZ, A., SICARD, M., RIU, J. & ROYO, S., 2013: Lidar with SiPM: Some capabilities and limitations in real environment. Optics and Laser Technology, **49**, 86-90.
- BRIESE, C., PFENNIGBAUER, M., LEHNER, H., ULLRICH, A., WAGNER, W. & PFEIFER, N., 2012: Radiometric calibration of multi-wavelength airborne laser scanning data. In ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 1, 335-340.
- DEGNAN, J., 2002: Photon-counting multikilohertz microlaser altimeters for airborne and spaceborne topographic measurements. Journal of Geodynamics, **34**(3-4), 503-549.
- DEGNAN, J., 2016: Scanning, Multibeam, Single Photon Lidars for Rapid, Large Scale, High Resolution, Topographic and Bathymetric Mapping. Remote Sensing, 8(11), 923-958.
- KIM, S., LEE, I. & KWON, Y. J., 2013: Simulation of a geiger-mode imaging LADAR system for performance assessment. Sensors (Switzerland), **13**(7), 8461-8489.
- Lehner, H. & Briese, C., 2010: Radiometric calibration of Full-Waveform Airborne Laser Scanning Data based on natural surfaces. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 38(7B), 360-365.
- MANDLBURGER, G. & JUTZI, B., 2018: Feasibility Investigation on Single Photon Lidar Based Water Surface Mapping. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences, 4, 109-116.
- PFEIFER, N., MANDLBURGER, G. & GLIRA, P., 2015: Laserscanning. In: Photogrammetrie und Fernerkundung, 1st ed., 1-51, Berlin Heidelberg: Springer.

- PFENNIGBAUER, M. & ULLRICH, A., 2018: Noisy lidar point clouds: impact on information extraction in high-precision lidar surveying. Laser Radar Technology and Applications, 33.
- PFENNIGBAUER, M., WOLF, C., WEINKOPF, J. & ULLRICH, A., 2014: Online waveform processing for demanding target situations. In Proc. SPIE, **90800J**.
- STOKER, J. M., ABDULLAH, Q. A., NAYEGANDHI, A. & WINEHOUSE, J., 2016: Evaluation of Single Photon and Geiger Mode Lidar for the 3D Elevation Program. Remote Sensing, 8(9), 716-767.
- SUGARBAKER, L. J., CONSTANCE, E. W., HEIDEMANN, H. K., JASON, A. L., LUKAS, V., SAGHY, D. L. & STOKER, J. M., 2014: The 3D Elevation Program Initiative — A Call for Action. US Geological Survey, https://pubs.usgs.gov/circ/1399/pdf/circ1399.pdf.
- ULLRICH, A. & PFENNIGBAUER, M., 2016: Linear LIDAR versus Geiger-mode LIDAR: impact on data properties and data quality. In Proc. SPIE, **9832**, 983204-983217.
- WAGNER, W., 2010: Radiometric calibration of small-footprint full-waveform airborne laser scanner measurements: Basic physical concepts. ISPRS Journal of Photogrammetry and Remote Sensing, **65**(6), 505-513.
- WAGNER, W., ULLRICH, A., MELZER, T., BRIESE, C. & KRAUS, K., 2004: From single-pulse to fullwaveform airborne laser scanners: potential and practical challenges. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 35(B3), 201-206.
Portables bildbasiertes Mobile Mapping System im Einsatz unter Tage – Systemaufbau, Georeferenzierungs- und Genauigkeitsuntersuchungen

STEFAN BLASER¹, STEPHAN NEBIKER¹ & DOMINIK WISLER¹

Zusammenfassung: Mit der fortschreitenden Urbanisierung gewinnen die Erstellung und der Werterhalt unterirdischer Bauten und Infrastrukturen zunehmend an Bedeutung. Mit bildbasierten Webdiensten mit 3D-Messfunktionalität – im Ansatz vergleichbar mit Street View – lassen sich aufwändige oft gefährliche Feldbegehungen und Messarbeiten ins Büro verlagern. Für eine effiziente 3D-Erfassung von Innenräumen wurde das portable bildbasierte Mobile Mapping System BIMAGE Backpack mit einer Mehrkopf-Panoramakamera, zwei Multiprofillaserscannern sowie einer Inertialmesseinheit entwickelt. Es erfolgten Aufnahmen und Genauigkeitsuntersuchungen im Versuchsstollen Hagerbach. Ähnlich wie im Außenraum, betrug nach anschließender bildbasierter Georeferenzierung die absolute Genauigkeit zwischen 5.0 und 7.2 cm. Die für die Praxis sehr wichtige relative Messgenauigkeit betrug nach der bildbasierten Georeferenzierung sogar 3-10 mm, was einer Genauigkeitssteigerung von etwa einer Zehnerpotenz gegenüber der Georeferenzierung mittels LiDAR-SLAM entspricht. Somit ist ein großes Einsatzpotenzial für die Instandhaltung und Inventarisierung von unterirdischen Anlagen gegeben.

1 Einleitung

Mit der fortschreitenden Urbanisierung und der baulichen Verdichtung gewinnt das Bauen im Untergrund zunehmend an Bedeutung. Dies gilt in besonderem Masse für die Erstellung und den Werterhalt von unterirdischen Infrastrukturbauten, wie etwa von U-Bahnnetzen. Die Digitalisierung führt zu tiefgreifenden Veränderungen im gesamten Planungs- und Bauprozess. Sie bietet insbesondere große Chancen im Management von Infrastrukturanlagen, indem zahlreiche aufwändige und potentiell gefährliche Mess-, Beurteilungs- und Planungsaufgaben vom Feld ins Büro verlagert werden können. Dies gilt in besonderem Masse für unterirdische Bauten und Verkehrsinfrastrukturen. Voraussetzung dafür bilden realitätsgetreue, genaue, zuverlässige und aktuelle 3D-Daten, die sich in kollaborativen digitalen Planungs- und Managementprozessen nutzen lassen. Im Außenraum werden dazu bildbasierte Webdienste mit 3D-Messfunktionalität - im Ansatz vergleichbar mit Street View - bereits verbreitet eingesetzt (NEBIKER 2017). Die Kombination von mobilem Reality Capturing in der Form von Multiview-Stereobilddaten und daraus abgeleiteten georeferenzierten 3D-Bilddatenbanken mit webbasierten Cloud Services bieten eine sehr intuitive Dateninterpretation und eine hohe Benutzerfreundlichkeit mit einem breiten Spektrum an Funktionalität. Damit entwickeln sich 3D-Bilddienste zur visuellen universellen Geoinformationsplattform für Stadträume und Infrastrukturen im Außenraum. Um zukünftig auch großräumige 3D-Bilddienste für Innenräume - wie etwa Gebäudekomplexe oder Infrastrukturen im Untergrund - realisieren zu können, wurden im Forschungsprojekt BIMAGE eine Reihe von Hardware- und Softwarekomponenten konzi-

¹ Fachhochschule Nordwestschweiz, Institut Geomatik, Hofackerstrasse 30,

CH-4132 Muttenz, Schweiz, E-Mail: [stefan.blaser, stephan.nebiker]@fhnw.ch, dwisler@gmx.ch

piert und untersucht. Insbesondere wurde mit dem *BIMAGE Backpack* ein vollständig funktionsfähiger Forschungsprototyp eines portablen bildbasierten Mobile Mapping Systems zur effizienten und genauen kinematischen 3D-Datenerfassung entwickelt. Die aktuelle Systemkonfiguration besteht aus einer Mehrkopf-Panoramakamera *Ladybug5*, zwei Multiprofillaserscannern *Velodyne VLP-16*, sowie einer Inertialmesseinheit *XSens MTI-300*. Während der Aufnahme berechnet ein LiDAR-SLAM-Algorithmus jeweils die Pose des Systems in einem lokalen Koordinatenbezugsrahmen. Dadurch kann die Mehrkopf-Panoramakamera nach geometrischen Bedingungen ausgelöst werden. Nach der Aufnahme liegen die Einzelbilder der Panoramakamera mit den zugehörigen Parametern der äußeren Orientierung basierend auf der SLAM-Lösung vor. Ebenso kann die Punktwolke aus dem LiDAR-SLAM für weitere Untersuchungen exportiert werden. Die Einzelbilder lassen sich mittels bildbasierter Georeferenzierung verbessern, indem die äußeren Orientierungen des SLAM als Näherungswerte in eine *Structure-from-Motion-Software* eingeführt werden und damit nachträglich eine Bündelblockausgleichung durchgeführt wird.

In diesem Beitrag wird das Potenzial eines bildbasierten portablen MMS unter Tage untersucht und aufgezeigt. Als Untersuchungsgebiet diente der Versuchsstollen Hagerbach (VSH) der Amberg Gruppe. Der VSH ist ein verzweigtes Stollensystem mit unterschiedlichen Tunnelausbaustufen. Darin sind zahlreiche im Bergbau tätige Firmen eingemietet, welche verschiedenartige Feldversuche, von der Sprengung über die Prüfung der Materialfestigkeit bis hin zu vollautomatisierten Tunnelausbauverfahren durchführen.

2 Bildbasiertes Mobile Mapping System

Mit Hilfe des neuen bildbasierten Mobile Mapping Systems (MMS) sollen Innenräume kinematisch erfasst und mittels georeferenzierter 3D-Bilder repräsentiert werden können, analog zu existierenden stereobildbasierten MMS im Außenraum (BURKHARD et al. 2012; BLASER et al. 2017). In NEBIKER et al. (2015) werden das Konzept und der Aufbau georeferenzierter 3D-Bildräume beschrieben sowie deren Vorteile u.a. in der Handhabung gegenüber 3D-Punktwolken diskutiert. Die für nahezu jedes Pixel verfügbare Tiefeninformation eines 3D-Bilds ermöglicht denkbar einfache, direkte 3D-Messungen im Bild per Mausklick. Dank bekannter äußerer Orientierung der Stereobilder, können aus den 3D-Messungen absolute 3D-Koordinaten berechnet werden. Der Ansatz des 3D-Bilds beschränkt sich jedoch nicht nur auf die zusätzliche Ebene mit Tiefeninformation, ein 3D-Bild lässt sich um weitere pixelbasierte Informationsebenen (z.B. Qualität der Tiefeninformation, Klassifikation, etc.) beliebig erweitern. Ein prototypisches, für Forschungszwecke ausgelegtes MMS für den Innenraum sollte einerseits portabel sein, andererseits eine ausreichende Flexibilität aufweisen, damit unterschiedliche Sensorkonfigurationen aufgebaut und untersucht werden können. Des Weiteren sollte die Umgebungssensorik den Raum möglichst vollständig abdecken. Die Navigationssensorik sollte eine präzise direkte Georeferenzierung liefern können. Im Außenraum kommt die Kombination von GNSS- und Inertialnavigationssensorik (INS) weit verbreitet zum Einsatz, was im Innenraum verunmöglicht wird, da keine GNSS-Signale empfangen werden können. Eine mögliche Alternative zur GNSS- und INS-Kombination für den Innenraum stellt 3D-LiDAR-SLAM dar (HESS et al. 2016). Mittels bildbasierter Georeferenzierung könnten die lokalen Kameraposen des LiDAR-SLAMs nochmals verbessert und in einen absoluten Koordinatenbezugsrahmen überführt werden (CAVEGN et al. 2016).

2.1 Systemkonfiguration

Alle Sensoren wurden an einen robusten Aluminiumrahmen montiert, welcher an einem Rucksacktraggestell befestigt ist. Die Umgebungssensorik der aktuellen Systemkonfiguration besteht aus der Mehrkopf-Panoramakamera Ladybug5 von FLIR. Sie besteht aus fünf radial angeordneten Kameraköpfen mit horizontaler Blickrichtung und einer nach oben gerichteten Kamera, welche eine Auflösung von jeweils 5 MP aufweisen. Die Panoramakamera wurde leicht schräg nach hinten geneigt an den Rahmen fixiert, was durch den Vorhalt des Körpers beim Tragen der Plattform entsprechend kompensiert wird (Abb. 1, rechts, Nr. 2). Dank der Bildüberlappung der einzelnen Kameraköpfe ist die Verdeckung durch den Aluminiumrahmen minimal. Als Navigationssensorik dienen zwei Mehrprofillaserscanner des Typs VLP-16 von Velodyne. Einer der beiden Laserscanner wurde oben am Rahmen nahezu horizontal (Abb. 1, rechts, Nr. 1), aber leicht schräg nach hinten geneigt befestigt, der zweite Laserscanner befindet sich weiter unten und wurde vertikal angeordnet (Abb. 1, rechts, Nr. 5). Diese Konfiguration verleiht der fusionierten Punktwolke die notwendige geometrische Stabilität – auch in beengten Verhältnissen wie etwa Treppenhäusern. Weiter zählt die INS MTI-300 von XSens zur Navigationssensorik (Abb. 1, rechts, Nr. 7). Die beiden Punktwolken und die INS-Rohdaten werden in einem 3D-LiDAR-SLAM fusioniert und liefern dadurch die relative Position, was in Kap. 2.2 im Detail erläutert wird. Die Speisung aller Systemkomponenten erfolgt mit einem 12 V Lithium-Ionen-Akkumulator, welcher eine Kapazität von 20 Ah aufweist. Das aktuelle Systemgewicht beträgt rund 20 kg. In BLASER et al. (2018) findet sich eine detaillierte Darlegung aller Systemkomponenten mit den zugehörigen Spezifikationen.



Abb. 1: Systemaufbau des tragbaren bildbasierten Mobile Mapping Systems *BIMAGE Backpack* mit Gehäuse (links) und mit der Bezeichnung aller Komponenten (rechts).

2.2 Sensorsynchronisation und Erfassungssoftware

Wegen der Mobilität der Messplattform, bildet die Sensorsynchronisation ein wichtiger Bestandteil, welche direkten Einfluss auf die Messgenauigkeit hat. Die Zeitreferenz des MMS befindet sich auf dem Computer II *Arduino Nano* (Abb. 2). Mit einem elektrischen Puls pro Sekunde synchronisiert er sämtliche Navigationssensoren (LiDAR & INS), die kontinuierlich Daten erfassen (Abb. 2). Diese Daten fließen zu Computer I und werden dort auf die Festplatte gespeichert. Die Kamerasteuerung erfolgt durch die Erfassungssoftware mit einem entsprechenden Befehl an Computer II. Dieser löst mittels elektrischem Signal die Kamera, sowie die Blitzlichter aus und sendet die genaue Bilderfassungszeit an Computer I zurück (Abb. 2).



Abb. 2: Schematische Darstellung des Datenflusses und der Sensorsynchronisation.

Analog zur Hardwarekonfiguration, wurde auch die Erfassungssoftware modular und dementsprechend flexibel aufgebaut (Abb. 3, links). Die Software fußt auf dem in der Robotik weit verbreiteten open-source Framework Robot Operating System (ROS). Neben zahlreichen Werkzeugen, Algorithmen und Treibern, weist ROS ein flexibles, Graph-basiertes Kommunikationskonzept auf. Einzelne, modular abgekapselte Anwendungen werden als ROS-Nodes bezeichnet (Darstellung als Ellipse). Das asynchrone Kommunikationskonzept wird unterstützt, indem die ROS-Nodes mit einem ROS-Publisher (Darstellung als roter Pfeil) Datenmitteilungen in einem vordefinierten Format an ein ROS-Topic (Darstellung als Rechteck) senden können. Im Gegenzug kann ein ROS-Node mit einem ROS-Subscriber (Darstellung als blauer Pfeil) ein ROS-Topic abonnieren und - im Falle von verfügbaren Datenmitteilungen - die Daten empfangen und entsprechend weiterverarbeiten. Auch das synchrone Kommunikationsparadigma wird unterstützt, indem ein ROS-Node einen ROS-Service (Darstellung als Rechteck) anbieten kann. Ein anderer ROS-Node kann diesen Service wiederum mit einer Anfrage aufrufen und anschließend eine Antwort empfangen (Darstellung mit grünen Pfeilen). QUIGLEY et al. (2009) bieten eine vertiefte Einführung in die Konzepte und Eigenschaften von ROS.

Die Rohdaten der Navigationssensorik (Laserscanner und INS) werden sowohl als ROS-*Bag*-Datei für die spätere Nachprozessierung gespeichert, als auch vom 3D-LiDAR SLAM *Google Cartographer* (HESS et al. 2016) verwendet, um die relative 3D-Systempose zu errechnen (Abb. 3, links). Aufgrund einer definierbaren Distanz- und oder Azimutdifferenz, lässt sich die Panoramakamera über geometrische Bedingungen auslösen. Das empfangene Bild der Panoramakamera besteht aus aneinandergereihten Einzelbildern der Kameraköpfe. Entsprechende ROS-*Nodes* zerschneiden das Bild in seine Einzelbilder, die anschließend auf die Festplatte gespeichert werden.

Das entwickelte *BIMAGE Backpack* kann direkt in Innenräumen initialisiert werden. Der Startpunkt definiert jeweils den Ursprung des lokalen Koordinatenbezugsrahmens. Der 3D-LiDAR SLAM ermittelt einerseits die 3D-Pose des Systems in Echtzeit. Dazu werden die Punktwolken über eine Voxel-basierte Karte, welche zeitgleich im Hintergrund erstellt wird, fusioniert. Da die Voxelkarte jeweils als 2D-Visualisierung zur Verfügung steht, kann der Aufnahmefortschritt in Echtzeit mitverfolgt werden (Abb. 3, oben rechts). Für die Kameraauslösung wurden sowohl zeitliche als auch geometrische Bedingungen implementiert. Neben der Auslösung enthält die Software einige vordefinierte Kamerakonfigurationssätze für die Aufnahme unterschiedlich hell ausgeleuchteter Umgebungen (Abb. 3, unten rechts).



Abb. 3: Schema der implementierten Erfassungssoftware auf ROS-Basis (links). Die ROS-Nodes sind als Ellipsen dargestellt, währenddessen ROS-Topics und ROS-Services rechteckig dargestellt sind. Die roten Pfeile repräsentieren ROS-Publishers, die blauen Pfeile ROS-Subscribers und die grünen Pfeile ROS-Services. Screenshot der Cartographer SLAM Voransicht (oben rechts), Screenshot mit der Bildvoransicht, den Kameraeinstellungen und mit den Auslösebedingungen (unten rechts).

2.3 Auswerteprozess

Die inneren Orientierungsparameter der einzelnen Panoramakameraköpfe, sowie die relativen Orientierungen zwischen den Panoramakameraköpfen werden vorgängig in einem mit photogrammetrischen Targets bestückten Kalibrierfeld geschätzt. Für die Panoramakameraköpfe wird das äquidistante Kameramodell verwendet (ABRAHAM & FÖRSTNER 2005). In BLASER et al. (2018) wird das eingesetzte Kalibrierverfahren im Detail abgehandelt. Die Hebelarme und Fehlausrichtungen der Panoramakamera, respektive Laserscanner bezüglich Navigationszentrum werden bisher aus dem mechanischen Konstruktionsplan übernommen. Deren analytische Kalibrierung ist zurzeit noch ausstehend.

Nach erfolgter Aufnahme wird der *Cartographer Status* des LiDAR-SLAM exportiert, welcher auch die optimierte SLAM-Trajektorie enthält. In einem ersten Schritt wird die SLAM-Trajektorie extrahiert. Um die Kameraposen zu erhalten, werden die Bildevents über den Erfassungszeitpunkt in die SLAM-Trajektorie interpoliert. Zusätzlich werden die bereits bekannten Hebelarme und Fehlausrichtungen der einzelnen Kameras bezüglich Navigationszentrum angebracht (Abb. 4). In einem davon unabhängigen Prozess werden die Bilder mit den vorkalibrierten inneren Orientierungsparametern geometrisch entzerrt. Die verzeichnungsfreien Bilder können nun den SLAM-basierten Bildposen zugeordnet werden und lassen sich bereits in einem cloudbasierten Bilddienst betrachten. Des Weiteren können damit bereits 3D-Messungen mittels räumlichem Vorwärtseinschnitt mit Bildern unterschiedlicher Auslösezeitpunkte vorgenommen werden.

In einem weiteren Schritt können jedoch die SLAM-basierten Bildposen mit einer Structurefrom-motion-Pipeline (SfM) weiter verbessert werden. Dabei dienen die SLAM-basierten Bildposen als Näherungswerte. CAVEGN et al. (2018) erweiterten die von SCHÖNBERGER & FRAHM (2016) entwickelte SfM-Software COLMAP um die Möglichkeit, um einerseits äußere Orientierungen und Passpunkte einzuführen und andererseits Kamerakonfigurationen mit relativen Orientierungsparametern zu fixieren. Ab Version 1.5 von *Agisoft PhotoScan* ist dieselbe Funktionalität ebenfalls implementiert. Mit der bildbasierten Georeferenzierung lassen sich die Bildposen nicht nur verbessern, sondern ebenso in einen absoluten Bezugsrahmen überführen. Die Qualität, sowie die Robustheit der bildbasierten Georeferenzierung sind maßgeblich von den Lichtverhältnissen, der Beschaffenheit der kartierten Umgebung (Strukturierung, Reflexionen, Schatten, etc.) und der Qualität der Näherungswerte abhängig.



Abb. 4: Flussdiagramm mit dem SLAM- und bildbasierten Post-Mission-Datenauswertungsprozess.

3 Untersuchungsgebiet

Der Versuchsstollen Hagerbach (VSH) ist eine Forschungs- und Entwicklungseinrichtung, welche sich für Praxistests diverser Innovationen im Untertagebau eignet. Der VSH wurde im Jahre 1970 gegründet und befindet sich auf dem Gemeindegebiet von Flums Hochwiese (CH). Die Amberg Gruppe bietet, als Eigentümerin des VSH, verschiedenen eingemieteten Partnerfirmen ideale Rahmenbedingungen, um ein breites Spektrum an Untersuchungen im Untertagebau durchzuführen. Zusätzlich unterhält sie ein Labor für die Baustoffprüfung (VERSUCHS-STOLLEN HAGERBACH AG 2018). Über die Jahre wuchs der VSH zu einem weit verzweigten Stollensystem an. Die einzelnen Stollen sind unterschiedlich ausgeprägt und weisen beispielsweise verschiedene Ausbaustufen und Stollendurchmesser auf (Abb. 5).

3.1 Referenzdaten

In RECHSTEINER & WISLER (2018) ist die tachymetrische Neubestimmung und Auswertung des gesamten Fixpunktnetzes des VSH dokumentiert. Das Fixpunktnetz des VSH besteht aus 127 mit Schraubbolzen versicherten Fixpunkten, welche mit mehreren zusammenhängenden Polygonzügen, eingemessen wurden. Die Fixpunkmessungen wurden in einem Guss nach der Methode der kleinsten Quadrate ausgeglichen und minimal gelagert, um zusätzliche Fehlereinflüsse des bestehenden, nicht streng ausgeglichenen Fixpunktnetzes zu vermeiden. Die maximale empirische Standardabweichung zwischen zwei Fixpunkten innerhalb des Netzes beträgt 4.1 mm in der Lage und 0.9 mm in der Höhe.

3.2 Versuchsstollen

Die beiden Untersuchungsgebiete «Sicherheitsstollen» (Abb. 5, (1)) und der Stollen des *Swiss Center of Applied Underground Technologies* (SCAUT) (Abb. 5, (2)) sind unterschiedlich ausgeprägt. Während dem der Sicherheitsstollen verschiedene Ausbaustufen eines Straßentunnels aufweist (Abb. 6, links), befindet sich SCAUT-Stollen ein innovativer Versuchsaufbau für den effizienten Einbau von Notausgängen in Eisenbahntunneln (Abb. 6, rechts).



Abb. 5: Grundriss des Versuchsstollens Hagerbach (VSH). Die dunkelgrauen Bereiche sind ohne Hindernisse zugänglich im Gegensatz zu den hellgrauen Bereichen, welche abgeschlossen oder unbegehbar sind. Die beiden Untersuchungsgebiete "Sicherheitsstollen" (1) und Stollen des SCAUT (2) sind strichpunktiert umrahmt.



Abb. 6: Untersuchungsgebiet "Sicherheitsstollen" (1), ausgelegt als Straßentunnel mit unterschiedlichen Ausbaustufen (links), Untersuchungsgebiet des SCAUT (2), ausgelegt als Eisenbahntunnel mit Oberleitungen und Querschlag zum Rettungsstollen (rechts).

4 Genauigkeitsuntersuchungen

Mit dem *BIMAGE Backpack* erfolgten Aufnahmen in beiden Untersuchungsgebieten des VSH. Aufgrund der schlechten Lichtverhältnisse im Stollen, wurden am MMS zusätzliche seitwärts gerichtete akkubetriebene Scheinwerfer angebracht. Die Panoramakamera wurde mit einer geometrischen Distanzbedingung von 2 m ausgelöst. Das erste Untersuchungsgebiet "Sicherheitsstollen" wurde schleifenförmig erfasst (Abb. 7, links), um den Betrag eines allfälligen Schleifenschlussfehlers zu beurteilen. Im Gegensatz dazu, wurde bei der Aufnahme des zweiten Untersuchungsgebiets "SCAUT-Stollen" zusätzlich ein weiterer Stollenabschnitt nach einer scharfen Kurve miterfasst.

Der 3D-SLAM-Algorithmus liefert neben der Trajektorie auch eine Punktwolke mit den registrierten Scans beider Multiprofillaserscanner. Diese Punktwolke eignet sich für eine erste qualitative Beurteilung der SLAM-basierten Georeferenzierung. Um die Georeferenzierung quantitativ zu beurteilen, können Differenzvektoren von Passpunktmessungen in den Bildern zu den tachymetrisch bestimmten Passpunktkoordinaten herbeigezogen werden. In einem mit *Python* entwickelten Passpunktmesswerkzeug können in mehreren aufeinanderfolgenden Bildern Messungen vorgenommen werden. Mit den zugehörigen Kameraposen lassen sich mittels Vorwärtseinschnitt die 3D-Koordinaten berechnen. Die 3D-Koordinaten befinden sich im lokalen Koordinatenbezugsrahmen der Kameraposen. Mit einem weiteren *Python*-Werkzeug können die lokalen Passpunktkoordinaten mit einer 3D-Koordinatentransformation über ausgewählte Stützpunkte in den globalen Bezugsrahmen überführt werden.

4.1 SLAM-basierte Georeferenzierung

Die schleifenförmige Aufnahme des Untersuchungsgebiets "Sicherheitsstollen" (Abb. 7, links) wurde bei der ersten Gabelung unten rechts initialisiert (S) und nach einem Umlauf bei der zweiten Gabelung mit einem Überlappungsbereich am Ende abgeschlossen (E). Die Initialisierung der linienförmigen Aufnahme des "SCAUT-Stollens" (Abb. 7, rechts) erfolgte oben rechts (S) und wurde, nach der Umkehr am Ende des SCAUT-Stollens wiederum in der Nähe des Startpunkts abgeschlossen (E).

Mit der Punktwolkenprojektion auf die XY-Ebene kann die Georeferenzierungsqualität in einem ersten Schritt visuell beurteilt werden (Abb. 7). Bei der Aufnahme des Sicherheitsstollens, welche mit den standardmäßigen LiDAR-SLAM-Parametereinstellungen erfasst wurde, ist ein großer Schleifenschlussfehler erkennbar (Abb. 7, 1a). Durch kontinuierliche Optimierung der SLAM-Parameter, insbesondere durch ein höheres Rotationsgewicht des *Ceres Scan-Matchers* und eine erhöhte Anzahl durchzuführender Graph-Optimierungen, konnte der Schleifenschluss vom Algorithmus detektiert und nachträglich korrigiert werden (Abb. 7, 1b). Die Parameteroptimierung der SLAM-Prozessierung ist in RECHSTEINER & WISLER (2018) im Detail abgehandelt. Die zweite linienförmige Aufnahme des SCAUT-Stollens wurde sowohl mit den Standard-SLAM-Parametern, als auch mit dem optimierten Parametersatz für den Sicherheitsstollen prozessiert. Die Punktwolkenprojektion des optimierten Parametersatzes (Abb. 7, 2b) weist ein geringeres Rauschen auf, als jene des Original-Parametersatzes (Abb. 7, 2a).



 Abb. 7: Projektionen von Punktwolken des 3D-LiDAR-SLAM-Algorithmus auf die XY-Ebene. Links: Punktwolken des "Sicherheitsstollens" (1); rechts: Punktwolken des "SCAUT-Stollens" (2). Die Punktwolken oben (a) wurden mit dem originalen SLAM-Parametersatz prozessiert, bei den unteren Punktwolken (b) kam der optimierte Parametersatz zur Anwendung.

4.2 Bildbasierte Georeferenzierung

Für die anschließende bildbasierte Georeferenzierung wurden die Bilder mit den SLAMbasierten Kameraposen als Näherungswerte in die SfM-Software *Agisoft PhotoScan* eingeführt. Die vorkalibrierten inneren Orientierungen der Kameraköpfe, sowie die Hebelarme und Fehlausrichtungen zwischen den einzelnen Kameraköpfen, wurden festgesetzt. Es wurden keine zusätzlichen Passpunktmessungen in die Bündelblockausgleichung eingeführt. Somit wurde das Bündel auf den genäherten Kameraposen gelagert. Die Prozessierungszeiten und die erzielten Rückprojektionsfehler sind vergleichbar mit denjenigen von CAVEGN ET AL. (2018), die das Konzept der bildbasierten Georeferenzierung von MMS-Daten im Detail behandeln.

Die resultierenden verbesserten Bildposen wurden als Basis für die 3D-Pass- und Kontrollpunktmessungen mittels Vorwärtseinschnitt verwendet. Anschließend wurden diese Punktmessungen jeweils über vier Stützpunkte vom lokalen in den globalen Bezugsrahmen transformiert (Abb. 7).

Die mittleren Kontrollpunktdifferenzen der bildbasierten Georeferenzierung zu den tachymetrisch bestimmten Referenzdaten betragen innerhalb des Sicherheitsstollens 60 mm und innerhalb des SCAUT-Stollens 72 mm. Außerhalb des SCAUT-Stollens treten stetig zunehmende systematische Abweichungen zu den Kontrollpunktkoordinaten von bis zu 540 mm auf (Abb.8, 2). Dies könnte auf eine fehlerhafte Bildverknüpfung in der starken Kurve hindeuten, oder aber durch die Lagerung im SCAUT-Stollen und die beträchtliche Extrapolation verursacht worden sein. Eine ähnlich große Abweichung tritt im Punkt außerhalb des Sicherheitsstollens auf (Abb. 8, 1). Für diese Abweichung könnten dieselben Ursachen wie im SCAUT-Stollen in Betracht gezogen werden. Zu bemerken ist, dass die Lichtverhältnisse abseits des Sicherheitsstollens deutlich schlechter waren, was die bildbasierte Georeferenzierung erschwerte.



Abb. 8: Übersichtsplan der bildbasierten Georeferenzierung. Die grünen Punkte repräsentieren die Bildaufnahmestandorte. Für die Genauigkeitsuntersuchungen wurden sämtliche lokalen Punktmessungen über die roten Stützpunkte mit einer 3D-Helmerttransformation in den absoluten Bezugsrahmen transformiert. Die roten und die blauen Verschiebungsvektoren repräsentieren die Lageabweichung in den Stütz- bzw. Kontrollpunkten, die schwarzen Vektoren die Höhenabweichung (Hinweis: alle Vektoren sind um Faktor 50 überzeichnet).

4.3 Resultate und Diskussion

Auf Basis a) der standardmäßigen Georeferenzierung mittels SLAM, b) der SLAM-Georeferenzierung mit optimierten Parametern, sowie c) der bildbasierten Georeferenzierung, wurden signalisierte Fixpunkte in je vier aufeinanderfolgenden Bildern gemessen und mittels Vorwärtseinschnitt deren 3D-Koordinaten berechnet.

4.3.1 Relative Genauigkeiten

Als erstes wurden die empirischen Standardabweichungen der Vorwärtseinschnitte ausgewiesen. Sie repräsentieren die Präzision einer 3D-Koordinatenbestimmung. Die Präzision wird hauptsächlich durch die relative Orientierungsgenauigkeit aufeinanderfolgender bzw. benachbarter Bilder beeinflusst. In CAVEGN ET AL. (2018) wurde ebenfalls die Güte der kalibrierten relativen Orientierungen zwischen den einzelnen Panoramakameraköpfen beurteilt, indem dieselben Punkte jeweils in unterschiedlichen Kameraköpfen gemessen wurden. Dabei wurden keine signifikanten Abweichungen in den Passpunkten festgestellt. In diesem Beitrag wurden für die Berechnung der Vorwärtseinschnitte teilweise auch unterschiedliche Kameraköpfe verwendet. Aufgrund der Präzision lässt sich auf die erreichbare Genauigkeit einer Relativmessung, beispielsweise einer 3D-Distanz, schließen. In beiden Untersuchungsgebieten verbesserte sich die Präzision der Standard-SLAM-Lösung a) von 79-150 mm durch die Verwendung optimierter SLAM-Parameter b) auf 66-101 mm und durch bildbasierte Georeferenzierung c) nochmals signifikant auf 3-10 mm (Tab. 1 und Tab. 2). Im Untersuchungsgebiet "SCAUT-Stollen" wurde gegenüber dem Gebiet "Sicherheitsstollen" eine um Faktor zwei verschlechterte Präzision bei Verwendung von SLAM-basierter Georeferenzierung nachgewiesen. Grund dafür könnte der größere Stollendurchmesser sein, wodurch sich die Punktwolkendichte am Objekt, zuungunsten der Merkmalsdetektion, verringert. Umgekehrt verhält es sich bei der bildbasierten Georeferenzierung, bei welcher eine offene Umgebung mehr gemeinsame Merkmale mit besseren Schnittbedingungen zulässt.

4.3.2 Absolute Genauigkeiten

Um die absolute Genauigkeit zu beurteilen, wurden zuerst die gemessenen Passpunkte sämtlicher Datensätze über jeweils vier Stützpunkte (SP) (Abb. 8) mit einer 3D-Koordinatentransformation ins Referenzkoordinatensystem überführt. Anschließend wurden jeweils die Differenzvektoren zu den tachymetrisch bestimmten Referenzkoordinaten gebildet. Pro Datensatz wurden jeweils die mittleren Abweichungen für die Stützpunkte (SP), die Kontrollpunkte innerhalb des Untersuchungsgebiets (ohne Extrapolation) (KP Kerngebiet) und für sämtliche Kontrollpunkte (alle KP) separat ausgewiesen. Mit optimierten SLAM-Parametern konnte die die mittlere Abweichung aller Kontrollpunkte im "Versuchsstollen" von 1.47 m auf 0.26 m um nahezu Faktor sechs (Tab. 1) und im "SCAUT-Stollen" von 0.63 m auf 0.21 m um Faktor drei verringert werden. Die große Verbesserung rührt daher, dass der SLAM-Algorithmus, durch automatisch detektierte Schleifenschlüsse, systematisch auftretende Abweichungen weitgehend eliminieren konnte (siehe Kap. 4.1).

Nach der bildbasierten Georeferenzierung bewegen sich die mittleren Abweichungen innerhalb beider Untersuchungsgebiete zwischen 50 und 72 mm (Tab. 1 und Tab. 2). Dies ist vergleichbar mit erreichbaren Genauigkeiten mit bildbasiertem MMS im urbanen Straßenraum unter erschwerten Bedingungen. Für das portable bildbasierte MMS bot der VSH ebenfalls erschwerte Bedingungen: einerseits wenig geometrische Bedingungen für den LiDAR-SLAM (geringe Anzahl ausgedehnter Ebenen und gerader Kanten) und andererseits ungünstige Lichtverhältnisse für die bildbasierte Georeferenzierung. Unter einfacheren Bedingungen (strukturreiche Umgebung in Gebäuden) wurden in BLASER ET AL. (2018) mit demselben MMS um Faktor zwei bessere Genauigkeiten erreicht.

Tab. 1: Zusammenstellung der Genauigkeitsuntersuchungen im «Sicherheitsstollen» (1). Die Präzision repräsentiert die Standardabweichung der räumlichen Vorwärtseinschnitte. Die Genauigkeit steht für den mittleren Betrag der Differenzvektoren zwischen den Passpunktmessungen und den tachymetrisch bestimmten Fixpunkten.

	Präzision	Genauigkeit [mm]		
	[mm]	SP	KP Kerngebiet	Alle KP
SLAM	79	177	257	1'470
SLAM opt.	66	130	153	263
Bildbasiert	10	61	60	155

Tab. 2: Zusammenstellung der Genauigkeitsuntersuchungen im «SCAUT-Stollen» (2). Die Präzision repräsentiert die Standardabweichung der räumlichen Vorwärtseinschnitte. Die Genauigkeit steht für den mittleren Betrag der Differenzvektoren zwischen den Passpunktmessungen und den tachymetrisch bestimmten Fixpunkten.

	Präzision	Genauigkeit [mm]		
	[mm]	SP	KP Kerngebiet	Alle KP
SLAM	150	428	950	629
SLAM opt.	101	237	289	211
Bildbasiert	3	50	72	157

5 Fazit und Ausblick

Mit dem vorgestellten portablen bildbasierten MMS *BIMAGE Backpack* wurden Aufnahmen im VSH in den zwei unterschiedlich ausgeprägten Untersuchungsgebieten "Sicherheitsstollen" und "SCAUT-Stollen" durchgeführt. Sowohl die SLAM-basierte, als auch die bildbasierte Georeferenzierung wurden qualitativ und quantitativ untersucht. Aufgrund der Punktwolken wurde die Georeferenzierung jeweils qualitativ beurteilt. Für die quantitative Untersuchung der Georeferenzierung wurden signalisierte Fixpunkte in mehreren Bildern gemessen und deren Koordinaten mithilfe des Vorwärtseinschnitts im lokalen Bezugsrahmen berechnet.

Die empirischen Standardabweichungen der Vorwärtseinschnitte als gutes Maß für die relative Messgenauigkeit lagen bei den SLAM-Lösungen in der Größenordnung von 66 bis 150 mm und mit der bildbasierten Georeferenzierung bei 3-10 mm. Dies entspricht einer Steigerung der relativen 3D-Messgenauigkeit in den resultierenden 3D-Bilddiensten um mehr als eine Zehnerpotenz. Die Ergebnisse zeigen auf, dass bei Relativmessungen wie etwa 3D-Distanzbestimmungen in Innenräumen, die einen Großteil der Messaufgaben ausmachen, Zentimetergenauigkeit erreichbar ist.

Für eine absolute Genauigkeitsuntersuchung wurden die in den Bildern gemessenen Fixpunkte mit einer 3D-Koordinatentransformation in den globalen Bezugsrahmen überführt und die Differenzen zu den tachymetrisch bestimmten Referenzkoordinaten ausgewiesen. Daraus ergab sich eine deutliche Genauigkeitssteigerung von der SLAM-basierten Georeferenzierung mit Standardeinstellungen zur optimierten SLAM-Lösung um einen Faktor drei bis sechs. Die anschließend durchgeführte bildbasierter Georeferenzierung ergab eine weitere Genauigkeitssteigerung um einen Faktor 2-4 mit absoluten Genauigkeiten zwischen 50 und 72 mm. Dies entspricht Genauigkeiten, die mit bildbasierten MMS im urbanen Außenraum unter erschwerten Bedingungen erreicht werden können.

Die Untersuchungsergebnisse zeigen das große Einsatzpotenzial portabler Mobile Mapping Systeme für die unterschiedlichsten Anwendungen unter Tage auf. Diese reichen von der Inventarisierung und Instandhaltung bis hin zur Erstellung eines digitalen Zwillings. Um diesem Ziel näher zu kommen umfasst ein aktueller Entwicklungsschritt die robuste Berechnung der Tiefeninformation in den erfassten Bildern, sodass künftig 3D-Koordinatenmessungen direkt im 3D-Bild per Mausklick vorgenommen werden können. Des Weiteren werden Schnittstellen des 3D-Bilddatendienstes zu BIM prototypisch implementiert sowie die Interaktionsmöglichkeiten mit entsprechender BIM-Software untersucht. Künftig sollen einerseits direkt im 3D-Bilddatendienst BIM-Objekte konstruiert werden können. Andererseits sollen auch bestehende BIM-Objekte – unter anderem zur Kontrolle oder zur Visualisierung – in die 3D-Bilder augmentiert werden können.

6 Dank

Das Forschungsprojekt BIMAGE (18493.2 PFES-ES) wurde von der Innosuisse, der Schweizerischen Agentur für Innovationsförderung, und von der Industriepartnerin iNovitas AG finanziell unterstützt. Die Messungen im Versuchsstollen Hagerbach wurden im Rahmen eines Blockprojekts von Geomatikstudierenden der Fachhochschule Nordwestschweiz FHNW durchgeführt. Das Blockprojekt im VSH wurde durch die Amberg Gruppe freundlicherweise ermöglicht und sowohl materiell als auch finanziell unterstützt.

7 Literaturverzeichnis

- ABRAHAM, S. & FÖRSTNER, W., 2005: Fish-eye-stereo calibration and epipolar rectification. ISPRS Journal of Photogrammetry and Remote Sensing **59**(5), 278-288.
- BLASER, S., CAVEGN, S. & NEBIKER, S., 2018: Development of a Portable High Performance Mobile Mapping System using the Robot Operating System. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 4(1), 13-20.
- BLASER, S., NEBIKER, S. & CAVEGN, S., 2017: System Design, Calibration and Performance Analysis of a Novel 360° Stereo Panoramic Mobile Mapping System. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 4(1/W1), 207-213.
- 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., BLASER, S., NEBIKER, S. & HAALA, N., 2018: Robust and Accurate Image-Based Georeferencing Exploiting Relative Orientation Constraints. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 4(2), 57-64.
- HESS, W., KOHLER, D., RAPP, H., & ANDOR, D., 2016: Real-Time Loop Closure in 2D LIDAR SLAM. IEEE International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, 1271-1278.
- NEBIKER, S., 2017: 3D Imagery for Infrastructure Management Mobile Mapping meets the Cloud. The 56th Photogrammetric Week. Stuttgart: University of Stuttgart, Institute for Photogrammetry.
- 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.
- QUIGLEY, M., CONLEY, K., GERKEY, B., FAUST, J., FOOTE, T., LEIBS, J., BERGER, E., WHEELER, R. & NG, A., 2009: ROS: an open-source Robot Operating System. ICRA workshop on open source software, 3(3.2).
- RECHSTEINER, F. & WISLER, D., 2018: Erstellung eines digitalen Zwillings des Versuchsstollens Hagerbach mittels des BIMAGE Backpacks und anderer 3D-Sensoren und Plattformen. Bachelorthesis, 2018(5), Fachhochschule Nordwestschweiz, Institut Geomatik, (unveröffentlicht).
- SCHÖNBERGER, J. L. & FRAHM, J.-M., 2016: Structure-from-Motion Revisited. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 4104-4113.
- VERSUCHSSTOLLEN HAGERBACH AG, 2018: Versuchsstollen Hagerbach. <u>http://hagerbach.ch/</u> <u>downloads/VSH_Broschuere_deutsch_online.pdf</u>, letzter Zugriff 21.12.2018.

Potenzial ultrahochauflösender und -genauer UAV-basierter 3D-Datenerfassung

MICHAEL CRAMER¹, GOTTFRIED MANDLBURGER¹, DOMINIK LAUPHEIMER¹, NORBERT HAALA¹ & PATRICK HAVEL²

Zusammenfassung: Dieser Artikel präsentiert eine Studie über das Potenzial ultrahochauflösender und -genauer UAV-basierter 3D-Datenerfassung zur Deformationsüberwachung einer Schiffsschleuse und deren Umgebung. Flächenhaft erfasste UAV-Luftbild- und -Laserdaten sollen dafür mit Ergebnissen von klassischen punktbasierten ingenieur-geodätischen Verfahren verglichen werden. Demnach sind 3D-Punkte bzw. Punktwolken mit einer Oualität im Bereich von besser 1 cm zu liefern. Im Rahmen des Beitrags werden Daten und Ergebnisse aus der ersten Flugkampagne analysiert und präsentiert. Die UAV-Daten wurden von zwei verschiedenen Multi-Copter-Systemen erfasst, die am selben Tag geflogen wurden. Die Ergebnisse zeigen das außerordentliche Potenzial qualitativ hochwertiger UAV-basierter Bild- und Lasersensoren und deren Auswertung sowie ggf. auch deren Fusion. Die bisherigen Untersuchungen konnten aber auch zeigen, dass die zu erzielenden Genauigkeiten der Georeferenzierung nicht einfach "skalierbar" sind. Um die geforderte Objektpunktgenauigkeit von besser 1 cm zu erreichen, ist die reine Trajektorieninformation aus GNSS/inertial-Daten nicht ausreichend. Aus diesem Grund spielt die Optimierung im Rahmen einer integrierten Ausgleichung von Bild- und Laserdaten eine wichtige Rolle. Die in diesem Artikel präsentierten Ergebnisse stellen lediglich einen aktuellen Zwischenstand der Auswertungen dar, da die Untersuchungen noch nicht abgeschlossen sind.

1 Einleitung

Der Artikel beschreibt das Potenzial von ultrahochauflösenden und -genauen UAV-basierten Laser- und Bilddaten zur 3D-Objektpunktbestimmung und Punktwolkengenerierung. Die konkrete Anwendung für diese bislang in UAV-Anwendungen in der Regel nicht angestrebten Genauigkeiten ist die 3D-Datenerfassung zur Deformationsüberwachung einer Schiffsschleuse (einschließlich Wehr- und Wasserkraftwerksgebäude) und ihrer Umgebung. Das Untersuchungsgebiet liegt in Hessigheim am Neckar nördlich von Stuttgart. In der Nähe der Schleuse wurde in den letzten Jahren eine Senkung von etwa 1 mm/Jahr bis 3 cm/Jahr relativ zur stabilen Umgebung beobachtet. Um solche Bewegungen zu überwachen, werden bislang die klassischen technisch-geodätischen Verfahren (Tachymetrie, präzises Nivellement, Extensometermessungen) genutzt. Diese liefern typischerweise punktweise Messung an den jeweiligen Strukturen. Im Gegensatz dazu zielen wir auf eine flächendeckende Messung durch UAV-basiertes Monitoring mittels Photogrammetrie und Laserscanning; auch vor dem Hintergrund, dass Veränderungen in bisher nicht überwachten Bereichen auftreten können und ggf. unbemerkt bleiben. Um die obigen Setzungen nachweisen zu können, müssen die Laser- und Bilddatenauswertungen an ihre Grenzen gebracht werden, auch

¹ Universität Stuttgart, Institut für Photogrammetrie (ifp), Geschwister-Scholl-Str. 24D, D-70174 Stuttgart, [vorname.nachname]@ifp.uni-stuttgart.de

² Bundesanstalt für Gewässerkunde (BfG), Referat M5 – Geodäsie, Am Mainzer Tor 1, D-56068 Koblenz, havel@bafg.de

Dreiländertagung der DGPF, der OVG und der SGPF in Wien, Österreich – Publikationen der DGPF, Band 28, 2019

um mit den obigen klassischen, geodätischen Vermessungsverfahren konkurrieren zu können. Die angestrebte Bodenpixelgröße (GSD) und Objektpunktgenauigkeit (für signalisierte und definierte Einzelpunkte) liegt bei 3-5 mm und die Punktdichte der 3D-Punktwolke bei bis zu 800 Punkten/m².



Abb. 1: Projektgebiet Hessigheim (geostete Darstellung). Der mit der PhaseOne zu erfassende Bereich ist durch das roten Umringpolygon markiert. Die Laserbefliegung erfasst ein geringfügig größeres Gebiet (blaues Polygon, 780 x 570 m²). Die signalisierten Passpunkte sind mit Rauten markiert (rot: Pfeilerpunkte, braun: Stativpunkte), die Passflächen für die Laserbefliegung mit gelben Quadraten.

2 Datenerfassung

2.1 Testgebiet Hessigheim

Das Projektgebiet umfasst neben der Schleusenanlage auch größere Teile des Uferbereichs mit Teilen der Ortschaft Hessigheim und einem großen Gärtnereibetrieb. Der Neckar zerteilt das Gebiet in einen westlichen und östlichen Teil. Die Wehranlange mit der Brückenstraße verbindet beide Uferteile. Die maximale Ausdehnung des Geländes beträgt etwa 570 m (Ost-West) x 780 m (Nord-Süd). Das Gelände ist in Neckarnähe überwiegend flach, steigt aber mäßig nach Nordwesten (Weinberge) bzw. stark nach Südosten (bewaldeter Prallhang) an. Diese Geländehöhenvariationen sollten bei den Flugplanungen berücksichtigt werden. Dieses Gebiet wird mittels Bild- und Laserdaten erfasst, wobei der Bereich für die Bilddatenerfassung geringfügig kleiner definiert ist, als der Bereich der Laserbefliegung (Abb. 1). Die Georeferenzierung der UAV-Bilder und Laserdaten basiert im Wesentlichen auf den verfügbaren signalisierten Objektpunkten. Der zentrale Teil des Gebiets, d. h. die Schleuse und das Neckarwehr selbst, wird von einem Netzwerk von neun betonierten Pfeilerpunkten umgeben, die auch als permanente Punkte für das ingenieurgeodätische Vermessungsnetz dienen. Um den deutlich größeren Bereich des Befliegungsgebietes abzudecken, wurden für die Flugkampagnen weitere 10 Punkte auf Vermessungsstativen installiert (sog. "mobile" Passpunkte). Diese Punkte sind nicht dauerhaft markiert, sondern gelten nur für die entsprechende Flugkampagne. Ihre genaue Position wird durch mehrfache statische GNSS-Basislinienmessung bestimmt, wobei die Höhenkomponente durch zusätzliches Nivellement eingemessen wird. Auf die gleiche Art werden auch die Pfeilerpunkte bestimmt. Die Referenzstation für die GNSS-Messungen liegt im Testgebiet. Die Koordinaten aller Punkte müssen für jede Befliegungskampagne neu bestimmt werden, da diese sich aufgrund von Setzungen verändern können. Lediglich der Pfeilerpunkt der GNSS-Referenzstation ist durch seine über 25 m tiefe Gründung als stabil zu betrachten. Aus der geodätischen Netzausgleichung ergeben sich die Passpunktkoordinaten mit einer Genauigkeit (Std.Abw.) von 2-5 mm.

Alle Pfeiler- und mobilen Passpunkte sind für die hochgenaue photogrammetrische Punktbestimmung mit zwangszentrierten Schachbretttafeln signalisiert, die für die geodätische Einmessung durch die GNSS-Antennen bzw. Prismen ersetzt werden können. Die Größe der runden Schachbretttafeln (27 cm Durchmesser) ist so gewählt, dass sie in den Bildern automatisch gemessen werden können (Abb. 2 (links)), um die Fehler der Bildpunktmessungen zu minimieren. Für die absolute Orientierung der Laserdaten wurden fünf zusätzliche satteldach-förmige Bezugsflächen installiert, die auf Vermessungsstativen mit bekannter Position und Orientierung im Raum befestigt sind (Abb. 1). Jede Seite dieser Referenzflächen hat eine Größe von 40 cm x 80 cm.

2.2 UAV-Datenerfassung

Die UAV-basierte Datenerfassung der ersten Messkampagne im März 2018 erfolgte mit dem Riegl VUX-1 Laser in Kombination mit zwei Sony Alpha 6000 Oblique Kameras und der PhaseOne iXU-RS 1000 für die 100 MPix großen Nadir-Bilddaten. Da es aktuell noch keine UAV-Trägerplattformen gibt, die sowohl VUX-1 Laser als auch PhaseOne-Kamera in einem System kombinieren, wurden zwei Multicopter-UAVs eingesetzt, die aber quasi parallel, also am gleichen Tag die Daten erfassten (CRAMER et al. 2018).

Die PhaseOne iXU-RS 1000 Nadirbilder (Sensorformat 11608 × 8708 Pixel, 4.6 µm Pixelgröße, 16 bit RGB-Bilder, für die Auswertung reduziert auf 8 bit-Tiefe) wurden mit dem CopterSystems CS-SQ8-Copter aufgenommen, der speziell für PhaseOne-Kamera-Nutzlasten optimiert ist. Es waren 18 aufeinander folgende Flugmissionen erforderlich, um das Projektgebiet vollständig abzudecken. Die mittlere Fluggeschwindigkeit lag bei 4 m/s. Die mittlere Flughöhe betrug etwa 40 m über dem Boden. Dies führt zu einer nominalen GSD von 3.7 mm für die Nadir-Bilder. 3775 Bilder mit 80/60 Soll-Überlappung sind in 146 zumeist in Nord-Süd-Richtung verlaufenden Fluglinien angeordnet, da die Flussrichtung des Neckar für dieses Testgebiet fast exakt in Nord-Süd-Richtung liegt. Der Fluss selbst wurde nicht überflogen (mit Ausnahme des Schleusenbereichs und der Wehranlage), da seine Breite etwa 100 m beträgt und die resultierende Bildabdeckung am Boden aus der vorgegebenen Flughöhe zu klein wäre um die Flussbreite komplett zu überbrücken. Die Dreiländertagung der DGPF, der OVG und der SGPF in Wien, Österreich – Publikationen der DGPF, Band 28, 2019

Auswertung der PhaseOne-Nadirbilder zerfällt also praktisch in einen westlichen und östlichen Bildverband.



Abb. 2: Signalisierter Passpunkt auf Vermessungsstativ (Durchmesser Schachbrettsignaltafel 27 cm) und zughörige Punktwolkenerfassung aus Lasermessung.

Für die Laserdatenerfassung reichten nur fünf Flüge, die insgesamt 17 Längsstreifen (d. h. Nord-Süd-Streifen), 4 Kreuzstreifen (Ost-West) und 4 Diagonalstreifen umfassen – letztere um den steil ansteigenden bewaldeten Prallhang in der südöstlichen Ecke des Untersuchungsgebiets abzudecken. Zwei weitere zusätzliche diagonale Fluglinien sollten zur weiteren Blockstabilisierung dienen. Bei einer Fluggeschwindigkeit von 8 m/s, einer nominalen Flughöhe von 50 m über Boden, einem Streifenabstand von 35 m, einer Pulswiederholungsrate von 820 kHz, einer Abtastzeilenrate von 133 Hz und einem verwendeten Scannerfeld von 70° beträgt die resultierende Laserimpulsdichte 300-400 Punkte/m² pro Streifen und mehr als 800 Punkte/m² für den gesamten Flugblock aufgrund der nominalen seitlichen Überlappung von 50 %. Diese Flugmissionsparameter garantieren einen Durchmesser des Laserfußabdrucks von weniger als 3 cm auf dem Boden, was eine hohe horizontale Auflösung von 5 cm ermöglicht. Ein Nachweis dieser horizontalen Punktdichte zeigt die Abb. 2 (rechts). Der Durchmesser der 27 cm großen Signalisierungsscheibe wird wie erwartet mit ca. 5-7 Laserpunkten erfasst. Die im Herstellerdatenblatt des Sensors angegebene Entfernungsgenauigkeit beträgt 10 mm (RIEGL 2018).

3 Auswertung der Daten

3.1 PhaseOne iXU-RS 1000 Nadirbilddaten

Die effektive Auflösung der PhaseOne-Bilder wurde durch separate Auflösungsuntersuchungen anhand der Analyse von Siemenssternen geschätzt. Der abgeleitete MTF10-Wert als Maß für die tatsächliche Auflösung entspricht im mittleren Bereich der Bilder in etwa der nominalen GSD. Die Auflösung nimmt zum Bildrand ab. Die Abb. 3 zeigt die Ergebnisse dieser Auflösungsuntersuchung. Die Bilder wurden in einem separaten Vortest erfasst, in dem das Auflösungsmuster aus zwei leicht unterschiedlichen Flughöhen mehrfach überflogen wurde. Der Siemensstern wurde so insgesamt 16 mal an unterschiedlichen Stellen im Bild abgebildet. Die nominelle Bodenpixelgröße im Bereich von 3.2-3.4 mm wird für die Bilder mit der Flughöhe zwischen 35-37 m knapp aufgelöst. Die höheren MTF10-Werte treten für diejenigen Bilder auf, in denen der Stern eher am Bildrand zu liegen kam. Für die niedrigere Flughöhe kann die nominelle Bodenpixelgröße aber nicht ganz erreicht werden. Hier überwiegt der Einfluss der Unschärfe aufgrund der Eigenbewegung der Kamera während der Bildaufnahme. Die Bilder wurden durchgängig mit 1/1000 s Belichtungszeit aufgenommen (ISO 800, maximale Blendenöffnung 1/5.6). Bei einer Fluggeschwindigkeit von 4 m/s ergibt sich daraus bereits eine theoretische Bewegungsunschärfe von 4 mm. Im statischen Umfeld löst die Kamera problemlos die nominelle Bodenpixelgröße auf, was durch einen separaten statischen Test nachgewiesen wurde. Zu berücksichtigen ist auch der Einfluss der Schärfentiefe. Um in der weiteren Auswertung eine möglichst stabile Kamerageometrie zu gewährleisten, wurde das 50 mm-Rodenstock-Objektiv auf eine Distanz von 40 m fixiert, was der im Projekt geplanten mittleren Flughöhe über Grund entspricht. Bei einer vorgewählten Blendenzahl 1/5.6 ergibt sich der Schärfentiefenbereich von ca. 28-68 m. Die niedrige Flughöhe liegt demnach am unteren Ende des Schärfentiefenbereichs, was die Auflösungsqualität beeinflusst.



Abb. 3: Auflösungsuntersuchung (MTF10 Wert [cm] bezogen auf Objektraum, RGB-Farbkanal) Phase-One iXU-RS 1000 Kamera aus zwei unterschiedlichen Flughöhen.

Im Folgenden werden nur die PhaseOne-Bilddaten des westlichen Blockteils bestehend aus 10 einzelnen Flügen mit insgesamt 1971 Bildern berücksichtigt. Die Hauptflugrichtung ist Nord-Süd, mit Ausnahme eines einzelnen Bildverbands, der über der Schleusenanlage in Ost-West-Richtung erfasst wurde. Die Georeferenzierung basiert auf 10 verfügbaren Passpunkten. Nur ein zusätzlicher Punkt wurde als unabhängiger Kontrollpunkt verwendet. Weitere Kontrollpunkte standen für diese Kampagne nicht zur Verfügung. Da alle Punkte mit Schachbretttafeln ausreichender Größe signalisiert sind, wurden deren Bildkoordinaten automatisch gemessen. Der mittlere Rückprojektionsfehler liegt für die signalisierten Punkte bei etwa 0.06 pix. Dies ist 5-fach besser als der Gesamtrückprojektionsfehler, der mit etwa 1/3 pix hauptsächlich durch die sonstigen Verknüpfungspunktmessungen bestimmt wird. Die Objektkoordinaten der Passpunkte werden gemäß ihren Standardabweichungen aus der vorherigen statischen GNSS-Netzmessung separat gewichtet. Diese liegen zwischen 2 und 5 mm (Standardabweichung).



Abb. 4: Differenz der Projektionszentrumskoordinaten (Nordkomponente = Flugrichtung, [m]): Geschätztes Projektionszentrum aus passpunktbasierter AT versus direkt gemessener Projektionszentrumskoordinate aus GNSS/inertial-Beobachtungen.

Es war vorgesehen die Bilder mittels GNSS-gestützter Aerotriangulation bzw. integrierter Sensororientierung zu georeferenzieren. Parallel zu den Bilddaten sind daher zusätzlich GNSS/inertial-Orientierungselemente über das Applanix APX-15 UAV-Board gemessen worden, deren Genauigkeit nach dem Post-Processing bei 2-5 cm für die Position und 0.025- 0.080° für die Orientierung liegt, vorausgesetzt die GNSS-Satellitenkonstellation ist ausreichend und es liegen kurze GNSS-Basislinien vor (APPLANIX 2018). Entgegen der ursprünglichen Planung erfolgte die Zeitsynchronisation zwischen dem Mittenimpuls der Kamera und dem APX-15 Board nur auf 0.1 s genau. Unter der Annahme, dass durch diese ungenaue Synchronisierung Zeitfehler von bis zu 0.05 s auftreten können, werden dadurch nicht reproduzierbare Abweichungen von bis zu 20 cm in der Position verursacht, hauptsächlich in Flugrichtung, bei 4 m/s angenommener mittlerer Fluggeschwindigkeit. Die Abb. 4 zeigt die Differenzen zwischen den aus der AT geschätzten Projektionszentrumskoordinaten und den direkt gemessenen aus GNSS/inertial-Sensorik für 17 ausgewählte Flugstreifen. Dargestellt ist nur die Differenz in der Nordachse, da diese mit der Hauptflugrichtung des Systems zusammenfällt. Die streifenabhängige Systematik ist deutlich zu erkennen. Diese kann durch streifenweise Offset- (und Drift-) Korrekturen kompensiert werden, wie aus der traditionellen GNSS-gestützten Bündelausgleichung bekannt. Der mittlere Offset pro Streifen deutet auf einen konstanten Zeitoffset hin, oder Restfehler in den Hebelarmen zwischen GNSS/inertial-Einheit und Kameraprojektionszentrum. Problematisch sind die Variationen innerhalb des jeweiligen Streifens, die mehr als doppelt so groß sind als die theoretisch zu erwartenden 2-5 cm. Bei den weiteren Auswertungen wurden die GNSS/inertial-Orientierungen daher nur mit einer geringen Gewichtung eingeführt.

Durch in-situ Selbstkalibrierung wird im Rahmen der AT die Kamerageometrie durch zusätzliche physikalische Parameter bestimmt (BROWN, 1971). Obwohl der Bildblock aus zehn Flugmissionen besteht, die an zwei aufeinanderfolgenden Tagen geflogen wurden, wurde die Kamera nur mit einem Satz von Kalibrierungsparametern modelliert.

Die hohen Bildüberlappungen und die geringe Flughöhe in Kombination mit dem gewählten Weitwinkelobjektiv ermöglichen eine theoretische 3D-Objektpunktqualität, die in horizontaler Richtung bei besser 1 Pixel und in vertikaler Richtung besser 1.5 Pixel liegen sollte. Da wegen der aufgrund der Zeitsynchronisation deutlich schlechteren GNSS/inertial-Orientierungselementen alle signalisierten Punkte (bis auf eine Ausnahme) als gewichtete Passpunkte verwendet wurden, lassen die Residuen an diesen Passpunkten nur eine grobe Qualitätsabschätzung zu. Die mittleren Abweichungen sind 2.0 mm für die Ost-, 4.1 mm für die Nord- und 6.6 mm für die vertikale Koordinate. Dies liegt bei der GSD von knapp 4 mm im Rahmen der Erwartungen. Wie bereits erwähnt, wurde nur ein einziger Punkt als unabhängiger Kontrollpunkt verwendet. Die Residuen betragen 5.5 mm, 0.9 mm und 5.6 mm für Ost-, Nord- und Vertikalkoordinate. Der Residuenvektor passt betragsmäßig gut mit den Residuen an den Passpunkten überein.

3.2 Riegl VUX-1 Laserdaten

Parallel zur photogrammetrischen Bündelausgleichung wurde eine Laser-Streifenausgleichung mit zeitabhängiger Bahnkorrektur durchgeführt (GLIRA et al. 2016, MANDLBURGER et al. 2017). Im Rahmen eines Kalibrierungsverfahrens wurden die Hebelarme und Boresight-Parameter, eine globale Offsetkorrektur sowie Flugbahnkorrekturen geschätzt, um die Abweichungen (Punkt-zu-Ebene-Abstände überlappender Flugstreifen) zu minimieren. Zur absoluten Orientierung werden die Referenzflächen berücksichtigt. Im Rahmen dieser Laserstreifenanpassung wurden zwei verschiedene Versionen untersucht. Die erste Version berücksichtigt nur einen konstanten Versatz ($\Delta x, \Delta y, \Delta z, \Delta roll, \Delta pitch, \Delta yaw$) pro Streifen, d. h. die Trajektorieninformationen aus GNSS/inertial-Daten werden nicht verformt, sondern nur verschoben und insgesamt leicht gedreht. In der zweiten Version werden zusätzliche Polynomkorrekturen eingeführt, die die a-priori-Flugbahninformationen aus der GNSS/inertial-Verarbeitung flexibler machen und die Unterschiede zwischen den Streifen weiter minimieren. Die Restdifferenz in den Vertikalen nach der Streifenanpassung betrug 4 mm, gemessen als robuste Standardabweichung der absoluten Streifenhöhenunterschiede (σ_{MAD}) in glatten und offenen Oberflächenbereichen.

Es sollte erwähnt werden, dass die Georeferenzierung von Laserpunktwolken in erster Linie von der Qualität der direkten Georeferenzierung abhängt. Bei dem APX-20 UAV GNSS/inertial-Board wird die Genauigkeit im Post-Processing mit 2-5 cm für die Positionierung festgelegt. Diese Genauigkeit des APX-15 UAV Boards, welches für die PhaseOne-Befliegung genutzt wurde. Da neben der Position vor allem auch die Genauigkeit der Orientierungswinkelbestimmung für die Orientierung von Scannern wesentlich ist, liefert das APX-20 Board hier eine höhere Genauigkeit: Die Orientierungswinkel werden mit einer Genauigkeit von 0.015-0.035° bestimmt, wobei sich der erste Wert auf die Roll- und Neigungswinkel bezieht, der zweite Wert auf die Qualität der Kurswinkelbestimmung (APPLANIX 2018). Da das Lasersystem in 50 m Höhe über dem Boden geflogen wurde, verursachen Fehler in Roll- / Neigungswinkel etwa 1.5 cm Positionsverschiebung

auf dem Boden. Dies Einfluss ist etwas kleiner als die angenommenen Fehler in der Positionsbestimmung, jedoch größer als die angestrebte 3D-Objekpunktgenauigkeit. Dies zeigt bereits, dass mit der ausschließlichen direkten Georeferenzierung die geforderte Genauigkeit nicht erreicht werden kann, sondern zusätzliche Referenzdaten und Streifenausgleichung erforderlich sind.

Die Genauigkeit der Laserstreifenausgleichung wurde an den vorhandenen Passpunkten ermittelt. Die kreisförmigen Schachbrettsignalisierungstafeln mit einem Durchmesser von 27 cm bilden eine definierte Fläche für die Lasermessungen (Abb. 2 rechts). Die Koordinaten dieser Punkte sind bekannt und können für den Genauigkeitsvergleich herangezogen werden. Acht Referenzpunkte stehen dafür zur Verfügung. Um die Differenzen zu ermitteln wird zunächst in einem Kugelvolumen (Radius 10 cm) um den Mittelpunkt der Signaltafel nach Laserpunkten gesucht. Aus den sechs nächstgelegenen Laserpunkten wird dann eine Ebene geschätzt. Da die Passmarken horizontal ausgerichtet sind, sollte diese geschätzte Ebene auch horizontal sein. Der vertikale Abstand (Normalrichtung) zwischen dieser Ebene und dem Referenzpunkt wird berechnet und kann somit als absolute Genauigkeit betrachtet werden. Die berechneten Standardabweichungen betragen 2.9 cm bzw. 3.0 cm, je nachdem ob für den Streifenausgleich nur Offset- bzw. zusätzliche Polynomkorrekturen verwendet wurden. Diese Abweichungen liegen deutlich über den angestrebten Genauigkeiten und auch über der Herstellerangabe zur Entfernungsmessgenauigkeit des Lasersystems. Das deutet auf verbleibende Restfehler bei der Georeferenzierung hin, die in den laufenden Untersuchungen noch festzustellen und zu korrigieren sind.

3.3 Punktwolkenvergleich

Sowohl Photogrammetrie durch dichte Bildpunktzuordnung als auch Laser liefern dichte Punktewolken. Die Abb. 5 zeigt einen ersten qualitativen Vergleich zwischen photogrammetrischer und Laserpunktwolke. Beide Methoden erlauben eine dichte und genaue Darstellung der jeweiligen Oberflächengeometrie. Der Profilschnitt in Abb. 5 (links) zeigt aber auch die Unterschiede beider Methoden. Ein dichtes Mehrbildstereo-Zuordnungsverfahren liefert 3D-Informationen für jedes Bildpixel, sofern von ausreichender Bildüberlappung und -textur ausgegangen werden kann. Im Gegensatz dazu ist das polare Messprinzip von Lasersensoren immer dann von Vorteil, wenn sich die Struktur des Objekts aus verschiedenen Blickrichtungen stark ändert. Dies gilt beispielsweise für semi-transparente Objekte wie Vegetation. Der Baum im Profilschnitt in Abb. 5 (links) wird detaillierter mit seiner ganzen Aststruktur erfasst. In der photogrammetrischen Punktwolke fehlt er vollständig. Anders die Situation bei definierten Oberflächen. Die Dachflächen z.B. können deutlich vollständiger erfasst werden. Außerdem ist die Punktdichte der photogrammetrischen Punktwolke durch die kleine GSD signifikant höher als das Abtastintervall des Lasers. Das ist deutlich in Abb. 5 (mittig, rechts) erkennbar. Das rechts dargestellte Laser-DSM wurde auf 5 cm Schrittweite abgetastet. Dies passt zu der nativen Abtastung der Laserpunkte, wenn nur ein Flugstreifen berücksichtigt wird, ist also im Vergleich zu der originalen Punktdichte (Streifenüberdeckung 50%) um den Faktor 2 heruntergerechnet. Die GSD der Nadir-Bilder liegt, wie bereits erwähnt, bei etwa 4 mm. Das photogrammetrisch abgeleitete DSM wird hier als 1 cm-Raster angegeben, also ebenfalls um den Faktor 2 verringert. Insbesondere bei den Detailvergrößerungen der Dachziegelstrukturen ist die höhere Auflösung der bildbasierten Rekonstruktion im Vergleich zum Ergebnis der Lasermessung deutlich zu erkennen. Auch in diesen beiden Bildern sind die systemimmanenten Unterschiede bei der Erfassung von Vegetation zu erkennen.



Abb. 5: Vergleich zwischen Laser- und photogrammetrischer Punktwolke. Profilschnitt (links) und schattiertes Oberflächenmodell aus dichter Punktzuordnung (Rasterweite 1 cm, mitte) bzw. Laser (Rasterweite 5 cm, rechts).

Der direkte Vergleich zwischen photogrammetrischer und Laserpunktwolke durch Differenzbildung erlaubt die umfassende Beurteilung der Qualität der Georeferenzierung beider Datensätze. Bislang wurden Bild- und Laserdaten mit den Standardverfahren getrennt ausgewertet. Wenn die Differenz beider Oberflächenmodelle klein ist, dann belegt das auch die Qualität der Georeferenzierung. In Abb. 6 (oben) wird die Differenz zwischen dem Laseroberflächenmodell (Streifenausgleich mit Splinekorrektur) und dem Oberflächenmodell aus der photogrammetrischen Punktwolke gezeigt. Die Abbildung zeigt deutliche systematische Unterschiede zwischen den beiden DOMs, die als eine Art globale Schwingung bezeichnet werden könnten. Dieses wird durch Restsystematiken in der Georeferenzierung der Sensoren verursacht. Die Abweichungen können einerseits durch den Laserstreifenausgleich oder Systematiken in der Bündeltriangulation der Bilddaten (inklusive in-situ Kamerakalibrierung) verursacht sein. Die Wahl der Parameter im Streifenausgleich wirkt sich deutlich auf die spätere 3D-Punktwolke aus. Eine gewisse Rolle spielt auch die Gewichtung der Passpunkte im Rahmen der Bündeltriangulation der Bilder. Beide Effekte werden ebenfalls in Abb. 6 gezeigt. Die in der Mitte gezeigte DOM-Differenz ergibt sich aus der unterschiedlichen Gewichtung der Passpunkte in der Bündelausgleichung der Bilder. In der Regel werden die Passpunkte mit der angenommenen Standardabweichung als beobachtete Größe in die Ausgleichung eingeführt. In diesem Fall wie oben beschrieben mit Werten von 2-5 mm, die sich aus der geodätischen Netzbestimmung ergaben. Im zweiten Fall wurden die Passpunkte als tatsächlich feste Größen behandelt, d.h. die Abbildungsstrahlen mussten sich an der vorgegebenen Passpunktkoordinate schneiden. In den von den Passpunkten am weitesten entfernten Bereichen ergeben sich so Differenzen im Bereich von ca. 2 cm, was bereits über der hier veranschlagten Projektgenauigkeit liegt. Deutlich größer sind aber die Unterschiede bei unterschiedlichen Freiheitsgraden im Laserstreifenausgleich. Die Abb. 6 (unten) zeigt den Vergleich des Laser-DOM nach Bias- bzw. Splinekorrektur. Die Differenzen erreichen bis zu ± 5 cm.

Da es aktuell an unabhängigen Kontrollinformationen vor Ort mangelt, können die Ursachen für die obigen Restfehler in der Georeferenzierung der Datensätze nicht klar zugewiesen werden. Zudem ist es momentan nicht möglich genau einzuordnen welche der berechneten Varianten nun die beste Lösung liefert. Alle hier durchgeführten Vergleiche sind in diesem Sinne nur relativ. Zusätzliche Referenzhöhen sollten in unmittelbarem zeitlichen Zusammenhang durch terrestrische Vermessung nachgemessen werden. Diese Punkte sind praktisch wie Profile angelegt, die das gesamte Projektgebiet durchziehen. Erst dann können die systematischen Fehler eindeutig der Photogrammetrie oder dem Laserscanning zugeordnen werden. Die Messung dieser Referenzprofile erfolgte im Herbst 2018, die finalen Koordinaten konnten für diese Auswertungen aber noch nicht berücksichtigt werden.



Abb. 6: Punktwolkendifferenzen ([m], geostete Kartendarstellung): Vergleich zwischen Laser-DOM (nach Splinekorrektur) und photogrammetrischem Oberflächenmodell (gewichtete Passpunktbeobachtungen in AT, (oben)). Vergleich photogrammetrische DOMs unter Verwendung von gewichteten bzw. festen Passpunktbeobachtungen (mittig). Vergleich zwischen Laser-DOMs nach Bias- bzw. Splinekorrekturmodell im Streifenausgleich (unten).

4 Diskussion und Ausblick

Die obigen Abschnitte präsentieren den aktuellen Stand der Auswertungen im Rahmen dieses Projekts. Mit den bisherigen Auswertungen konnte die avisierte geometrische 3D-Objektpunktgenauigkeit noch nicht erreicht werden. Die Adressierung einer Genauigkeitsstufe von besser 1 cm für derartige UAV-Anwendungen ist definitiv nicht trivial. Die ausschließliche Trajektorieninformation aus GNSS/inertial-Daten erscheint dafür nicht ausreichend. Daher spielt die nachgeschaltete integrierte Ausgleichung eine wichtige Rolle bei der Optimierung der Georeferenzierung. Bisher werden die beiden Datensätze als separate Daten betrachtet, aber im nächsten Schritt werden die Bilder und Laserdaten in einem gemeinsamen Laserbildanpassungsprozess zusammengeführt. Den erfolgreichen Einsatz der hybriden Ausgleichung hat GLIRA (2018) bereits für ein Teilgebiet des Hessigheim Flugblocks demonstriert.

5 Literaturverzeichnis

- APPLANIX, 2018: APX UAV product specifications. https://www.applanix.com/products/dguavs.htm.
- BROWN, D.C., 1971: Close range camera calibration. Photogrammetric Engineering, **37**(8):855-866, https://www.asprs.org/wp-content/uploads/pers/1971journal/aug/1971_aug_855-866.pdf.
- CRAMER, M., HAALA, N., LAUPHEIMER, D., MANDLBURGER, G. & HAVEL, P., 2018: Ultra-high Precision UAV-based LiDAR and Dense Image Matching. The Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., 42(1), 115-120, https://doi.org/10.5194/isprs-archives-XLII-1-115-2018.
- GLIRA, P., PFEIFER, N. & MANDLBURGER, G., 2016: Rigorous Strip Adjustment of UAV-based Laser Scanning Data Including Time-Dependent Correction of Trajectory Errors. Photogrammetric Engineering & Remote Sensing, 12, 945-954.
- GLIRA, P., 2018: Hybrid Orientation of LiDAR Point Clouds and Aerial Images. Dissertation, Technische Universität Wien.
- MANDLBURGER, G., WENZEL, K., SPITZER, A., HAALA, N., GLIRA, P. & PFEIFER, N., 2017: Improved topographic models via concurrent airborne LiDAR and dense image matching. IS-PRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., IV-2/W4, 259-266, https://doi.org/10.5194/isprs-annals-IV-2-W4-259-2017.
- RIEGL, 2018: RIEGL VUX-1UAV product specifications. http://www.riegl.com/up-loads/tx_pxpriegldownloads/RIEGL_VUX-1UAV_Datasheet_2017-09-01_01.pdf.

Dreiländertagung der DGPF, der OVG und der SGPF in Wien, Österreich – Publikationen der DGPF, Band 28, 2019

Untersuchungen zur Leistungsfähigkeit aktueller bildgebender UAV-Sensoren

HEINZ JÜRGEN PRZYBILLA¹

Zusammenfassung: Das geodätisch-photogrammetrische Testfeld auf dem Areal des Industriedenkmals Zeche Zollern in Dortmund bietet ein Szenario zur Durchführung geometrischer und radiometrischer Tests von UAV-Systemen. Die Basis hierzu bildet ein geodätisches Präzisionsnetz (Lage- und Höhengenauigkeit ca. 2 mm) mit insgesamt 45 Bodenpasspunkten, verteilt über eine Fläche von ca. 7 ha. Im Rahmen einer im Herbst 2017 durchgeführten Kampagne wurden verschiedene UAV-Sensorsysteme unter vergleichbaren Bedingungen getestet. Dieser Beitrag stellt geometrische Untersuchungen zweier aktueller DJI-Kameras, Zenmuse X4S (20Mpix) und X5S (20,8 Mpix), sowie einer Phase One IXU 1000 (100Mpix) vor. Die Zenmuse Kameras geben den aktuellen Entwicklungsstand des Herstellers DJI wieder, während das Mittelformat-Kamerasystem von Phase One vorrangig im klassischen Luftbildsegment zu finden ist. Allerdings macht der Wunsch nach gesteigerten Messgenauigkeiten (z. B. für Ingenieuranwendungen) einen derartigen Hochleistungssensor auch für UAV-Applikationen interessant.

Voraussetzung für den Systemvergleich bilden, neben der Konfiguration des Testfeldes, identische Parameter für die Flugplanung, hier insbesondere Längs- und Querüberdeckung, die Durchführung einer vollständigen Kreuzbefliegung in unterschiedlichen Flughöhen sowie die Festlegung einer einheitlichen Bodenauflösung (GSD=14mm).

Die Untersuchung zeigt deutliche Unterschiede in der erreichbaren Qualität der Kameras. Das hochpreisige Phase One System zeigt die besten Ergebnisse, allerdings liefert das preiswerteste System, die Zenmuse X4S, nur geringfügig schlechtere Resultate, während die Zenmuse X5S deutlich abfällt. Hierfür verantwortlich ist vorrangig das mechanisch wenig stabile Kamerakonzept mit Wechselobjektiven. Der abschließende Vergleich der Software-Produkte Pix4Dmapper und Agisoft PhotoScan zeigt zum Teil signifikante Unterschiede in den Ergebnissen der Bildorientierung.

1 Einleitung

Die UAV-basierte Bildaufnahme stellt mittlerweile ein etabliertes geodätisches/photogrammetrisches Datenerfassungsverfahren dar, welches in vielen Anwendungsbereichen mit unterschiedlichen Anforderungen an die Qualität der Ergebnisse – vom Dezimeter bis zum Millimeter – und wegen der großen Flexibilität eingesetzt wird. Im Rahmen der System-Beschaffung sind die Nutzer in der Regel an hochentwickelten UAV-Plattformen interessiert, vernachlässigen dabei vielfach sich hinreichend über die Qualität der integrierten Kamera zu informieren. Dabei stellt die Auswahl und Qualität/Geometrie der Kamera den wichtigsten Einflussfaktor mit Blick auf die zu erreichenden Ergebnisse dar (z. B. 3D-Punktwolken, Orthofotos, etc.). Der Markt digitaler Kameras, die in UAV-basierten Szenarien eingesetzt werden, ist einem kontinuierlichen Wandel unterlegen; es erscheinen sehr häufig neue Kameras, so dass bestehende Systeme schnell veralten.

¹ Hochschule Bochum, Labor für Photogrammetrie, Lennershofstr. 140, D-44801 Bochum, E-Mail: [Heinz-Juergen.Przybilla]@hs-bochum.de

UAV des Herstellers DJI verfügen über eine erhebliche Markt Präsenz und finden dabei vorrangig Verwendung in Film- und Videoproduktionen. Allerdings haben sie zwischenzeitlich auch bei geodätischen Anwendungen eine weite Verbreitung gefunden. Die DJI-Kameras Zenmuse X4S (20Mpix) und X5S (20,8 Mpix) geben den aktuellen Entwicklungsstand des Herstellers wieder, sind konzeptionell jedoch sehr unterschiedlich aufgebaut (Tab. 1). Während die proprietären DJI-Kameras in Verbindung mit der Phantom 4 Pro- und Inspire 2-Serie verwendet werden, ist die Phase One IXU 1000 (100Mpix) ein Mittelformat-Kamerasystem, das typischerweise auf klassischen Luftbildplattformen eingesetzt wird. Für die vorliegenden Tests wurde ein leistungsstarkes UAV des Anbieters Coptersystems (COPTERSYSTEMS 2018), mit einem maximalen Startgewicht von 10 kg, eingesetzt. Die erreichbare Flugdauer mit diesem Kopter beträgt ca. 20 Minuten und liegt damit in einer mit den DJI-Flugplattformen vergleichbaren Größenordnung (Abb. 1).

Hersteller	Phase One / Copter- systems		DJI
Kamera	Phase One IXU1000	Zenmuse X5S	Zenmuse X4S (Phantom 4 Pro)
Objektiv	Rodenstock 50/5.6	DJI MFT ASPH 15/1.7	Integriert 8.8/2.8
Wechselobjektiv	Ja (mechanisch stabi- lisiert)	Ja	Nein
Fokussierung	mechanisch: ∞	elektronisch: ∞	elektronisch: ∞
Verschluss (Shutter)	Zentral (mechanisch)	Rolling	Zentral (mechanisch)
Pixelauflösung [MPx]	100	20,8	20
Sensorformat [mm]	53.4×40.0	17.3×13.0	13.2×8.8
Pixelanzahl	11.608×8708	5.280×3956	5.472×3648
Pixelgröße [µm]	4,6	3,28	2,4
Brennweite [mm]	50	15	8,8
Öffnungswinkel (FOV) / (Diagonal)	67,4°	73,6°	83,8°
Preis [€]	50 - 60.000(*)	2.200	800

|--|

(*) abhängig von der verwendeten Optik

Dreiländertagung der DGPF, der OVG und der SGPF in Wien, Österreich – Publikationen der DGPF, Band 28, 2019



DJI Zenmuse X4S (20Mpix)

DJI Zenmuse X5S (20,8Mpix)

Phase One IXU1000 (100Mpix)



DJI Phantom 4 Pro

DJI Inspire 2

Coptersystems Multicopter

Abb. 1: Eingesetzte Kameras und Trägerplattformen

In der manntragenden Luftbild-Photogrammetrie ist die Einrichtung und Nutzung von Testfeldern seit vielen Jahrzehnten ein probates Mittel für Untersuchungen zur Qualität photogrammetrischer Aufnahmesysteme. Exemplarisch ist hier das Testfeld Vaihingen/Enz zu nennen (CRAMER & KRAUß 2008). Auch für die in-situ Kalibrierung digitaler Luftbildkameras kommen Testfelder zum Einsatz (MUELLER & NEUMANN 2016). Den Ansprüchen zur Prüfung von UAV-Systemen werden diese Testfelder jedoch nicht gerecht: neben der (an den klassischen Bildflug angepassten) Testfeldgröße fehlt es vorrangig an der Qualität der Referenzpunkte, die im Falle eine UAV-Testfeldes im Sub-Zentimeter Bereich liegen müssen.

Ziel der vorliegenden Untersuchung ist es, die beteiligten Kameras hinsichtlich ihrer geometrischen Qualität vergleichend zu untersuchen. Im Fokus der Untersuchungen stehen dabei auch Parameter, die Einfluss auf die geometrische Qualität der Bildverbände ausüben können.

2 Das UAV-Testfeld Zeche Zollern

Das UAV-Testfeld auf dem Areal des Industriemuseums Zeche Zollern in Dortmund (LWL 2018) wurde 2014 durch die Hochschule Bochum (HSBO) eingerichtet und bisher bei verschiedenen Kampagnen genutzt (NEX et al. 2015; PRZYBILLA et al. 2015; GERKE & PRZYBILLA 2016; CRAMER et al. 2017; PRZYBILLA et al. 2017).

Die Zeche Zollern wurde 1898-1904 am westlichen Stadtrand von Dortmund ($\delta 51^{\circ} 31' 4'' N, 7^{\circ} 20' 5'' O$) im Stadtteil Bövinghausen im Jugendstil errichtet. Nach ihrer Stilllegung (Ende der 1960er Jahre) integrierte der Landschaftsverband Westfalen-Lippe (LWL) die Zeche 1981 in das

dezentrale Westfälische Industriemuseum (Wikipedia 2018). Die Zeche ist heute nicht nur Museum mit festen und wechselnden Ausstellungen, sondern auch Ort für Naherholung sowie kulturelle und wissenschaftliche Veranstaltungen.

Die Fläche des von der HSBO eingerichteten UAV-Testfeldes umfasst nahezu das komplette Areal der Zeche. Seine Ausdehnung beträgt $320 \text{ m} \times 220 \text{ m}$ (ca. 7 ha). Die höchste vertikale Ausdehnung ist über zwei ca. 40 m hohe Fördergerüste gegeben. Es besteht aus 45 rasterförmig angeordneten signalisierten Bodenpunkten (Abb. 2)

Eine Übersicht des UAV-Testfeldes mit der geodätischen Passpunktkonfiguration zeigt Abb. 3 (PRZYBILLA et al. 2018).



Abb. 2: UAV-Testfeld Zeche Zollern. Oben: Target zur Signalisierung der GCP sowie Blick auf das Areal Unten: Punktwolke mit signalisierten Bodenpasspunkten (UAV-Befliegung vom 16.10.2017) Dreiländertagung der DGPF, der OVG und der SGPF in Wien, Österreich – Publikationen der DGPF, Band 28, 2019



Abb. 3: Workflow der geodätischen Messungen und Datenprozessierung

3 Aufnahmekonzept

Um eine Vergleichbarkeit der Kamerasysteme im Testszenario zu gewährleisten, wurde vorab für alle Systeme eine einheitliche Bodenauflösung (GSD) von 14 mm festgelegt (Fluganordnung: Regular / R). Daraus resultierend ergaben sich unterschiedliche Flughöhen für die Plattformen (Tab. 2). Ergänzt wurde dies durch Flüge in einer Kreuzanordnung (Cross / C), wobei die Flughöhe jeweils um 20% von den Normalanordnungen abweicht (Abb. 4). Hintergrund für die Erweiterung des Aufnahmekonzepts um die Kreuzanordnung ist deren positive Auswirkung auf die in-situ Kalibrierung der jeweiligen Kamera (PRZYBILLA et al. 2015; GERKE & PRZYBILLA 2016).

Tab. 2: Systeme,	Fluganordnung	und -höhe
------------------	---------------	-----------

Kamera	Phase One IXU1000	DJI Zenmuse X5S	DJI Zenmuse X4S
Regular (R)	120m	60m	50m
Cross (C)	148m	72m	60m



Abb. 4: Fluganordnung Regular (R) und Cross (C) (Block Zenmuse X4S/Phantom 4Pro)

4 Untersuchungsergebnisse

Im Fokus der Untersuchungen stehen Parameter, die Einfluss auf die geometrische Qualität der Bildverbände ausüben (Abb. 5). Ziel ist es daher, den Einfluss der Parameter auf die Geometrie der Bildverbände zu evaluieren.



Abb. 5: Einflussgrößen auf geometrische Qualität der Bildverbände

Bewertungsgrößen sind die Restabweichungen an Pass- und Kontrollpunkten (RMSE-Werte) als Ergebnisse der Bündelausgleichungen mit der Auswertesoftware Agisoft PhotoScan und Pix4Dmapper. Die Berechnungsvarianten gliedern sich bezüglich der Parameter:

- Blocklagerung (Passpunktanzahl und -anordnung)
- Anzahl der Kameras im Prozess der Selbstkalibrierung (Innere Orientierung)
- Fluganordnung (Regular R / Cross C)
- Eingesetzte Auswertesoftware

Die Parametrisierung der Inneren Orientierung der jeweils verwendeten Kamera wurde für alle nachfolgenden Berechnungen mit Agisoft PhotoScan identisch gewählt (Kamerakonstante: f; Hauptpunktlage: cx, cy; radial-symmetrische Verzeichnung: k1 - k3; Affinität und Nichtorthogonalität: b1, b2; tangential-asymmetrische Verzeichnung: p1, p2).

4.1 Auswirkungen der Blocklagerung

Die Anzahl und Anordnung der Passpunkte übt in der Regel einen signifikanten Einfluss auf die Blockgeometrie aus. Auswirkungen von verschiedenen Passpunktkonfigurationen auf die Endprodukte (DGM) wurden u. a. von LINDSTAEDT & KERSTEN (2018) in mehreren UAV-basierten Befliegungen in Äthiopien gezeigt. Eine hohe Anzahl von Passpunkten stabilisiert einerseits den Block, erfordert andererseits jedoch einen erheblichen Aufwand bei der terrestrischen Vermessung vor Ort. Da UAV-Systeme differentielles GNSS (RTK) zur Blocklagerung bisher nur in geringem Maße zur Verfügung stellen, orientiert sich die Verteilung der Passpunkte häufig an Schemata aus der Zeit des "analogen" Bildflugs (Abb. 6). Abweichend von dem hier Dargestellten werden in der Praxis alle Punkte als 3D-Koordinaten – häufig mittels GNSS – gemessen.



Abb. 6: Schematische Anordnung von Passpunkten beim (analogen) Bildflug (Kraus 1994)

Die Abb. 7-9 zeigen die Auswirkungen der variierenden Passpunktverteilungen für die drei untersuchten Kameras. Die dargestellten Ergebnisse basieren auf einer Lagerung der Blöcke mit 45, 22, 12 sowie 5 Passpunkten (GCP). Diese sind im Wesentlichen rasterförmig über das Areal verteilt (Abb. 2). Die Reduktion der Passpunkte führt im Gegenzug zu einer Erhöhung der Anzahl der Kontrollpunkte, von 0 über 23, 33 bis hin zu 40 CP und bietet somit die Möglichkeit einer Bewertung der Blöcke.

Die Basis für alle Berechnungen bildet eine Kreuzbefliegung (RC), die sich jeweils aus zwei separaten Teilflügen zusammensetzt. Das verwendete Bilddatenformat ist im Folgenden ein unkomprimiertes TIF, welches mit der Software Capture One aus den aufgezeichneten RAW-Bilddaten abgeleitet wurde. Für die in Abb. 7 dargestellten Ergebnisse wird EINE zu berechnende Innere Orientierung (UNIFIED – identisch für beide Teilflüge) eingeführt.



Abb. 7: RMSE-Werte an Passpunkten (GCP, links) und Kontrollpunkten (CP, rechts). Fluganordnung: RC. EINE Innere Orientierung (UNIFIED). Berechnung: Agisoft PhotoScan. HINWEIS: die Skalierung der RMSE-Werte variiert zwischen GCP und CP.

Der Vergleich der RMSE-Werte an den Kontrollpunkten (an denen eventuell vorliegende Modellverbiegungen detektiert werden können) zeigt signifikante Unterschiede zwischen den verschiedenen Systemen. Das großformatige Phase One Kamerasystem liefert die geringsten RMSE-Werte – ein Ergebnis, dass bei allen Passpunktanordnungen erkennbar ist.

Selbst in der Minimalanordnung (5 GCP – in den Blockecken und der -mitte) führt dies zu 3D-Genauigkeiten, die deutlich unterhalb der GSD von ca. 14 mm liegen. Sowohl das große Bildformat als auch die mechanische Stabilität der Kamera sind hierfür verantwortlich.

Die im Phantom 4 PRO verbaute Zenmuse X4S zeigt nur geringfügig schlechtere Ergebnisse. Der 3D RMSE-Wert liegt auch bei minimaler Passpunktbesetzung unterhalb der GSD. Das in der Zenmuse X4S realisierte Kamerakonzept mit einem Fixfokusobjektiv und mechanischem Zentralverschluss verfügt über eine hohe Stabilität und zeigt – zumindest im Rahmen der hier zeitnah durchgeführten Flüge – nahezu metrische Eigenschaften.

Die Resultate für die Zenmuse X5S fallen deutlich schlechter aus. Die vorhandene Wechseloptik in Verbindung mit der elektronischen Fokussierung auf ∞ führen zu einem in sich wenig stabilen Kamerasystem, das am Bajonett mechanisches Spiel zwischen Objektiv und Kamerakörper aufweist. Die RMSE-Werte sind gegenüber den Vergleichssystemen um einen Faktor von 2-3 schlechter und liegen nur bei dichter Passpunktanordnung (hier 22 GCP) unterhalb der GSD.

4.2 Auswirkungen separater Parameter für die Innere Orientierung

Die Durchführung der Kreuzbefliegung über separate Flüge hat ihren wesentlichen Grund in der Tatsache, dass die verfügbaren Bildflug-Planungstools (z. B. Map Pilot, Pix4DCapture) eine Kreuzbefliegung mit variierender Flughöhe nicht zulassen. Darüber hinaus bedingt die Datenspeicherung im RAW-Format bei den DJI-Systemen die Reduktion der Fluggeschwindigkeit, damit genügend Zeit für das Speichern der Bilder auf der SD-Karte verfügbar ist (ca. Faktor 3 höherer Zeitbedarf im Vergleich zum "DJI-JPG"). Infolge dessen ist für die DJI-Systeme vor jedem Flug ein neues Setup durchzuführen, da diese durch den Wechsel der Akkus zwischenzeitlich stromlos sind und ihre Einstellungen verlieren.

Dieser Fakt lässt es sinnvoll erscheinen, im Rahmen der Kreuzbefliegungen die Bündelblockausgleichungen mit separaten Inneren Orientierungsparametern für die Teilflüge durchzuführen. Die Resultate sind in Abb. 8 zusammengefasst. Zwar sind die Ergebnisse der Varianten "RC UNI-FIED" bzw. "RC SEPARATE" ähnlich, es ist jedoch erkennbar, dass die Variante "RC SEPA-RATE" zu Genauigkeitssteigerungen führt, insbesondere für die Zenmuse X5S. Dies ist ein deutliches Indiz für die im Vergleich zur X4S und zur Phase One weniger stabilen Innere Orientierung. Für die DJI Systeme wäre es aus den o. a. Gründen sinnvoll, beim Wechsel der Stromquellen die Kameras durch einen Puffer-Akku eingeschaltet lassen zu können, so dass das Setup erhalten bleibt (ein leider bisher unerhörter Wunsch seitens photogrammetrischer Anwender an die DJI Entwickler!).

4.3 Auswirkungen der Einfach-Befliegung

Eine vollständige Kreuzbefliegung stellt in der Regel einen erheblichen Mehraufwand über die gesamte Prozesskette – von der Aufnahme bis zur Auswertung – dar. Beim Einsatz von Kameras mit stabilen metrischen Eigenschaften (typisch: digitale Luftbildkameras) kann jedoch auf diese erweiterte Fluganordnung verzichtet werden.



Abb.8: RMSE-Werte an Passpunkten (GCP, links) und Kontrollpunkten (CP, rechts). Fluganordnung: R+C. ZWEI Innere Orientierungen (SEPARATE). Berechnung: Agisoft PhotoScan. HINWEIS: die Skalierung der RMSE-Werte variiert zwischen GCP und CP.



Abb. 9: RMSE-Werte an Passpunkten (GCP, links) und Kontrollpunkten (CP, rechts). Fluganordnung: R. Berechnung: Agisoft PhotoScan. HINWEIS: die Skalierung der RMSE-Werte variiert zwischen GCP und CP.
Die Auswirkungen einer Einfach-Befliegung (hier: Regular – R) auf die untersuchten UAV-Kameras sind sehr deutlich (Abb. 9). Während die mechanisch stabileren Systeme – Phase One und Zensmuse X4S – zwar mit tendenziell schlechteren Ergebnissen aufwarten, sind mit der Zenmuse X5S deutliche Genauigkeitsverluste erkennbar. Besonders betroffen ist die Höhengenauigkeit in Verbindung mit einer reduzierten Passpunktanzahl zur Lagerung des Blocks. Die größeren RMSE-Werte (in Z-Richtung) sind im Wesentlichen auf Probleme der numerischen Bestimmung des Parameters "Kamerakonstante" zurückzuführen, für den der vorliegende R-Block offensichtlich zu wenig Tiefeninformation bietet.

Die positive Wirkung einer sehr dichten Anordnung von Passpunkten ist in diesem Block für die instabile X5S deutlich erkennbar. Die Ergebnisse der großformatigen Phase One sind dagegen nahezu unabhängig von der Passpunktverteilung, wie an den RMSE-Werten der Kontrollpunkte erkennbar ist. Prinzipiell gilt diese Aussage auch für die Zenmuse X4S.

4.4 Vergleich der eingesetzten Software-Produkte

Die Menge der Software-Produkte zur Generierung bildbasierter Punktwolken steigt stetig. Zwar kann davon ausgegangen werden, dass im Rahmen von Arbeitsprozessen vielfach auf die gleiche Software zurückgegriffen wird, jedoch sind die Ergebnisse unterschiedlichen Software-Einsatzes häufig nicht identisch.

Auf der Grundlage der Zollern-Testdatensätze wurden daher identische Projekte mit der Software Agisoft PhotoScan (AGISOFT 2018) und Pix4Dmapper (PIX4D 2018) aufgesetzt. Beide Softwares sind kommerzielle Produkte und verfügen über eine nicht unerhebliche Marktpräsenz.

Die erstellten Projekte entsprechen in ihrer Konfiguration den Vorgaben für die jeweilige Software. Soweit verfügbar wurde auf implementierte "Vorlagen" (Standardparameter) zurückgegriffen. Die Messung der signalisierten Bodenpunkte erfolgte auf der Grundlage der in den Programmen installierten Messroutinen manuell durch denselben Auswerter.

Die eingeschlagene Vorgehensweise kann aus diesem Grunde als typisch für eine Projektbearbeitung betrachtet werden. Die resultierenden Ergebnisse sind somit weitestgehend unbeeinflusst von "erweitertem Expertenwissen".

Abb. 10 zeigt die dreidimensionalen RMSE-Werte nach der Bündelblockausgleichung, wobei hier auf eine getrennte Darstellung von Lage- und Höhenabweichungen verzichtet wird. Die Unterschiede der RMSE-Werte sind für alle untersuchten Kameras und den jeweiligen Blocklagerungen deutlich erkennbar. Dabei resultieren aus den Berechnungen mit Pix4Dmapper in der Regel deutlich geringere RMSE-Werte als aus den Vergleichsprojekten mit Agisoft PhotoScan. Diese Feststellung trifft sowohl auf die Passpunkte, als auch auf die Kontrollpunkte zu. Grundsätzlich fällt es schwer, hierfür Gründe zu finden, da Verfahrens-technisch gleiche Methoden vorliegen sollten. Software-intern sind jedoch unterschiedliche Vorgehensweisen zu vermuten, die aber von Anwendern eher weniger evaluiert werden können.

Davon unabhängig sind die in Abb. 10 dargestellten Ergebnisse von gewisser Relevanz, da sie eine der wesentlichen Bewertungsgrundlagen für die Bildorientierung darstellen. Insbesondere im Rahmen von Ingenieurprojekten, die im Genauigkeitsniveau weniger Millimeter anzusiedeln sind, fällt eine Qualitätsbewertung der erzielten Ergebnisse eher schwer.



Abb. 10 Vergleich der 3D RMSE-Werte an Passpunkten (GCP, oben) und Kontrollpunkten (CP, unten) als Ergebnisse der Bündelblockausgleichungen mit PIX4Dmapper und Agisoft PhotoScan. HIN-WEIS: die Skalierung der RMSE-Werte variiert zwischen GCP und CP.

5 Fazit & Ausblick

Die auf der Grundlage von vergleichbaren Datensätzen des UAV-Testfelds Zeche Zollern durchgeführten in-situ Prüfungen verschiedener UAV-Systeme zeigen erkennbare Unterschiede in der Leistungsfähigkeit der eingesetzten Kameras. Bemerkenswert sind dabei die sehr guten Ergebnisse der preiswertesten Kamera im Test. Die positiven Auswirkungen der Kreuzbefliegung (mit Unterschieden in der Flughöhe von ca. 20%) konnten bestätigt werden. Leider bieten die Bildflug-Planungstools diese Variante in der Regel nicht als Standardkonzept an, so dass hier ein gewisser Handlungsbedarf seitens der Anbieter bestehen sollte. Damit einher geht die Notwendigkeit der Selbstkalibrierung für die eingesetzten Kameras. Keine der untersuchten Kameras kann – unter den spezifischen Anforderungen eines hochpräzisen UAV-Bildflugs – als metrische Kamera bezeichnet werden. Unter Berücksichtigung des individuellen Kamerakonzepts kann es daher auch sinnvoll sein, Teilflüge mit eigenen Parametern der Inneren Orientierung in die Bündelblockausgleichung einzuführen.

Der abschließende Vergleich von Berechnungsergebnissen mit unterschiedlichen Softwaren ist in gewisser Weise erstaunlich, da er mit den vorliegenden Datensätzen deutliche Qualitätsunterschiede zeigt. Hier sollten weitere Untersuchungen (auch der Software-Anbieter) ansetzen.

Dank

Der Autor dankt allen Teilnehmern der Zollern Herbstkampagne für ihre tatkräftige Unterstützung. Neben den Partnern von PhaseOne und Coptersystems gilt ein besonderer Dank den Studierenden der Wahlpflichtveranstaltung "Ingenieurvermessung" 2017 des Studiengangs Vermessung an der HSBO (Leitung: Prof. Dr. Manfred Bäumker) für ihre motivierte Mitwirkung bei der geodätischen Netzmessung!

6 Literaturverzeichnis

AGISOFT, 2018: Webseite http://www.agisoft.com/, letzter Zugriff 28.12.2018.

COPTERSYSTEMS, 2017: Webseite https://coptersystems.com/, letzter Zugriff 28.12.2018.

- CRAMER, M., 2016: Welche ist am besten? Ein paar Anmerkungen zur Auswahl von Kamerasystemen in der UAS-Luftbildphotogrammetrie. UAV2016 – Vermessung mit unbemannten System, Schriftenreihe des DVW, 82, 97-118, Wißner-Verlag, Augsburg.
- CRAMER, M.; PRZYBILLA, H.-J., MEIBNER, H. & STEBNER, K., 2017: Kalibrierung und Qualitätsuntersuchungen UAV-basierter Kamerasysteme. Unmanned Aerial Vehicles 2017 (UAV 2017), Schriftenreihe des DVW, 86, 67-84, Wißner-Verlag, Augsburg.
- GERKE, M. & PRZYBILLA, H.-J., 2016: Accuracy analysis of photogrammetric UAV image blocks: influence of onboard RTK-GNSS and cross flight patterns. Zeitschrift für Photogrammetrie, Fernerkundung und Geoinformation (PFG), **2016**(1), 17-30.
- KRAUS, K., 1994: Photogrammetrie, Band 1. Grundlagen und Standardverfahren. Ferdinand Dümmlers Verlag, Bonn, 263 S.
- LINDSTAEDT, M. & KERSTEN, T., 2018: Zur Bedeutung von Passpunkten bei der Aerotriangulation UAV-basierter Bildflüge. UAV 2018 - Vermessung mit unbemannten Flugsystemen, Schriftenreihe des DVW, 89, 81-101, Wißner-Verlag, Augsburg.
- LWL 2018: Homepage des Industriemuseums Zeche Zollern, https://www.lwl.org/industriemuseum/standorte/zeche-zollern, letzter Zugriff: 28.12.2018.
- MUELLER, C. & NEUMANN, K., 2016: LEICA DMC III Calibration and Geometric Sensor Accuracy. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **40**(3/W4), 1-9.

- PHASEONE, 2018: Webseite https://industrial.phaseone.com/Mapping_Surveying_ Applications.aspx, letzter Zugriff 28.12.2018.
- PIX4D, 2018: Webseite https://www.pix4d.com/, letzter Zugriff 28.12.2018.
- PRZYBILLA, H.-J., REUBER, C., BÄUMKER, M. & GERKE, M., 2015: Untersuchungen zur Genauigkeitssteigerung von UAV-Bildflügen. Publikationen der Deut. Gesellschaft für Photogrammetrie, Fernerkundung u. Geoinformation e.V., 35. Wissenschaftlich-Technische Jahrestagung der DGPF, 24, 45-54.
- PRZYBILLA, H.-J., 2017: Kameras für UAS Eine Marktübersicht. In: Unmanned Aerial Vehicles 2017 (UAV 2017), Schriftenreihe des DVW, **86**, 51-66, Wißner-Verlag, Augsburg.
- PRZYBILLA, H.-J., BÄUMKER, M. & VIETEN, J., 2018: Das UAV-Testfeld Zeche Zollern in Dortmund. UAV 2018 - Vermessung mit unbemannten Flugsystemen, Schriftenreihe des DVW, 89, 61-80, Wißner-Verlag, Augsburg.
- WIKIPEDIA, 2018: Industriemuseum Zeche Zollern, Dortmund-Bövinghausen. https://de.wikipedia.org/wiki/Zeche Zollern, letzter Zugriff: 28.12.2018.

Abgrenzung homogener Waldstücke in ALS-Punktwolken

MORITZ BRUGGISSER¹, MARKUS HOLLAUS¹, DI WANG² & NORBERT PFEIFER¹

Zusammenfassung: Wir stellen ein flexibles Rahmenwerk vor, mit dem sich homogene Waldflächen in ALS-Punktwolken abgrenzen lassen. Für die Segmentierung werden vorab Waldstrukturmetriken aus der Punktwolke berechnet. Kernstück der Segmentierung ist ein iterativer k-means Clustering-Schritt im Merkmalsraum. Beginnend mit der gesamten Punktwolke als initiales Cluster, werden die Punkte solange in zwei Sub-Cluster aufgespaltet, bis jedes Sub-Cluster eine gewünschte Homogenität aufweist. Wir demonstrieren die Funktionsweise der entwickelten Methode für zwei Testgebiete für die Segmentierung von Waldstücken mit homogenen Eigenschaften in Bezug auf den Wasserkreislauf. Die Segmentierungen zeigen Konsistenzen von R^2 =0.65-0.91 in den relevanten Waldstrukturmetriken.

1 Einleitung

Studien zu Segmentierungen homogener Waldflächen aus flugzeuggestütztem Laserscanning (ALS) zielten bisher vorwiegend darauf ab, Waldbestände einheitlicher Spezies, Baumhöhen, Altersklasse und Bestandesdichten abzugrenzen (DIEDERSHAGEN et al. 2004; KOCH et al. 2009; MUS-TONEN et al. 2008; SULLIVAN et al. 2009; WU et al. 2014). Die so gewonnenen Wald-Einheiten folgen der Bestandesdefinition von KOIVUNIEMI & KORHONEN (2006) und zeichnen sich durch Homogenität hinsichtlich des Holzvorrats und der Biomasse aus. Da Waldbestandeskarten typischerweise auf historischen Managementaktivitäten beruhen, sind solche aus ALS-Daten abgeleiteten Bestandeskarten, welche die tatsächliche Waldstruktur widergeben, wertvoll für die Forstwirtschaft (KOCH et al. 2009; HOLLAUS et al. 2015). Naturnähere Waldbewirtschaftungsformen und Sukzessionsprozesse führen jedoch zu einer Diversifizierung der Wälder hinsichtlich der Baumhöhen, des Baumalters und der Schichtigkeit, sowie zur Koexistenz mehrerer Baumarten (KOCH et al. 2014). Um die erhöhte Komplexität erfassen zu können, muss daher die vertikale Waldstruktur besser berücksichtigt werden, welche in den oben genannten Studien zu Bestandesabgrenzungen nicht berücksichtigt ist. Das Potential von ALS zur Beschreibung der vertikalen Struktur wurde bereits aufgezeigt (z.B. COOPS et al. 2007; LEITERER et al. 2015; PALACE et al 2015). Da der Laser in die Kronenschicht einzudringen vermag, lassen sich durch ALS die kroneninternen Verhältnisse im dreidimensionalen Raum detailliert erfassen.

In dieser Studie stellen wir einen Segmentierungsansatz vor, um homogene Waldbereiche innerhalb einer ALS-Punktwolke abzugrenzen. Durch die freie Festlegung der Waldstrukturparameter, welche für die Segmentierung verwendet werden, ist der Ansatz sehr flexibel einsetzbar. Dies trägt der Tatsache Rechnung, dass eine Abgrenzung von homogenen Waldbereichen nicht eine einzige, eindeutige Lösung ergeben kann, sondern dass die Homogenität immer in Bezug auf eine Anwen-

¹ TU Wien, Department of Geodesy and Geoinformation, Gußhausstraße 27-29/E120,

A-1040 Wien, E-Mail: [Moritz.Bruggisser, Markus.Hollaus, Norbert.Pfeifer]@geo.tuwien.ac.at ² Aalto University, Department of Built Environment, Otakaari 24, FI-00076 Aalto, E-Mail:

di.wang@aalto.fi

dung zu sehen ist. Da unser Ansatz insbesondere den Einbezug vertikaler Waldstruktureigenschaften für die Abgrenzungen erlaubt, können Waldsegmentierungen hinsichtlich homogener ökologischer Prozesseigenschaften vorgenommen werden. Wir zeigen dies am Beispiel der Abgrenzung von Waldflächen mit homogenen Wasserkreislauf-Eigenschaften. Die resultierenden Entitäten folgen damit nicht der strengen Bestandesdefinition von KOIVUNIEMI & KORHONEN (2006), sondern sind homogene Waldstücke.

2 Untersuchungsgebiet

Für die Studie wurden zwei ALS Punktwolken-Datensätze aus Österreichischen Bundesland-Befliegungen verwendet, welche eine große Breite von Wald- und Gebietscharakteristiken abdecken.

2.1 Burgenland

Die Burgenlandszene ist auf 5280900 N, 606300 E (UTM 33N) zentriert und deckt eine Fläche von 1200 m x 2400 m (Nord x West) ab. Die Gebietshöhe reicht von 415 m bis 590 m ü. M. Das 95%-Quantil der Vegetationshöhe liegt bei 24.1 m, wobei die Szene eine große Heterogenität in der Kronenhöhe umfasst, unter anderem auch Schlagflächen. Weiter zeichnet sich das Gebiet durch eine Vielfalt in den Baumarten aus. Die verwendeten Daten wurden im April 2010 unter *leaf-off* Bedingungen mit den Riegl Scannern LMS-Q560 und LMS-Q680 erhoben. Die mittlere Flughöhe betrug 400 m über Grund, der Median der Punktdichte liegt bei 24 Punkten / m2.

2.2 Ötscher

Die Ötscher-Szene ist auf 304540 N, -84970 E (MGI/Austria GK East) zentriert und deckt eine Fläche von 1415 m x 1550 m (Nord x West) ab. Charakteristisch für die Szene ist der große Höhengradient mit Gebietshöhen zwischen 800 m und 1650 m ü. M., was im südlichen Bereich der Szene die Baumgrenze mit einschließt. Die Szene deckt eine Vielfalt an Waldtypen hinsichtlich der Baumhöhen, Kronendichte und Kronenschichtmächtigkeit ab. Das 95%-Höhenquantil der Vegetationshöhe liegt bei 28.8 m. Die Datenerhebung fand im Januar 2007 unter *leaf-off* Bedingungen statt. Verwendet wurde ein Riegl LMS-Q560 Sensor, die mittlere Flughöhe betrug 620 m über Grund. Der Median der Punktdichte liegt bei 19 Punkten / m².

3 Methodik

Der entwickelte Ansatz besteht aus einem Vorprozessierungsschritt, in dem Attribute berechnet werden, welche die Waldstruktur beschreiben, und dem eigentlichen Segmentierungsschritt, in dem die Abgrenzung der homogenen Waldstücke erfolgt (Abb. 1). Der Segmentierungsschritt stellt eine geschlossene Pipeline dar und besteht wiederrum aus drei Schritten, nämlich i) einem *Splitting Schritt*, ii) der *Elimination zu kleiner Cluster* und iii) der *Eliminiation nicht-spezifischer Cluster*. In jedem dieser drei Schritte findet eine Iteration statt und der nächste Schritt wird erst dann erreicht, wenn die Iteration im vorgängigen Schritt abgeschlossen ist. Die Segmentierung wird im Merkmalsraum und auf Punkt-Entitäten durchgeführt.



Abb. 1: Ablaufdiagramm zur Abgrenzung homogener Waldstücke. Dargestellt ist die Attributberechnung (Kap. 3.1) und das eigentliche Segmentierungs-Framework (Iterative Splitting Segmentation framework, Kap. 3.2). Das Ablaufdiagramm zeigt die drei Prozessierungsschritte innerhalb der *Iterative Splitting Segmentation* und benennt die dabei verwendeten Grenzwerte und Attributbezeichnungen.

3.1 Attributberechnung

Die Kronenstruktur eines Waldes beeinflusst den Wasserkreislauf durch Interzeption und Evapotranspiration. Die Interzeptionsfähigkeit wird dabei vorwiegend durch die Blattfläche, den funktionellen Baumtyp und die Kronendichte bestimmt (BONAN 2015; MIRALLES et al. 2010). Für die Transpirationsfähigkeit der Kronenschicht ist vorwiegend die Blattfläche entscheidend (BONAN 2015; ALMEIDA & SANDS 2016). Diese Charakteristiken können aus ALS-Punktwolken nicht direkt ermittelt werden. Hingegen lassen sich aus ALS-Punktwolken Metriken ableiten, welche als Proxies für die gesuchten Größen dienen können. Wir verwendeten die *fractional cover* und die Anzahl Layer in der Kronenschicht (*vg_nLayers*) als Proxy für die Blattfläche, die Kronendichte *d*40 und die Kronenmächtigkeit (*vg_layerDiff*) als Proxy für die Dichte der Kronenschicht und die Kronendichte *d*50 für Unterscheidung der funktionellen Baumtypen (HoLLAUS et al. 2009).

3.1.1 Waldstrukturmetriken

Um die Waldstrukturmetriken aus der ALS-Punktwolke abzuleiten, wird ein Suchzylinder eines bestimmten Radius auf einen Punkt der Punktwolke zentriert. Basierend auf den Punkten innerhalb der zylindrischen Nachbarschaft lässt sich für jeden Punkt die entsprechende Metrik berechnen:

Fractional cover

Die *fractional cover* ist ein dimensionsloser Parameter, der den Anteil des Bodens innerhalb einer Referenzfläche angibt, der durch Vegetation bedeckt wird. Die Berechnung aus der Punktwolke erfolgt wie in MORSDORF et al. (2006) dargestellt als:

$$fractional \ cover = \frac{N_{Kronenschicht}}{N_{total}}$$

wobei *N_{Kronenschicht}* die Anzahl Echos aus der Kronenschicht bezeichnet und *N_{total}* die gesamte Anzahl der Echos innerhalb des Suchzylinders.

Kronendichte

Gemäß HOLLAUS et al. (2009) wird die Kronendichte berechnet als:

$$dQQ = \frac{N_{hQQ}}{N_{total}}$$

Wie bei der *fractional cover* entspricht N_{total} der gesamten Anzahl der Punkte innerhalb des Suchzylinders. N_{hQQ} bezeichnet die Anzahl der Punkte oberhalb eines gewissen Höhenanteils, bezogen auf die Höhe des obersten Punktes innerhalb des Suchzylinders. Wir verwenden das 40%und das 50%-Höhenmaß, um die Kronendichten *d*40 und *d*50 zu berechnen.

Vegetationsprofile

COOPS et al. (2007) folgend, wurden zuerst gap-Wahrscheinlichkeitsprofile Pgap(h) berechnet als:

$$P_{gap}(h) = \frac{N_{\{h>z\}}}{N_{total}}$$

wobei $N_{\{h>z\}}$ die Anzahl Punkte oberhalb der Höhe z angibt. Die gap-Wahrscheinlichkeitsprofile wurden für ein Interval dz = 1 m berechnet und anschließend in plant area index-Profile (PAI(h)) transformiert über PAI(h) = - ln(P_{gap}(h)). Die erste Ableitung von PAI(h) wiederrum entspricht den Vegetationsprofilen. Wie in PALACE et al. (2015) besprochen, haben wir daraus die Anzahl Layer (vg_nLayers) als die Anzahl der lokalen Maxima, die Höhe des obersten Layers (vg_topLayer) als Höhe des obersten lokalen Maximums, sowie die Kronenmächtigkeit (vg_layerDiff) als Differenz zwischen unterstem und oberstem lokalen Maximum bestimmt.

3.1.2 Maßstabsebene

Die Abgrenzung homogener Waldstücke erfordert eine gewisse Generalisierung von Variabilitäten auf kleiner Maßstabsebene, wie sie beispielsweise durch Lücken zwischen Bäumen eingeführt werden. Solche Variabilitäten sind der Abgrenzung hinderlich und sollen deshalb möglichst unterdrückt werden (DIEDERSHAGEN et al. 2004). Wir erreichten dies, indem wir Suchradien von 2 m, 5 m und 10 m für die Suchzylinder für die Berechnung der Waldstrukturmetriken verwendet haben. Eine Ausnahme bildet das Attribut *vg_nLayers*, welches für die Segmentierung nur für einen 10 m Suchradius berechnet wurde. Die Kehrseite solch großer Suchradien ist die rechenintensive Suche benachbarter Punkte in der Punktwolke. Um eine effiziente Rechenzeit gewährleisten zu können, haben wir die Strukturmetriken deshalb nicht für alle Punkte berechnet, sondern nur für eine gesampelte Punktwolke. Diese gesampelte Punktwolke wurde gewonnen, indem ein 1 m-Raster über die Punktwolke gelegt wurde und innerhalb jeder Rasterzelle derjenige Punkt aus der ursprünglichen Punktwolke selektiert wurde, der am nächsten am Rasterzentrum liegt. Für die eigentliche Berechnung der Waldstrukturmetriken werden alle Punkte der ursprünglichen Punktwolke hinzugezogen, womit für die Berechnung der volle Informationsgehalt der Punktwolke genutzt wird. Für die Segmentierung wird hingegen die gesampelte Punktwolke verwendet, in der jeder Punkt die Waldstrukturmetriken als Attribute angehängt hat.

3.2 Iterative Splitting Segmentierung

3.2.1 Splitting Schritt

Die gesampelte Punktwolke wird iterativ in zwei Subsegmente gespaltet. Für dieses iterative Zweiteilen der Punktwolkensubsegmente wird der k-means-Algorithmus verwendet (LLOYD 1982). Da dieser Algorithmus auf der Auswertung der Euklidischen Distanz beruht, wurden die berechneten Waldstrukturmetriken auf den Wertebereich [0,1] normalisiert. Für jedes der entstehenden Subsegmente wird wiederum eine Entscheidung getroffen, ob weiter zweigeteilt werden oder ob der Prozess für das jeweilige Subsegment gestoppt werden soll. Dazu werden zwei Kriterien untersucht: i) wird mit dem Prozess für ein Subsegment Seg0, welches einer Teilmenge der ursprünglichen gesampelten Punktwolke entspricht, nur dann weiter verfahren, wenn eine Mindestanzahl an Punkten in Seg0 vorhanden ist. Ist dieses Kriterium erfüllt, durchläuft Seg0 den k-means Clustering-Schritt, was in den Subsegmenten {Seg1, Seg2} resultiert. Die Segmente {Seg1, Seg2} werden dann ii) auf ihre Spezifizität in einem Attribut ftr_thSplit hin überprüft, für welches vom Benutzer eines der Attribute aus den berechneten Waldstrukturmetriken definiert wird. Die Spezifizität wird über die Mahalanobis-Distanz ermittelt, die die Distanz des Mittelwerts im Attribut ftr_thSplit von Seg1 gegenüber der Verteilung von Seg2 misst und umgekehrt. Unterschreitet eine dieser beiden Distanzen einen vom Nutzer definierten Grenzwert th_split, werden die Subsegmente {Seg1, Seg2} als nicht spezifisch genug erachtet, der letzte Teilungsschritt rückgängig gemacht, und die weitere Aufteilung des Subsegments Seg0 gestoppt. Kann keines der Subsegmente weiter aufgespaltet werden, wird der Splitting Schritt gestoppt und die Subsegmente gelangen in den nächsten Prozessierungsschritt.

3.2.2 Elimination zu kleiner Cluster

Der Splitting Schritt führt typischerweise zu einer leichten Übersegmentierung der Punktwolke. Die Cluster zeichnen sich dann zwar durch einen hohen Grad an Homogenität im Attributraum aus, sind aber für die meisten Anwendungen zu klein, um als eigenständige Waldstückklassen verwendet zu werden. Zu kleine Cluster werden deshalb eliminiert, wobei vom Nutzer über den Schwellwert *th_minSegSize* bestimmt werden kann, wie viele Punkte ein Cluster mindestens umfassen muss. Der Eliminationsschritt erfolgt, in dem Cluster, welche *th_minSegSize* unterschreiten, mit jenem andern Cluster verschmolzen werden, zu welchem sie die größte Ähnlichkeit in einem vom Nutzer definierten Attribut *ftr_mergeSegments* aufweisen. Rechnerisch ist dies so umgesetzt, dass ein zu kleines Cluster mit allen andern Clustern verglichen und mit jenem Cluster verschmolzen wird, bei dem die Standardabweichung nach dem Verschmelzen die kleinste Änderung erfährt. Begonnen wird mit dem kleinsten Cluster, welches mit allen andern Clustern aus dem *Splitting* Schritt verglichen wird, also auch mit andern Clustern, welche als zu klein erachtet werden. Der Eliminations-Schritt wird nun solange durchlaufen, bis alle Cluster die Mindestgröße *th_min-SegSize* erreicht haben.

3.2.3 Elimination nicht-spezifischer Cluster

Die Cluster weisen nach Durchlaufen der ersten zwei Prozessierungsstufen zwar einen hohen Homogenitätsgrad auf und verfügen zudem über die definierte Mindestgröße. Allerdings sind sie teilweise nicht spezifisch genug, sondern zeigen Ähnlichkeiten zu andern Clustern. In einem letzten Prozessierungsschritt wird deshalb die Überlappung der Cluster im Merkmalsraum des vom Nutzer definierten Attributs *ftr_mergeOverlap* analysiert. Die Überlappung wird wiederrum über die Mahalanobis-Distanz gemessen wie unter Kap. 3.2.1 für den *Splitting Schritt* dargestellt. Cluster, die einen Distanz-Schwellwert *th_mergeOverlap* zwischen den Clusterverteilungen unterschreiten, werden verschmolzen.

Für die Generierung der homogenen Waldstücke mithilfe des vorgeschlagenen Segmentierungs-Frameworks, müssen den oben genannten und in Abb. 1 gezeigten Schwellwerten Attribute aus den berechneten Waldstrukturmetriken zugewiesen werden. Die Segmentierung wird durch deren Auswertung gesteuert. Für den Wasserkreislauf haben wir die *fractional cover* für *ftr_thSplit* und *d*50 für *ftr_mergeSemgents* und *ftr_mergeOverlap* verwendet, berechnet jeweils basierend auf der Nachbarschaft mit einem Radius von 5 m.

3.3 Validierung

Für die Validierung der abgegrenzten Waldstücke werteten wir die Konsistenz der Segmentierung aus. Diese lässt sich einerseits durch die Betrachtung der Attributverteilungen der abgegrenzten Waldstücke abschätzen, anderseits verwendeten wir die R^2 -Metrik, wie sie von MUSTONEN et al. (2008) und von WU et al. (2014) für die Auswertung von Waldbestandesabgrenzungen genutzt wurde:

$$R^2 = 1 - \frac{Var_{innerhalb}}{Var_{gesamt}}$$

Varinnerhalb misst die Variabilität innerhalb einer Waldstückklasse als:

$$Var_{innerhalb} = \sum_{i=1}^{\kappa} \sum_{j=1}^{n_i} (x_{ij} - \overline{x_i})^2$$

und Vargesamt die Gesamtvarianz der Szene:

$$Var_{innerhalb} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \overline{x})^2$$

In den obigen Gleichungen bezeichnet k die Anzahl der Waldstückklassen, n_i die Anzahl der Punkte innerhalb der Waldstückklasse i, x_{ij} den Attributwert des Punktes j in Klasse i, $\overline{x_i}$ den Attributmittelwert aller Punkte in Klasse i und \overline{x} den mittleren Attributwert innerhalb der gesamten Szene. Weiterhin wurde das Konsistenzmaß den Größen der resultierenden Waldstückklassen gegenüber gestellt.

4 Resultate & Diskussion

In den Abb. 2 und Abb. 3 werden die Resultate der Abgrenzungen für Waldstücke mit homogenen Eigenschaften im Wasserkreislauf für die Burgenland- und die Ötscher-Szene gezeigt. Die Verteilungen der Attribute *fractional cover* und *d*50 innerhalb der resultierenden Klassen sind in Abb. 4 und Abb. 5 dargestellt. Die Tab. 1 und Tab. 3 geben die Konsistenz der einzelnen Segmentierungen für die Attribute *fractional cover* und *d*50 für die beiden Szenen wider. In Tab. 2 und Tab. 4 ist die Größe der resultierenden Waldstück-Klassen bezüglich der darin enthaltener Punkte aufgelistet.



Abb. 2: Abgegrenzte Waldstücke mit homogenen Wasserkreislaufeigenschaften in der Burgenland-Szene. Dargestellt ist die Segmentierung in 4 Klassen (links) und in 5 Klassen (rechts).



Abb. 3: Abgegrenzte Waldstücke mit homogenen Wasserkreislaufeigenschaften in der Ötscher-Szene. Dargestellt ist die Segmentierung in 4 Klassen (links) und in 5 Klassen (rechts).



Abb. 4: Verteilungen der *fractional cover* (obere Zeile) und der Kronendichte *d*50 (untere Zeile), berechnet für eine Punktnachbarschaft mit 5 m Radius, in den vier (links) und fünf (rechts) resultierenden Waldstückklassen in der Burgenland-Szene. Die Whisker sind auf das Minimum, bzw. das Maximum gesetzt. Eine Ausnahme stellt die blaue Klasse in der Verteilung der *fractional cover* dar, welche nur wenige Ausreißer hat. Die Klassen-Label entsprechen den Farben in Abb. 2. Die Boxplots zeigen jeweils auch die Gesamtverteilung der Attribute (als *all* bezeichnet).

Tab. 1: Konsistenz der Waldstück-Klassen für die Attribute *fractional cover* (R^{2}_{fc}) und *d*50 (R^{2}_{d50}) für eine Punktnachbarschaft mit 5 m Suchradius für die Burgenland-Szene und für vier, fünf bzw. sechs finale Waldstück-Klassen.

	Anzahl Waldstück-Klassen				
	4 5 6				
R ² fc	0.67	0.80	0.82		
<i>R</i> ² _{d50}	0.65	0.76	0.79		

Tab. 2: Resultierende Segmentgrößen aus der Waldstück-Abgrenzung für die Burgenland-Szene für vier, bzw. fünf Waldstück-Klassen. Die Tabelle weißt die relative Anzahl enthaltener Punkte pro Klasse aus bei einer Gesamtheit von 2'874'780 Punkten in der Szene.

	Relative Anzahl an Punkten [%]			
Klasse / Anzahl Waldstück-Klassen	4	5		
Blau (blue)	1.04	1.04		
Orange (orange)	5.20	5.20		
Gelb (yellow)	17.94	17.94		
Grün (green)	75.83	47.80		
Violett (purple)		28.03		



Abb. 5: Verteilungen der *fractional cover* (obere Zeile) und der Kronendichte *d*50 (untere Zeile), berechnet für eine Punktnachbarschaft mit 5 m Radius, in den vier (links) und fünf (rechts) resultierenden Waldstückklassen in der Ötscher-Szene. Die Whisker sind jeweils auf das Minimum, bzw. das Maximum gesetzt. Die Klassen-Label entsprechen den Farben in Abb. 3. Die Boxplots zeigen jeweils auch die Gesamtverteilung der Attribute (als *all* bezeichnet).

Tab. 3: Konsistenz der Waldstück-Klassen für die Attribute *fractional cover* (*R*²_{*fc*}) und *d*50 (*R*²_{*d*50}) für eine Punktnachbarschaft mit 5 m Suchradius für die Ötscher-Szene und für vier, fünf bzw. sechs finale Wald-stück-Klassen.

	Anzahl Waldstück-Klassen					
	4 5 6					
R ² fc	0.90	0.91	0.91			
R^{2}_{d50}	0.88	0.90	0.90			

Tab. 4: Resultierende Segmentgrößen aus der Waldstück-Abgrenzung für die Ötscher-Szene für vier, bzw. fünf Waldstück-Klassen. Die Tabelle weißt die relative Anzahl enthaltener Punkte pro Klasse aus bei einer Gesamtheit von 1'983'117 Punkten in der Szene.

	Relative Anzahl an Punkten [%]			
Klasse / Anzahl Waldstück-Klassen	4	5		
Blau (blue)	19.20	19.20		
Orange (orange)	16.99	16.99		
Gelb (yellow)	29.48	29.48		
Grün (green)	34.33	17.51		
Violett (purple)		16.81		

Die Attributverteilungen (Abb. 4 und Abb. 5) lassen eine klare Differenzierung in den Attributen fractional cover und d50 innerhalb des 25%-75%-Quantils der Attributwerte pro Waldstück-Klasse erkennen. Die gut ersichtliche Unterscheidung der Waldstückklassen ist in den Konsistenzmetriken für die beiden Attribute widergespiegelt (Tab. 1 und Tab. 3). Als Vergleich ist jeweils auch die Konsistenz für eine Segmentierung in sechs Klassen angegeben. Die Konsistenz der abgegrenzten Waldstücke nimmt zu, je mehr Klassen erhalten werden. Allerdings zeigt sich speziell in der Segmentierung in der Burgenlandszene, dass der Unterschied in der Konsistenz zwischen vier und fünf Klassen größer ist (0.13 für die fractional cover, 0.09 für d50, Tab. 1) als der Unterschied zwischen fühf und sechs Klassen (0.02 für die fractional cover, 0.03 für d50). In der Ötscher-Szene sind die Konsistenzunterschiede weniger deutlich sichtbar, insbesondere ist die Konsistenz identisch für Segmentierung in fünf und sechs Klassen. Für eine Beurteilung, wie viele Waldstück-Klassen unterschieden werden sollen, kann zusätzlich die Größe der resultierenden Segmente betrachtet werden (Tab. 2 und Tab. 4). Im Falle der Burgenland-Szene tritt im Falle der Segmentierung in vier Klassen ein Segment (grün) auf, das dreiviertel der Punkte beinhaltet. In der Ötschers-Szene sind sich die Segmente ähnlicher hinsichtlich ihrer Größe, auch im Falle der Unterscheidung in vier Waldstück-Klassen. Eine Entscheidung, wie viele Klassen sinnvollerweise unterschieden werden sollen, hängt aber letztlich von der Anwendung ab, für die Waldstück-Polygone generiert werden. Neben dem Konsistenzmaß bietet die Segmentgröße dafür eine gute Entscheidungsgrundlage.

Ein offener Aspekt für eine spätere Anwendung betrifft den räumlichen Zusammenhang der Segmente. Wie die räumliche Verteilung der Segmente zeigt, ist dieser noch nicht gegeben (Abb. 2 und Abb. 3), stellt jedoch ein Kriterium in der Waldbestandesdefinition dar (KOIVUNIEMI & KORHONEN 2006). Der fehlende räumliche Zusammenhang bei unserem Framework kommt dadurch zustande, dass die Lage-Information der Punkte nicht explizit in die Segmentierung einfließt, sondern dass die Abgrenzung vollständig durch Cluster-Bildung im Merkmalsraum erfolgt. Dies ist ein gangbarer Ansatz, da die Attribute, welche den Merkmalsraum aufspannen, die Waldstruktur beschreiben. Die Waldstruktur stellt bis auf Lücken oder Übergängen zu Schlagflächen ein relativ kontinuierliches Attribut-Feld dar. Räumlich benachbarte Punkte, welche wir hier für die Segmentierung verwenden, werden deshalb tendenziell auch im Merkmalsraum näher beieinanderliegen als räumlich weiter entfernt liegende.

Um letztlich auf räumlich zusammenhängende Waldstück-Polygone zu kommen, wäre ein Nachbearbeitungsschritt über Energie-Minimierungsansätze denkbar (DECHESNE et al. 2016).

5 Fazit & Ausblick

Wir haben ein Framework vorgestellt, das die Abgrenzung von homogenen Waldstücken innerhalb von ALS-Punktwolken ermöglicht. Die Stärke des Ansatzes ist, dass das Homogenitätskriterium, nach dem die Abgrenzung erfolgt, vom Nutzer definiert werden kann. Dafür werden vorgängig Waldstrukturmetriken aus der Punktwolke abgeleitet, welche für die eigentliche Abgrenzung verwendet werden. Welche Metriken das sind, kann vom Nutzer spezifiziert und auf die spätere Anwendung der abgegrenzten Waldstücke angepasst werden. Dies erlaubt die Abgrenzung hinsichtlich einheitlicher ökologischer Prozesse, was hier am Beispiel von Waldstücken mit homogenen Bedingungen im Wasserkreislauf gezeigt wurde. Die Auswertung der resultierenden Waldstücke zeigte einen hohen Grad an Konsistenz (R^2 =0.65-0.91) in den relevanten Strukturmerkmalen. Eine Validierung mit Referenzdaten ist jedoch ausstehend. Da hier der Fokus auf einen ökologischen Prozess gelegt wird, lassen sich entsprechende Referenzdaten nicht einfach aus Luftbildern ableiten, sondern müssten mittels Bodenmessungen generiert werden. Dies steht im Gegensatz zu Bestandesabgrenzungen, für die sich die Informationen traditionellerweise aus Luftbildern gewinnen lassen (DIEDERSHAGEN et al. 2004; DECHESNE et al. 2016). Schließlich muss je nach Anwendung die segmentierte Punktwolke noch in ein Raster transformiert und die Durchmischung der Klassen reduziert werden. Unser Framework liefert ein geeignetes Ausgangsprodukt für solche Nachbearbeitungsschritte.

6 Literaturverzeichnis

- ALMEIDA, A.C. & SANDS, P.J., 2016: Improving the ability of 3-PG to model the water balance of forest plantations in contrasting environments. Ecohydrology, 9(4), 610-630. doi:10.1002/eco.1661.
- BONAN, G., 2015: Ecological Climatology: Concepts and Applications (3 ed.), Cambridge University Press, 2015. doi:10.1017/CBO9781107339200.
- COOPS, N.C., HILKER, T., WULDER, M.A., ST-ONGE, B., Newnham, G., Siggins, A. & Trofymow, J.A., 2007: Estimating canopy structure of Douglas-fir forest stands from discrete-return Li-DAR. Trees, 21(3), 295-310. doi:10.1007/s00468-006-0119-6.
- DECHESNE, C., MALLET, C., LE BRIS, A., GOUET, V. & HERVIEU, A., 2016: Forest stand segmentation using airborne lidar data and very high resolution multispectral imagery. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 41(B3), 207-214.
- DIEDERSHAGEN, O., KOCH, B. & WEINACKER, H, 2004; Automatic segmentation and characterisation of forest stand parameters using airborne lidar data, multispectral and fogis data. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 2004, 36(8), 208-212.
- HOLLAUS, M., MÜCKE, W., HÖFLE, B., DORIGO, W., PFEIFER, N., WAGNER, W., BAUERHANSL, C. & REGNER, B., 2009: Tree species classification based on full-waveform airborne laser scanning data. Proceedings of SilviLaser 2009, 14-16 Oktober 2009, College Station, Texas, USA, 54-62.
- HOLLAUS, M., EYSN, L., MAIER. B. & PFEIFER, N., 2015: Site index assessment based on multitemporal ALS data. Proceedings of SilviLaser 2015, 28-30 September 2015, La Grande Motte, France, 159-161.
- KOCH, B., STRAUB, C., DEES, M., WANG, Y. & WEINACKER, H., 2009: Airborne laser data for stand delineation and information extraction. The International Journal of Remote Sensing, 30(4), 935-963. doi:10.1080/01431160802395284.
- KOCH, B., KATTENBORN, T., STRAUB, C. & VAUHKONEN, J., 2014: Segmentation of Forest to Tree Objects. Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies. Managing Forest Ecosystems 27, Maltamo, M., Næsset, E., Vauhkonen, J. (Hrsg.), Springer, Dordrecht, The Netherlands, 89-112.

- KOIVUNIEMI, J. & KORHONEN, K.T., 2006: Inventory by Compartments. Forest Inventory Methodology and Applications. Managing Forest Ecosystems 10, Kangas, A. & Maltamo, M. (Hrsg.), Springer, Dordrecht, The Netherlands, 271-278.
- LEITERER, R., TORABZADEH, H., FURRER, R., SCHAEPMAN, M.E. & MORSDORF, F., 2015: Towards automated characterization of canopy layering in mixed temperate forests using airborne laser scanning. Forests, 6(11), 4146-4167. http://dx.doi.org/10.3390/ f6114146.
- LLOYD, S., 1982: Least squares quantization in PCM. IEEE Transactions on Information Theory **28**(2), 129-137. doi:10.1109/TIT.1982.1056489.
- MIRALLES, D.G., GASH, J.H., HOLMES, T.R., DE JEU, R.A. & DOLMAN, A., 2010: Global canopy interception from satellite observations. Journal of Geophysical Research, **115**(D16). doi:10.1029/2009JD013530.
- MORSDORF, F., KÖTZ, B., MEIER, E., ITTEN, K. & ALLGÖWER, B., 2006: Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction. Remote Sensing of Environment, **104**(1), 50-61. doi:10.1016/j.rse.2006.04.019
- MUSTONEN, J., PACKALEN, P. & KANGAS, A., 2008: Automatic segmentation of forest stands using a canopy height model and aerial photography. Scandinavian Journal of Forest Research, 23(6), 534-545. doi:10.1080/02827580802552446.
- PALACE, M.W., SULLIVAN, F.B., DUCEY, M.J., TREUHAFT, R.N., HERRICK, C., SHIMBO, J.Z. & MOTA-E-SILVA, J., 2015: Estimating forest structure in a tropical forest using field measurements, a synthetic model and discrete return lidar data. Remote Sensing of Environment, 161, 1-11. doi:10.1016/j.rse.2015.01.020.
- SULLIVAN, A.A., MCGAUGHEY, R.J., ANDERSEN, H.E. & SCHIESS, P., 2009: Object-oriented classification of forest structure from light detection and ranging data for stand mapping. Western Journal of Applied Forestry 2009, **24**(4), 198-204.
- WU, Z., HEIKKINEN, V., HAUTA-KASARI, M., PARKKINEN, J. & TOKOLA, T., 2014: ALS data based forest stand delineation with a coarse-to-fine segmentation approach. IEEE 7th International Congress on Image and Signal Processing (CISP), 14–16 October 2014, Dalian, China, 547– 552. doi:10.1109/CISP.2014.7003840.

Waldinventur aus Luftbildern und LiDAR-Daten - Mit welcher Genauigkeit und Präzision lassen sich Baumhöhe, Kronenradius und Kronenansatz von Douglasien ableiten?

SELINA GANZ¹, YANNEK KÄBER¹ & PETRA ADLER¹

Zusammenfassung: In dieser Studie wurde untersucht, wie gut sich die in Forstwirtschaft und Forstwissenschaft relevanten Einzelbaumattribute Baumhöhe, Kronenradius und Kronenansatzhöhe mithilfe verschiedener Sensorsysteme ableiten lassen. Im Fokus standen dabei die Genauigkeit und Reproduzierbarkeit der automatisierten Messungen und die Identifizierung von Unterschieden zwischen den Sensorsystemen. Als Sensorsysteme kamen unbemannte Luftfahrtsysteme (UAV) mit Light Detection And Ranging (LiDAR)-Sensor bzw. RGB-Kamera sowie Gyrocopter mit RGB-Kamera zum Einsatz. Ergänzend wurden Luftbilder aus photogrammetrischen Bildflügen des Landesamtes für Geoinformation und Landentwicklung verwendet. Das Untersuchungsgebiet war eine 50-jährige Douglasien-Versuchsfläche. Die Baumhöhe konnte durch alle Sensorsysteme mit einem RMSE zwischen 0,36 m und 2,89 m abgeleitet werden. Auch der Kronenradius ließ sich mit einem akzeptablen RMSE zwischen 0,45 m und 0,76 m relativ zuverlässig berechnen. Für die Kronenansatzhöhe wurde ein RMSE zwischen 3,88 m und 11,26 m erreicht. Folglich ließen sich Baumhöhe und Kronenradius zuverlässig ableiten, während die Ableitung des Kronenansatzes auf Einzelbaumebene sich als nicht anwendbar erwiesen hat. Selbst die räumlich hochaufgelösten LiDAR-Daten ermöglichten keine erfolgreiche Ableitung des Kronenansatzes. Die Unterschiede zwischen den Ergebnissen aus LiDAR- und photogrammetrischen Daten fielen geringer aus als erwartet, wobei sich bei den photogrammetrischen Daten Hinweise dafür finden, dass eine hohe Punktdichte ein wichtiger Faktor für eine erfolgreiche Ableitung der Einzelbaumattribute ist.

1 Einleitung

Mit modernen Methoden der Fernerkundung ist es möglich, äußere Merkmale von stehenden Bäumen abzuleiten um Waldinventuren zu unterstützen und zu verbessern. Traditionelle Waldinventuren zielen auf den Volumen- und Wertertrag von Waldbeständen. Mit der vermehrten Nutzung von Einzelbaumwuchsmodellen rückt aber auch die Beschreibung der Struktur der Kronen in den Fokus (WEIDENBACH et al. 2012). Die Struktur des Kronenraumes und die sich daraus ergebende Wuchsraumeffektivität und Konkurrenzverhältnisse gehören zu den wesentlichen Untersuchungsgegenständen waldbaulicher bzw. ertragskundlicher Forschungen (PRETZSCH 2014; MARTENS et al. 2012). Detaillierte Versuchsaufnahmen beinhalten daher sowohl Baumhöhen und Brusthöhendurchmesser als auch Baumposition, Kronenradien und die Höhe des Kronenansatzes (MARTENS et al. 2012, WEIDENBACH et al. 2012). Mit den Eingangsgrößen Baumhöhe, Kronenansatzhöhe und Kronenradius kann die räumliche Ausdehnung einer Krone sowie

¹ Forstliche Versuchs- und Forschungsanstalt Baden-Württemberg (FVA), Abteilung Biometrie und Informatik, Wonnhaldestraße 4, D-79111 Freiburg, E-Mail: [Selina.Ganz, Yannek.Käber, Petra.Adler]@Forst.bwl.de

die Kronenschirmfläche, Kronenmantelfläche und das Kronenvolumen bestimmt werden. Des Weiteren sind Aussagen über die Kronenkonkurrenz möglich.

Bisherige Untersuchungen haben meist einen Sensor auf einer spezifischen Fläche angewandt (siehe z.B. DEAN et al. 2009; JUNG et al. 2011; KATO et al. 2009; LUO et al. 2018; MUSS et al. 2011; POPESCU & ZHAO 2008; SOLBERG et al. 2006; VAUHKONEN 2008; WALLACE et al. 2014; XU et al. 2013). Aufgrund der Unterschiede zwischen den fernerkundlichen Systemen und den Rahmenbedingungen in diesen Untersuchungen ist es nicht möglich allgemeingültige Aussagen zur Anforderung an die Eingangsdaten zu machen. Dies ist hauptsächlich der Tatsache geschuldet, dass sich die Untersuchungsgebiete bezüglich Baumarten, Waldtyp oder Bestandesalter unterscheiden. In der vorliegenden Studie liegt ein direkter Vergleich zwischen den fernerkundlichen Systemen mit den gleichen Referenzdaten, teilweise dem gleichen Befliegungsdatum und bei Wiederholungsbefliegungen der gleichen Sensorik und Plattform und teilweise der gleichen Flugplanung vor. Dadurch kann mit einer höheren Sicherheit das Anwendungspotenzial der verschiedenen Sensoren und Plattformen aufgezeigt werden. In dieser Studie wurden die Einzelbaumattribute Baumhöhe, Kronenradius und Kronenansatzhöhe aus unterschiedlichen Systemen abgeleitet. Für den Vergleich dienten sowohl Light Detection And Ranging (LiDAR)-Daten als auch Luftbilder, welche von unbemannten Luftfahrtsystemen (Unmanned Aerial Vehicle, UAV), Gyrocopter und Flugzeug aufgenommen wurden. Der Fokus der Studie lag dabei auf dem Vergleich der verschiedenen Fernerkundungsdaten und der Validierung sowie der Untersuchung von Genauigkeit und Präzision bei der automatisierten Ableitung von Kronenparametern. Die Anwendung der unterschiedlichen Systeme unter gleichen Bedingungen ermöglichte uns für die 50jährige Douglasien-Versuchsfläche die Beantwortung folgender Fragen. (1) Mit welcher Genauigkeit und Präzision können die Kronenattribute mit fernerkundlichen Methoden abgeleitet werden? (2) Gibt es bei der Bestimmung von Kronenattributen Unterschiede zwischen den angewandten Systemen? (3) Was sind die Anforderungen an die Eingangsdaten für die erfolgreiche Ableitung dieser Kronenattribute?

2 Material und Methoden

Dieses Kapitel beschreibt das Untersuchungsgebiet, die Referenz- und Fernerkundungsdaten sowie die Erfassungsmethoden für die Einzelbaumparameter mittels Feldmessungen und Fernerkundung, welche in der vorliegenden Studie Verwendung fanden.

2.1 Untersuchungsgebiet und Versuchsbäume

Das Untersuchungsgebiet war eine 2 ha große Versuchsfläche bei Kandern im Südschwarzwald, welche Teil des koordinierten Standraumversuchs Baden-Württemberg ist. Zum Zeitpunkt der Auswahl der Versuchsbäume wies die homogene Fläche keinerlei Naturverjüngung und kaum Bodenbewuchs auf (siehe Abbildung 1). Im September 2014 standen 227 Douglasien auf der Fläche. Diese hatten einen Brusthöhendurchmesser (BHD) von 40,9 cm (±13,5 cm, Standardabweichung (SD)) in einem Wertebereich von 13,8 cm bis 63,9 cm. Aus diesen Bäumen wurden für die Studie im März 2017 aus allen Durchmesserklassen 30 Bäume ausgewählt. Es wurden Bäume mit unterschiedlich vielen Ästen und Kronengrößen ausgewählt. Im Mittel wiesen die

ausgewählten Bäume eine Baumhöhe von 34,2 m ($\pm 2,8$ m, SD), einen Kronenradius von 3,9 m ($\pm 0,9$ m, SD) und eine Kronenansatzhöhe von 15,3 m ($\pm 2,7$ m, SD) auf (siehe Abbildung 2).



Abb. 1: Versuchsbestand bei der Auswahl der Bäume.



Abb. 2: Histogramme von Baumhöhe, Kronenradius und Kronenansatzhöhe der 30 Versuchsbäume.

2.2 Fernerkundungsdaten

Im Sommer 2017 und im Frühjahr 2018 fand die Datenaufnahme mit verschiedenen fernerkundlichen Systemen statt. Darunter UAVs, welche die Versuchsfläche mithilfe eines LiDAR-Sensors und einer RGB-Kamera erfasst haben. Zusätzlich wurden mit einem Gyrocopter Luftbilder aus zwei unterschiedlichen Flughöhen aufgenommen. Die Daten mit der geringsten räumlichen Auflösung lieferte eine amtliche Standardbefliegung (photogrammetrischer Bildflug) des Landesamts für Geoinformation und Landentwicklung (LGL) aus dem Jahr 2016 (LGL, 2018a). Eine Übersicht über die durchgeführten Befliegungen (UAV-LiDAR 2017, UAV-LiDAR 2018, UAV-Kamera 2017, UAV-Kamera 2018, Gyrocopter 2 cm, Gyrocopter 5 cm) und den Standardbildflug von 2016 ist in Tabelle 1 ersichtlich.

	Datum	Fluggerät	Flughöhe über Grund	Überlappung	
	19.06.17 - 20.06.17	HEIGHT TECH HT-8	~ 80 m – 100 m	50%	
UAV-LIDAR	15.03.18	Tarot X6	~ 50 m – 60 m	60%-70%	
	19.06.17 - 20.06.17	HEIGHT TECH	. 90 m 100 m	80%/80%	
UAV-Kamera	15.03.18	HT-8 C180	~ 80 m – 100 m		
Gyrocopter 2 cm	20.06.17	AutoGyro	~ 300 m	85%/70%	
		MTOSport (D-MDOB)			
Gvrocopter 5 cm	20.06.17	AutoGyro	~ 800 m	90%/70%	
	20.00.11	MTOsport (D-MDOB)			
Standard	10.07.16	Flugzeug einer Standardbefliegung	2700 m – 2750 m	68%/40%	

Tab. 1: Daten zu den Befliegungen 2017/2018 und dem Standardbildflug von 2016.

Für die Laserbefliegungen wurde 2017 ein HEIGHTECH HT-8 und 2018 ein Tarot X6 verwendet. Die UAV wurde jeweils mit dem Laserscanner Mapper I der Firma YellowScan ausgerüstet. Der LiDAR-Sensor verfügte über die Multi-Echo Technologie und konnte pro Laserstrahl 3 Echos erfassen (First, Middle und Last Return). Die aufgenommenen Laserscan-Punkte wurden mithilfe der Software *TerraScan* in Bodenpunkte und Nicht-Bodenpunkte klassifiziert und orientiert.

Bei den Bildflügen mittels UAV wurde als Sensor die Kamera Sony Alpha 7R mit einer Brennweite von 35 mm eingesetzt. Es wurden die drei Kanäle rot, grün und blau (RGB) mit einer radiometrische Auflösung von 8 bit aufgenommen. Die Luftbilder hatten eine geometrische Auflösung von 1 cm. Die Befliegungen mittels Gyrocopter wurden mit der Kamera D800E der Marke NIKON mit einer Brennweite von 85 mm durchgeführt. Es wurden ebenfalls RGB-Bilder mit einer radiometrischen Auflösung von 8 bit aufgenommen. Die geometrische Auflösung lag hierbei – abhängig von der Flughöhe – bei 2 cm und 5 cm. Der Standardbildflug wurde mit der UltraCam XP mit einer Brennweite von 100,5 mm durchgeführt. Die Luftbilder beinhalteten die vier Kanäle rot, grün, blau und infrarot (RGBI) bei 16 bit. Die geometrische Auflösung betrug 20 cm. Tabelle 2 beinhaltet die technischen Angaben zu den Bildflügen.

Tab. 2:	Technische	Angaben z	u den	Bildflügen.

	Kamera	Brennweite	Radiometrische Auflösung	Geometrische Auflösung
UAV-Kamera	Sony Alpha 7R	35 mm	RGB, 8 bit	1 cm
Gyro 2 cm	NIKON D800E	85 mm	RGB, 8 bit	2 cm
Gyro 5 cm	NIKON D800E	85 mm	RGB, 8 bit	5 cm
Standard	UltraCam XP	100,5 mm	RGBI, 16 bit	20 cm

Die Verortung der Bild- und Laserdaten war über die im System verbaute GNSS-Einheit möglich. Zusätzliche wurden für die Bildflüge Passpunkte verwendet. Für die Befliegungen mittels Gyrocopter wurden 14 (5 cm) bzw. 12 (2 cm) und mittels UAV wurden 4 (2017) bzw. 12 (2018) Passpunkte verwendet.

Die Luftbilder von 2017 und 2018 wurden mithilfe der Software Agisoft Photoscan verarbeitet, sodass photogrammetrische Punktwolken und Orthophotos erzeugt wurden. Die Luftbilder aus den Standardbildflügen wurden mit der Software SURE zu Punktwolken und Orthophotos mit einer Bodenauflösung von 20 cm und 50 cm verarbeitet. Die Namen der Datensätze sind in Tabelle 3 mit der entsprechenden Punktdichte dargestellt. Die LiDAR-Punktwolken (2017/2018) bestanden aus 75,5%/67,6% First Returns, 21,5%/26,7% Middle Returns und 3,0%/5,7% Last Returns. Die Durchdringungsrate (Bodenpunkte geteilt durch alle Punkte der Punktwolke) lag bei 3,0% (2017) bzw. 13,5% (2018).

	Punktdichte [Punkte/m ²]
UAV-LIDAR 2017	350
UAV-LIDAR 2018	620
UAV-Kamera 2017	1690
UAV-Kamera 2018	2118
Gyro 2 cm	502
Gyro 5 cm	66
Standard 20 cm	25
Standard 50 cm	4

Tab. 3: Punktwolken-Dichte der verschiedenen Datensätze

Abbildung 3 zeigt beispielhaft die Punktwolken einer segmentierten Baumkrone. Es ist zu erkennen, dass bei höherer Punktdichte mehr Details innerhalb der Baumkrone ersichtlich sind. Die photogrammetrisch erzeugten Punktwolken beschreiben die Oberfläche der Baumkrone, während die LiDAR-Punktwolke die Struktur des Baumes von der Kronenspitze (hohe Punktdichte) bis zum Boden (niedrige Punktdichte) beschreibt.

2.3 Erfassungsmethoden für die Einzelbaumparameter Baumhöhe, Kronenradius und Kronenansatzhöhe mittels Feldmessungen

Die Baumhöhen der 30 Versuchsbäume wurden nach der Fällung am liegenden Baum mittels Maßband gemessen. Zur Bestimmung des Kronenradius wurden, wie in Röhle (1986) beschrieben, Kronenprojektionen am stehenden Baum erstellt. Dafür wurden acht Kronenrandpunkte in gleichmäßig verteilten Winkeln (von Norden aus in 45° Abstand) bestimmt und mit Fluchtstäben markiert. Die Entfernung der Fluchtstäbe von der Stammachse wurde trigonometrisch (ultraschallbasiert mittels Vertex) ermittelt. Die 8 Kronenradien wurden gemittelt. Die Kronenansatzhöhe wurde als die Höhe des ersten Grünastes definiert und am stehenden Baum mittels Vertex gemessen.



Abb. 3: Vergleich der Punktwolken einer automatisch segmentierten Baumkrone.

2.4 Erfassungsmethoden für die Einzelbaumparameter Baumhöhe, Kronenradius und Kronenansatzhöhe mittels Fernerkundung

Dieses Kapitel beschreibt die verwendeten Methoden zur Ableitung der Einzelbaumparameter mittels Fernerkundung. Für alle Datensätze wurden die Parameter aus dem Oberflächenmodell berechnet (rasterbasierter Ansatz). Für die LiDAR-Daten wurden zusätzlich Methoden entwickelt, um die Kronenparameter direkt aus den LiDAR-Punktwolken abzuleiten (LiDAR-Ansatz). Alle Berechnungsschritte wurden mithilfe der Programmiersprache R (R Core Team, 2017) durchgeführt.

2.4.1 Erstellung eines Oberflächenmodells und Einzelbaumdelinierung

Als Grundlage für die Entwicklung der bildanalytischen Verfahren dienten digitale Oberflächenmodelle (DSM), welche sowohl aus den photogrammetrisch erzeugten Punktwolken als auch aus LiDAR-Daten erzeugt wurden. In R wurde mittels *grid_canopy* aus dem Paket *lidR*, Version 1.5.0 (ROUSSEL et al. 2018), ein DSM erstellt. Diese Funktion gab für jedes Pixel den Z-Wert des am höchsten gelegenen Punktes wieder. Leere Pixel wurden mit einer Kombination aus dem *k-nearest-neighbor*-Ansatz und *Inverse Distance Weighting* interpoliert. Für die Daten mittels UAV und Gyrocopter wurde jeweils ein DSM mit 10 cm Auflösung, aus den Punktwolken des Standardbildfluges ein DSM mit 20 cm bzw. 50 cm Auflösung erstellt.

Um Kronenparameter abzuleiten sind einzelbaumbezogene Informationen notwendig. Folglich müssen alle Pixel eines Bildes bzw. alle Punkte einer Punktwolke, welche zu einer Baumkrone gehören, von den anderen Pixeln bzw. Punkten isoliert werden. Für die Segmentierung der Einzelbäume wurde der Watershed-Algorithmus mithilfe des Tools *lastrees* aus *lidR* angewandt. Segmentiert wurde auf Basis des Oberflächenmodells. Die Punktwolken der LiDAR-Daten wurden mithilfe der resultierenden Watershed-Segmente zu Einzelbäumen segmentiert.

2.4.2 Baumhöhe

Um die Baumhöhen zu ermitteln, mussten zunächst die Koordinaten der Baumspitzen identifiziert werden. Da bei der Segmentierung mittels Watershed nicht alle relevanten Bäume automatisch erkannt werden konnten, wurden manuell Polygone erstellt, welche die Baumspitzen aller Testbäume beinhalteten. Der höchste DSM-Wert innerhalb der Polygone wurde abgefragt und dessen Koordinaten ausgelesen. Um die Baumhöhe zu erhalten wurde die Differenz zwischen höchstem DSM-Wert und der entsprechenden Geländehöhe aus einem digitalen Geländemodell (DTM) berechnet. Als DTM wurde das landesweite DTM verwendet, welches vom LGL auf Basis von LiDAR-Daten mit einer Auflösung von 1 m erzeugt wurde (LGL 2018b). Für die Li-DAR-Punktwolken wurde die Baumhöhe zusätzlich anhand der mit den Bodenpunkten normalisierten Punktwolke bestimmt (höchster Punkt innerhalb der Krone), nachdem mit den Watershed-Segmenten auf einzelne Kronen zugeschnitten wurde.

2.4.3 Kronenradius

Die rasterbasierte Ableitung des Kronenradius konzentrierte sich auf den Vergleich des Verfahrens der Fernerkundung mit der herkömmlichen terrestrischen Messmethode zur Erstellung von Kronenprojektionen. Basierend auf der Arbeit von POPESCU et al. (2003) wurde ein Algorithmus entwickelt, welcher ausgehend von der Baumspitze die Kronenbreite eines Baumes automatisch erfasst. Für eine Baumkrone wurden vier Höhenprofile aus dem Oberflächenmodell extrahiert: Ost-West, Nord-Süd, Südwest-Nordost und Nordwest-Südost. Durch die Anpassung einer Kurve vierten Grades wurde die Kronenbreite modelliert. Die Minima der Kurve bildeten die Enden der Krone ab. Dies resultierte in acht Kronenradien, welche - simultan zu den terrestrischen Aufnahmen - von Norden aus mit 45° Abstand um die Krone verteilt waren. Die Höhenprofile wurden iterativ erstellt, angefangen mit einem Mindestradius von 3 m. Das Höhenprofil verlängerte sich schrittweise solange, bis an beiden Seiten durch eine Kurve vierten Grades ein lokales Minimum gefunden wurde und bestimmte Abbruch-Kriterien erfüllt waren oder bis ein Radius von 6 m erreicht wurde. Aus der LiDAR-Punktwolke wurde eine zweidimensionale von oben projizierte konvexe Hülle um die segmentierten Punkte gelegt. Der Radius eines Kreises mit der Fläche dieser konvexen Hülle wurde als mittlerer Kronenradius definiert. In Abbildung 4 sind die Methoden zur Ableitung des Kronenradius abgebildet.



Abb. 4: Ableitung des Kronenradius. Links: Rasierbasierter Ansatz. Höhenprofil nach den Höhenwerten eingefärbt von grün (niedrig) über gelb bis rot (hoch), resultierende Kronenprojektion in rot. Rechts: LiDAR-Ansatz. Berechnung des Kronenradius aus der zweidimensionalen konvexen Hülle.

2.4.4 Kronenansatzhöhe

Zur Bestimmung der Kronenansatzhöhe (rasterbasierter Ansatz) wurden normalisierte Oberflächenmodelle (nDSM) berechnet. Ein nDSM wurde auf detaillierte Kronenabgrenzungen zugeschnitten, welche eine Weiterentwicklung der achteckigen Kronenprojektionen aus 2.4.3 darstellten. Der niedrigste Pixelwert des zugeschnittenen nDSMs wurde als Kronenansatzhöhe definiert. Für die Bestimmung der Kronenansatzhöhe mittels LiDAR wurde eine Polynomialfunktion vierten Grades auf die vertikale Häufigkeit der Punkte gefittet. Die Höhe beim Minimum dieser Funktion war als Kronenansatzhöhe definiert (siehe Abbildung 5).



Abb. 5: Mittels LiDAR wurde die Kronenansatzhöhe mithilfe eines Histogramms berechnet. Die Position des Minimums der angepassten Kurve wurde als Kronenansatzhöhe definiert.

3 Ergebnisse

In diesem Kapitel werden die aus verschiedenen Fernerkundungsdaten abgeleiteten Kronenparameter mit den Feldmessungen verglichen und validiert. Für jeden Kronenparameter werden die Ergebnisse gesondert dargestellt. Die Genauigkeit der abgeleiteten Parameter wurde mit der Wurzel des mittleren quadratischen Fehlers (RMSE, Root Mean Square Error) und dem mittleren Fehler (ME, Mean Error) validiert. Der Variationskoeffizient über alle automatisch erkannten Bäume innerhalb der 6 Felder galt als Maß für die Reproduzierbarkeit (Präzision). Für die Baumhöhe wurden die Fernerkundungsdaten mit den Felddaten verglichen, welche im gleichen Zeitraum aufgenommen wurden. Dies ergab einen Vergleich von jeweils 15 Bäumen pro Datensatz. Für Kronenradius und Kronenansatzhöhe wurden jeweils alle 30 Bäume zur Validierung verwendet unter der Annahme, dass sich die Kronenausdehnung zwischen Juni 2017 und März 2018 nicht verändert hat. Tabelle 4 zeigt die Ergebnisse für alle Parameter und Datensätze. Die Reproduzierbarkeit der Kronenparameter wurde anhand aller auf der Versuchsfläche automatisch erkannten Bäume abgeleitet, indem der Variationskoeffizient über alle Datensätze einzelbaumweise berechnet und über alle Bäume gemittelt wurde.

Tab. 4: RMSE und ME für Baumhöhe, Kronenradius und Kronenansatzhöhe. Für die LiDAR-Daten wird das jeweils beste Ergebnis (Raster-/LiDAR-Ansatz) angezeigt.

	Baumhöhe (m)		Kronenradius (m)		Kronenansatzhöhe (m)	
Datensatz	RMSE	ME	RMSE	ME	RMSE	ME
UAV-LiDAR 2017	0,36	-0,13	0,53	0,11	4,46	1,00
UAV-LiDAR 2018	0,49	-0,3	0,46	-0,08	4,49	-1,83
UAV-Kamera 2017	1,09	1,03	0,45	0,12	4,52	1,08
UAV-Kamera 2018	0,49	0,18	0,53	0,06	5,18	1,23
Gyrocopter 2 cm	0,68	0,44	0,49	0,07	3,88	-0,21
Gyrocopter 5 cm	1,44	1,29	0,76	-0,11	8,81	-8,48
Standard 20 cm	1,86	1,63	0,68	0,06	9,49	-8,73
Standard 50 cm	2,89	2,71	0,67	0,09	11,26	-11,06

3.1 Baumhöhe

Wie in Abbildung 6 und Tabelle 4 dargestellt, liegen die berechneten Baumhöhen sehr nah an den Feldmessungen. Die Ergebnisse variieren zwischen einem RMSE von 0,36 m (RMSE%: 1,05) für UAV-LiDAR 2017 bis 2,89 m (RMSE%: 8,46) für Standard 50 cm. Vergleicht man bei den LiDAR-Daten den rasterbasierten mit dem LiDAR-Ansatz, so ergeben sich mit einem RMSE für den rasterbasierten Ansatz von 0,36 m bzw. 0,63 m (2017/2018) und 0,38 m bzw. 0,49 m (2017/2018) für den LiDAR-Ansatz nur geringe Unterschiede. Der Vergleich der verschiedenen Sensoren ergibt für UAV-LiDAR, gefolgt von UAV-Kamera 2018 und Gyrocopter 2 cm die besten Ergebnisse mit einem RMSE < 0,7 m. Die Genauigkeit für UAV-Kamera 2017, Gyrocopter 5 cm und die Standard-Luftbilder liegen mit einem ME > 1 m zu niedrig. Die Li-DAR-Daten liegen tendenziell eher höher als die Feldmessungen, während die photogrammetrischen Daten bis auf wenige Ausnahmen niedriger als die Feldmessungen liegen.



Abb. 6: Ergebnisse zur Ableitung der Baumhöhe aus verschiedenen Fernerkundungsdaten. Datensätze: st50/st20 = Standard 20/50 cm, gyro2/gyro5 = Gyrocopter 2/5cm, cam17/cam18: UAV-Kamera aus 2017/2018, lid17/lid18: UAV-LiDAR aus 2017/2018.

3.2 Kronenradius

Für den Kronenradius unterscheiden sich die Ergebnisse der Datensätze nur geringfügig (siehe Abbildung 7 und Tabelle 4). Der RMSE variiert zwischen 0,45 m (RMSE%: 11,88 m) und 0,76 m (RMSE%: 20,05). Der Kronenradius wird weder über- noch unterschätzt; der Fehler schwankt um die Werte der Feldmessungen. Die Unterschiede zwischen Raster-Ansatz und Li-DAR-Ansatz sind mit einem RMSE von 0,53 m/0,68 m (2017) und 0,46 m/0,71 m (2018) gering. Auch zwischen photogrammetrischen und LIDAR-Daten sind keine Unterschiede erkennbar.



Abb. 7: Ergebnisse zur Ableitung des Kronenradius aus verschiedenen Fernerkundungsdaten. Datensätze: st50/st20 = Standard 20/50 cm, gyro2/gyro5 = Gyrocopter 2/5cm, cam17/cam18: UAV-Kamera aus 2017/2018, lid17/lid18: UAV-LiDAR aus 2017/2018.

3.3 Kronenansatzhöhe

Für die Ableitung der Kronenansatzhöhe ergibt sich eine Genauigkeit von einem RMSE zwischen 3,88 m (RMSE%: 25,41) bei Gyrocopter 2 cm und 11,26 m (RMSE%: 73,74) bei Standard 50 cm. Für die Datensätze Standard und Gyrocopter 5 cm ist eine hohe systematische Überschätzung zu beobachten, während die anderen Datensätze im Rahmen von \pm 5 m (ohne Ausreißer) um die Feldmessung schwanken. Der LiDAR-Ansatz kann die Genauigkeit mit einem RMSE von 4,50 m/4,89 m (2017/2018) gegenüber dem Raster-Ansatz mit einem RMSE von 4,46 m/4,49 m (2017/2018) nicht verbessern. Die Ergebnisse sind in Abbildung 8 und Tabelle 4 dargestellt.



Abb. 8: Ergebnisse zur Ableitung der Kronenansatzhöhe aus verschiedenen Fernerkundungsdaten. Datensätze: st50/st20 = Standard 20/50 cm, gyro2/gyro5 = Gyrocopter 2/5cm, cam17/cam18: UAV-Kamera aus 2017/2018, lid17/lid18: UAV-LiDAR aus 2017/2018.

3.4 Reproduzierbarkeit

Der Variationskoeffizient (in %, \pm SD) liegt für die Ableitung von Baumhöhe, Kronenradius und Kronenansatzhöhe aus dem Oberflächenmodell bei jeweils 3,37 \pm 1,12, 7,07 \pm 2,93 und 34,74 \pm 7,7. Aus den mit LiDAR-Daten abgeleiteten Parametern ergibt sich ein Variationskoeffizient von jeweils 1,05 \pm 1,15, 4,83 \pm 5,75 und 11,58 \pm 10,23. Die Ergebnisse für die drei Kronenparameter werden in Abbildung 9 dargestellt. LiDAR-basiert ist die Variation geringer als mit rasterbasiertem Ansatz, wobei LiDAR-basiert die Variation von zwei LiDAR-Datensätzen analysiert wird, während mittels rasterbasiertem Ansatz 8 Datensätze verglichen wurden.



🛑 rasterbasiert 🗮 LiDAR-basiert

Abb. 9: Variationskoeffizient zur Beurteilung der Reproduzierbarkeit (Präzision).

4 Diskussion

Vergleichbare Studien erreichten für die Ableitung der Baumhöhe mittels LiDAR-Daten einen RMSE von 0,92 m (WALLACE et al. 2016) bzw. 1,2 m (SOLBERG et al. 2006). Mittels Luftbildern wurde ein RMSE von 0,35 m (ZARCO-TEJADA et al. 2014), 0,91 m (JENSEN & MATHEWS 2016), 1,30 m (WALLACE et al. 2016), 1,38 m (GIFTTHALER 2017) und 3.00 m - 3,08 m (PANAGIOTIDIS et al. 2016) erreicht. Aufgrund der Abweichungen, die bei der Bestimmung der Referenzwerte auftreten, ist es fraglich, ob im Feld eine genaue Reproduktion der terrestrischen Höhenmessungen überhaupt möglich ist. Auch die Ergebnisse zur Reproduzierbarkeit konventioneller Baumhöhenmessungen von LUOMA et al. (2017) stützen diese Zweifel. Bei Wiederholungsbefliegungen sind bei robusten Methoden generell ähnliche Ergebnisse zu erwarten. Zwischen der UAV-Befliegung mit Kamera im Jahr 2017 und 2018 ergaben sich bei der Baumhöhe jedoch signifikante Unterschiede. Grund hierfür war vermutlich die unzureichende Verteilung und Anzahl der Passpunkte bei der Befliegung 2017. Weil sich die Orientierung der UAV-Luftbilder von 2017 mit nur 4 Passpunkten schwierig gestaltete, wurden im Jahr 2018 12 über das Befliegungsareal verteilte Passpunkte eingesetzt. Aufgrund der zeitlichen Differenz zwischen der Befliegung des Standardbildfluges Anfang Juli 2016 und den Feldmessungen Ende Juni 2017 müssen die Ergebnisse der Standard-Luftbilder kritisch betrachtet werden. Die Aufnahme der Quirle an den gefällten Testbäumen ergab für die letzten 5 Jahre einen durchschnittlichen Quirl-Abstand von 0,5 m und maximal 1,4 m. Unter der Annahme, dass pro Jahr ein neuer Quirl entstanden ist, kann von einem jährlichen Höhenwachstum von etwa 50 cm und im Extremfall von bis zu 1,4 m ausgegangen werden. Ein mittlerer Fehler von bis zu 1,4 m ist bei den Ergebnissen der Standard-Luftbilder folglich zu tolerieren. Abzüglich dessen bleibt ein Fehler von 1 - 2 m.

POPESCU et al. (2003) erreichten für die Ableitung der durchschnittlichen Kronenbreite von dominanten Bäumen mittels LiDAR-Daten einen RMSE von 1,36 m – 1,41 m. Vergleichbare Studien erzielten für den durchschnittlichen Kronendurchmesser einen RMSE von 0,9 m (KATO et al. 2009), 0,82 m – 1,04 m (PANAGIOTIDIS et al. 2016), 1,10 m (SOLBERG et al. 2006), 1,8 m (POPSECU & ZHAO 2008) bzw. 1,8 m (SOLBERG et al. 2006). In diesen Studien wurde der RMSE für den Kronendurchmesser anstelle des mittleren Kronenradius angegeben. Daher sollte der RMSE dieser Studien halbiert werden, um vergleichbare Werte zu erhalten. In der vorliegenden Studie waren die Ergebnisse über alle Datensätze hinweg ähnlich. Dies wurde dadurch gefördert, dass die Parameter der rasterbasierten Methode bei der Erstellung der Kronenprojektionen für den jeweiligen Datensatz soweit optimiert wurden, um einen möglichst geringen RMSE zu erhalten. Folglich war der niedrigste RMSE 45 cm. Dies deckt sich in etwa mit der Genauigkeit, welche bei terrestrischen Aufnahmen erreicht werden kann.

Die Kronenansatzhöhe kann auf unterschiedlichste Weisen definiert sein (DEAN et al. 2009; JUNG et al. 2011; KATO et al. 2009; LUO et al. 2018; MUSS et al. 2011; POPESCU & ZHAO 2008; SOLBERG et al. 2006; VAUHKONEN 2008; WALLACE et al. 2014; XU et al. 2013). Jede dieser Studien verwendete hierbei eine andere Definition der Kronenansatzhöhe. Auch die Messfehler variierten von einem RMSE von 1,62 m in SOLBERG et al. (2006) bis zu 3,5 m in LUO et al. (2018). In der vorliegenden Studie konnte selbst der LiDAR-Datensatz mit einem RMSE von 4,5 m keine vergleichbaren Ergebnisse erzielen. Dies kann zum einen daran liegen, dass die Definition des ersten grünen Astes eine sehr schwierig zu reproduzierende Vergleichsgröße liefert. Es kann aber auch auf die Tatsache zurückzuführen sein, dass der Kronenansatz in seiner räumlichen Erscheinung von Baum zu Baum eine zu große Variabilität aufweist. So kann der Kronenansatz durch einen kleinen grünen Ast charakterisiert sein oder durch mehrere starke Äste. Zudem können Äste nach oben oder nach unten geneigt sein, was die Bestimmung der Kronenansatzhöhe noch weiter erschwert. Im Mittel konnten jedoch teilweise sehr gute Ergebnisse erzielt werden. Tatsächlich wichen die Mittelwerte aller 30 Versuchsbäume für die Datensätze UAV-LiDAR, UAV-Kamera und Gyrocopter 2 cm mit rasterbasiertem Ansatz maximal 2 m vom Mittelwert der Feldmessungen ab. Der mittlere Fehler dieser Datensätze lag zwischen -1,83 m und 1,23 m, während der mittlere Fehler für Gyrocopter 5 cm und Standard kleiner als -8,48 m war (siehe Tabelle 4). Dies weist darauf hin, dass der Kronenansatz auf Bestandesebene besser zu beschreiben ist. Auf Einzelbaumebene haben die Ergebnisse jedoch gezeigt, dass eine Bestimmung des Kronenansatzes mit den verfügbaren Daten stark fehlerbehaftet war.

Die Datensätze dieser Studie können auch durch eine unterschiedliche Punktdichte charakterisiert werden. YU et al. (2004) beschrieben für die Ableitung der Baumhöhe eine abnehmende Genauigkeit bei abnehmender LiDAR-Punktdichte bzw. zunehmender Flughöhe, wobei die Flughöhen 400 m, 800 m und 1500 m über Grund getestet wurden. Bei NÆSSET (2004) variierte hingegen nur 1 von 54 LiDAR-Höhenmetriken signifikant zwischen den Flughöhen von 530 m -540 m und 840 m - 850 m über Grund. Mit den Datensätzen dieser Studie konnte ab einer Punktdichte von 350 Punkten/m² eine höhere Genauigkeit erzielt werden (Vergleich Gyrocopter 2 cm und 5 cm). Für den Kronenradius hatten die Ausreißer bis zu einem Kronenradius von 4 m fast ausschließlich eine Punktdichte kleiner 350 Punkten/m². Da der RMSE bei allen Datensätzen jedoch ähnlich war, konnte auf keinen Zusammenhang zur Punktdichte geschlossen werden. Abbildung 10 zeigt, dass der Kronenansatz sich mit einer zu niedrigen Punktdichte nicht bestimmen lässt und die Abweichung zu den Feldmessungen ab einer Punktdichte von 502 Punkten/m² (Gyrocopter 5 cm) substanziell geringer ist. Bei niedrigeren Punktdichten werden kleine Lücken im Bestand bzw. Kronenränder vermutlich zu schlecht erfasst, sodass das DSM an diesen Stellen interpoliert wird und der untere Kronenbereich nicht ausreichend abgebildet wird. Dies führt zu einer systematischen Überschätzung der Kronenansatzhöhe.



Abb. 10: Streudiagramme zwischen den Feldmessungen und jeweils den abgeleiteten Parametern Baumhöhe, Kronenradius und Kronenansatzhöhe. Die Punktdichte der einzelnen Datensätze wird farbig von grün (= niedrige Punktdichte) bis rot (=hohe Punktdichte) markiert.

Es gab bereits eine Vielzahl an Untersuchungen zur Ableitung von Baumhöhe, Kronendurchmesser und -ansatzhöhe. Vergleichende Studien zielten zwar auf die gleiche Zielgröße, basierten jedoch auf einer unterschiedlichen Datengrundlage und Methodik, was einen Vergleich erschwert, mit einer großen Unsicherheit behaftet bis zu unmöglich macht. Beispielsweise variierten Baumarten, Waldstrukturen, Anzahl und Aufnahmemethodik der Referenzdaten, Bedingungen bei den Befliegungen, Sensorik, Plattform, Kameraeinstellungen, Flugplanung oder die Definitionen der Kronenparameter. Frühere Studien bewerteten meist einen Fernerkundungs-Datensatz (PANAGIOTIDIS et al. 2016; GIFTTHALER 2017, SIBONA et al. 2017), verglichen zwei Datensätze wie LiDAR- und photogrammetrischen Daten (WALLACE et al. 2016; THIEL & SCHMULLIUS 2016) oder untersuchten beispielhaft den Einfluss von zwei bis drei unterschiedlichen Flughöhen (YU et al. 2004; NÆSSET 2004) oder den Einfluss verschiedener Versuchsflächen (ZARCO-TEJADA et al. 2014; PANAGIOTIDIS et al. 2017). In diesem Beitrag wurden insgesamt 8 Datensätze miteinander verglichen, welche sowohl in der Sensorik (Kamera, LiDAR) als auch in der Aufnahmemethodik (Gyrocopter, UAV, Flugzeug, Flughöhe, Überlappung, Jahreszeit...) variierten, jedoch auf die gleichen Referenzdaten und die gleiche Methodik bei der Ableitung der Parameter zurückgriffen. Dadurch konnte mit einer höheren Sicherheit das Anwendungspotenzial der verschiedenen Sensoren und Plattformen aufgezeigt werden. Solche Vergleiche, wie in der vorliegenden Studie beschrieben, tragen dazu bei das Bewusstsein zu schärfen, dass verschiedene Metriken mehr oder weniger empfindlich auf unterschiedliche Datensätze reagieren können (ROUSSEL et al. 2017).

5 Fazit und Ausblick

Die Ableitung inventurrelevanter Einzelbaumcharakteristika ist mit Einschränkungen möglich. Allerdings setzt die erfolgreiche Anwendung eine korrekte Segmentierung der Einzelbäume voraus. Automatisiert ist dies leider immer noch mit gewissen Unsicherheiten verbunden. Die Genauigkeiten für die Bestimmung der Baumhöhe und des Kronenradius sind vielversprechend und die entwickelten Methoden ausreichend robust. Der Vergleich von automatisch abgeleiteter Baumhöhe bzw. Kronenradius und terrestrischer Messung fällt für alle Datensätze sehr positiv aus. Für die Ableitung der Baumhöhe sind LiDAR-Daten oder photogrammetrische Daten mit einer hohen Punktdichte und einer guten absoluten Orientierung optimal. Bei der Bestimmung von Kronenradius ist die Wahl des Sensors von geringerer Bedeutung. Die Höhe des Kronenansatzes auf Einzelbaumebene hingegen ist mit einer für die Rundholzsortierung relevanten Definition nicht automatisch bestimmbar. Selbst unter der Verwendung der LiDAR-Daten kann bei der Bestimmung des Kronenansatzes keine höhere Genauigkeit erzielt werden. Die Gründe für diese Ungenauigkeit und auch die im Vergleich zu anderen Untersuchungen entstehenden Wiedersprüche der Ergebnisse bei der Bestimmung des Kronenansatzes liegen vermutlich in der Definition der Kronenansatzhöhe. Bildet man jedoch den Mittelwert der Kronenansatzhöhen über alle Versuchsbäume, so kann für Datensätzen mit einer ausreichend hohen Punktdichte (UAV-LiDAR, UAV-Kamera, Gyrocopter 2 cm) auf Bestandesebene der Kronenansatz mit höherer Genauigkeit abgeleitet werden. Bei den photogrammetrischen Punktwolken können die besten Ergebnisse mit einer möglichst hohen Punktdichte erzielt werden. Allgemein kann man festhalten, dass die Unterschiede in der Genauigkeit zwischen LiDAR-Daten und photogrammetrischen Daten relativ gering ausfallen. Die Verwendung von Luftbildern wird daher als adäquate und kostengünstige Alternative für die Vermessung von einzelnen Kronen erachtet. WALLACE et al. (2016) stützen diese Bewertung durch eine Untersuchung von LiDAR- und photogrammetrischen Daten zur Analyse von Waldstrukturen. Die Analyse der Reproduzierbarkeit lassen die gleichen Schlussfolgerungen zu wie der Vergleich der Daten mit den Feldmessungen mittels RMSE: Die Baumhöhe lässt sich am zuverlässigsten ableiten. Der Kronenradius lässt sich mit einem akzeptabler Streuung reproduzieren, während die Ableitung des Kronenansatzes je nach Datensatz zu substanziell anderen Ergebnissen führt. Die Anwendung der untersuchten Verfahren kann dazu beitragen, Waldinventuren zu optimieren oder waldwachstumskundliche Versuchsflächen effizienter und standardisiert zu Vermessen.

6 Danksagung

Das in diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Ernährung und Landwirtschaft unter dem Förderkennzeichen 22023114 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren. Die Studie wurde an der Forstlichen Versuchs- und Forschungsanstalt Baden-Württemberg (FVA) durchgeführt. Die in dieser Arbeit verwendeten Datensätze (UAV-LiDAR, UAV-Kamera, Gyrocopter) sind im Rahmen Projektes "Weiterentwicklung statistischer Holzaufkommensprognoseverfahren zur Differenzierung von Rohholzsorten und Produktqualität (Pro-Qual-Tools)" entstanden. Die Luftbilder

des Standardbildflugs sowie das Geländemodell wurden vom LGL zur Verfügung gestellt (LGL 2018a; LGL 2018b).

7 Literaturverzeichnis

- DEAN, T., CAO, Q., ROBERTS, S. & EVANS, D., 2009: Measuring heights to crown base and crown median with LiDAR in a mature, even-aged loblolly pine stand. Forest Ecology and Management, 257(1), 126-133.
- JUNG, S., KWAK, D., PARK, T., LEE, W. & YOO, S., 2011: Estimating Crown Variables of Individual Trees Using Airborne and Terrestrial Laser Scanners. Remote Sensing, 3(11), 2346-2363.
- KATO, S., MOSKAL, L., SCHIESS, P., SWANSON, M., CALHOUN, D. & STUETZLE, W., 2009: Estimating Crown Variables of Individual Trees Using Airborne and Terrestrial Laser Scanners. Remote Sensing of Environment, 113(6), 1148-1162.
- LANDESAMT FÜR GEOINFORMATION UND LANDENTWICKLUNG BADEN-WÜRTTEMBERG, LGL, 2018a: Geobasisdaten © www.lgl-bw.de Az.: 2851.9–1/19.
- LANDESAMT FÜR GEOINFORMATION UND LANDENTWICKLUNG BADEN-WÜRTTEMBERG, LGL, 2018b: Digitale Geländemodelle (DGM). Online: https://www.lgl-bw.de/lglinternet/opencms/de/05_Geoinformation/Geotopographie/Digitale_Gelaendemodelle/ (Zugriff am 20 Dezember 2018).
- LUO, L., ZHAI, Q., SU, Y., MA, Q., KELLY, M. & GUO, Q., 2018: Simple method for direct crown base height estimation of individual conifer trees using airborne LiDAR data. Optics Express, 26(10), A562-A578.
- MARTENS, S., 2012: Erste Erfahrungen bei der Herleitung ertragskundlicher Parameter für Einzelbäume aus LIDAR-Daten. Deutscher Verband Forstlicher Forschungsanstalten, Sektion Ertragskunde, Jahrestagung 2012, 171-179.
- MUSS. J., MLADENOFF, D. & TOWNSEND, P., 2011: A pseudo-waveform technique to assess forest structure using discrete lidar data. Remote Sensing of Environment, **115**(3), 824-835.
- NÆSSET, E., 2004: Effects of different flying altitudes on biophysical stand properties estimated from canopy height and density measured with a small-footprint airborne scanning laser. Remote Sensing of Environment, **91**(2), 243-255.
- PRETZSCH, H., 2014: Canopy space filling and tree crown morphology in mixed-species stands compared with monocultures. Forest Ecology and Management, **327**, 251-264.
- PANAGIOTIDIS, D., ABDOLLAHNEJAD, A., SUROVÝ, P., & CHITECULO, V., 2017: Determining tree height and crown diameter from high-resolution UAV imagery. International journal of remote sensing, 38(8-10), 2392-2410.
- POPESCU, S. C., WYNNE, R. H. & NELSON, R. F., 2003: Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. Canadian journal of remote sensing, 29(5), 564-577.

- POPESCU, S. C., & ZHAO, K., 2008: A voxel-based lidar method for estimating crown base height for deciduous and pine trees. Remote Sensing of Environment, **112**(3), 767-781.
- R CORE TEAM, 2017: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- RÖHLE, H., 1986: Vergleichende Untersuchungen zur Ermittlung der Genauigkeit bei der Ablotung von Kronenradien mit dem Dachlot und durch senkrechtes Anvisieren des Kronenrandes (Hochblick-Messung). Forstarchiv, 57(2):67–71.
- ROUSSEL, J.-R., CASPERSEN, J., BÉLAND, M., THOMAS, S., ACHIM, A., 2017: Removing bias from LiDAR-based estimates of canopy height: Accounting for the effects of pulse density and footprint size. Remote Sensing of Environment, **198**, 1-16.
- ROUSSEL, J.-R. & AUTY D., 2018: lidR: Airborne LiDAR Data Manipulation and Visualization for Forestry Applications. R package version 1.5.0.
- SOLBERG, S., NAESSET, E. & BOLLANDSAS, O., 2006: Single Tree Segmentation Using Airborne Laser Scanner Data in a Structurally Heterogeneous Spruce Forest. Photogrammetric Engineering & Remote Sensing, 72(12), 1369-1378.
- THIEL, C. & SCHMULLIUS C., 2017: Comparison of UAV photograph-based and airborne lidarbased point clouds over forest from a forestry application perspective. International journal of remote sensing, **38**(8-10), 2411-2426.
- VAUHKONEN, J., 2008: Estimating crown base height for Scots pine by means of the 3D geometry of airborne laser scanning data. Publikation der University of Joensuu, Faculty of Forest Sciences, Hill R.A., Rosette, J. and Suárez, J. (Editor), 8th international conference on Li-DAR applications in forest assessment and inventory, 17.-19. September 2008 in Edinburgh, UK, 616-624.
- WALLACE, L., WATSON, C. & LUCIEER, A, 2017: Detecting pruning of individual stems using airborne laser scanning data captured from an Unmanned Aerial Vehicle. International Journal of Applied Earth Observation and Geoinformation, 30, 76-85.
- WEIDENBACH, M., WEZYK, P., TOMPALSKI, P., HOFFMANN, K. & MARTENS, S., 2012: Erfassung von Einzelbaumparametern mit Airborne-Laser-Scanning-Daten. AFZ-DerWald, 21, 12-15.
- XU, W., SU, Z., FENG, Z., XU, H., JIAO, Y. & YAN, F., 2012: Comparison of conventional measurement and LiDAR-based measurement for crown structures. Computers and Electronics in Agriculture, 98, 242-251.
- YU, X., HYYPPÄ, J., HYYPPÄ, H., & MALTAMO, M., 2004: Effects of flight altitude on tree height estimation using airborne laser scanning. Proceedings of the Laser Scanners for Forest and Landscape Assessment–Instruments, Processing Methods and Applications, 02-06.
- ZARCO-TEJADA, P. J., DIAZ-VARELA, R., ANGILERI, V., & LOUDJANI, P., 2014: Tree height quantification using very high resolution imagery acquired from an unmanned aerial vehicle (UAV) and automatic 3D photo-reconstruction methods. European journal of agronomy, 55, 89-99.

Deep Learning-basierte Baumartenklassifizierung auf Basis von ALS-Daten

SEAD MUSTAFIC¹ & MATHIAS SCHARDT¹

Zusammenfassung: In diesem Beitrag werden Methoden zur Baumartentrennung (Fichte, Kiefer und Laubbäume) auf Basis von ASL-Daten und den darin enthaltenen Strukturmerkmalen mit dem Deep Learning (DL) Ansatz unter Verwendung von Convolutional Neural Network (CNN) vorgestellt. Dazu wurden bereits bestehende Modelle, die auf andere Aufgabenstellungen optimiert und trainiert wurden, mittels Transfer Learning (TL) an die gegenständliche Themenstellung angepasst. Darüber hinaus wurde eine eigene Netzarchitektur entwickelt und mehrere CNN-Modelle (für 1D- bzw. 2D-Inputbilder) von Grund auf neu trainiert. Die Evaluierung der Ergebnisse hat gezeigt, dass in beiden Fällen (TL bzw. eigene Netzarchitektur) eine mittlere Klassifizierungsgenauigkeit von 74 % erreicht werden kann.

1 Motivation und Stand der Technik

In der jüngsten Zeit hat sich Deep Learning (DL) zu einer Kerntechnologie im Bereich des Machine Learning (ML) entwickelt und in verschiedenen Anwendungsbereichen bereits eine große Popularität erreicht. Ob ein traditioneller Machine Learning- oder Deep Learning- Ansatz (LIU & COCEA 2018; NILSSON 1998) gewählt wird, hängt von vielen Faktoren ab, wie z. B. der Komplexität des Problems, von der Datenmenge, aber zu einem nicht geringen Teil auch von den zur Verfügung stehenden Rechenressourcen. So ist der Einsatz von ML für die Lösung von "einfacheren" Problemen oft eine bessere Wahl, insbesondere wenn die Datenmenge, mit der das Modell trainiert werden soll, klein ist. Mit DL hingegen können komplexere Aufgaben gelöst werden, da das Artificial Neural Network (ANN) direkt aus den Trainingsdaten selbstständig bestimmte Muster und Eigenschaften, die in den Daten vorkommen, selektiert und daraus lernt. Dies erfordert jedoch eine weitaus höhere Menge an Trainingsdaten und in den meisten Fällen auch die Verfügbarkeit von mehr Rechenressourcen. Stehen nicht genügend Daten und Rechenressourcen zum Training zur Verfügung, kann ein sogenanntes Transfer Learning (TL) angewendet werden. Dabei wird ein Modell, das bereits mit riesigen Datenmengen von Grund auf für eine bestimmte Aufgabe trainiert wurde, herangezogen und nur bestimmte Bereiche des ANNs werden auf die neue Aufgabe mit weniger Trainingsdaten und Rechenressourcen optimiert (retraining) (CHEN et al. 2018; HUANG et al. 2017; ZHAO et al. 2017; TORREY & SHAVLIK 2010). Aufgrund der großen Erfolge in anderen Anwendungsfeldern (Computer Vision, Robotik, etc.) fanden die Deep Learning Ansätze auch Einzug in verschiedene Bereiche der Fernerkundung. Dies gilt vor allem für die Lösung komplexerer Fragestellungen der Fernerkundung, die mit den traditionellen Klassifizierungsmethoden bisher nur schlecht oder gar nicht gelöst werden konnten.

¹ Technische Universität Graz, Institut für Geodäsie, Arbeitsgruppe Fernerkundung und Photogrammetrie, Steyrergasse 30, A-8010 Graz, E-Mail: [Mustafic, Mathias.Schardt]@tugraz.at

Ein Beispiel hierfür ist die Ableitung von forstlichen Parametern wie Baumart oder die Segmentierung einzelner Baumkronen aus ALS-Daten.

Für eine Inventur in der Forstwirtschaft werden heutzutage verschiedene Fernerkundungsdaten (Satelliten- bzw. Luftbilder oder ALS (Airborne Laser Scanning) Daten) herangezogen und miteinander kombiniert. Dabei werden in der Regel aus Luftbildern die Baumart erfasst und aus ALS-Daten die Baumhöhe abgeleitet oder die Kronen segmentiert (DECHENSE et al. 2016; ABBASI 2015; EYSN et al. 2015; MUSTAFIC et al. 2014; BUDDENBAUM 2010; Heinzel et al. 2008). Nur in wenigen Studien wurde eine Baumartentrennung allein anhand der ALS-Daten (mit einer relativ geringen Punkdichte) bzw. der daraus gewonnenen *features* auf Einzelbaumbasis durchgeführt. ØRKA et al. (2007) haben für die Trennung verschiedener Baumarten die ASL-Intensitätswerte verwendet und dabei eine Genauigkeit von 74 % erreicht. In einer weiteren Studie konnte anhand der Strukturmerkmale einer Baumkrone eine Klassifizierungsgenauigkeit von 77 % erreicht werden (ØRKA et al. 2009). Durch die Kombination von Intensität und Strukturmerkmalen konnte die Klassifizierungsgenauigkeit sogar auf ca. 90 % gesteigert werden (REITBERGER et al. 2010; ØRKA et al. 2009).

In der Untersuchung wurde die Tauglichkeit verschiedener *Deep Learning* Techniken für die Baumartenklassifizierung auf der Basis von ALS-Daten, die eine relativ geringe Punktdichte aufweisen, analysiert und miteinander verglichen.

2 Daten und Untersuchungsgebiet

Das Untersuchungsgebiet befindet sich in der Oststeiermark (Österreich) in der Nähe von Burgau und liegt auf eine Meereshöhe zwischen 260 und 370 Meter. Das Gebiet ist ca. 2 km breit und 9 km lang und ist die Heimat von verschiedenen Baumarten. Die meistvorkommenden Baumarten sind Fichte, Kiefer, Lerche, Buche, Eiche, Esche und Erle.

In der Vergangenheit wurde das Gebiet von Joanneum Research sehr oft als Untersuchungsgebiet herangezogen, woraus eine Vielzahl an verschiedenen Datensätzen (Satelliten- und Luftbilder, ALS-Daten sowie auch Felddaten) entstanden ist. Für diese Studie besonders von Bedeutung ist eine Zeitreihe von ALS-Daten, die in einer Zeitspanne von ca. 17 Jahren aufgenommen wurden. Die letzte Befliegung fand im Jahr 2016 im Sommer (im belaubten Zustand) statt. Die aus dieser Befliegung gewonnenen Daten wurden für die Untersuchung herangezogen. Die Aufnahme wurde aus einer Höhe von ca. 600 m über Grund durchgeführt, woraus eine mittlere Punktdichte von ca. 15 Punkte/m² resultiert. Es wurde das erste, mittlere und letzte Echo eines Laserstrahls, die für diese Untersuchung gleichwertig betrachtet wurden, aufgenommen.

3 Methode

Die hier vorgestellte auf ALS-Daten basierte Baumartenklassifizierung wurde auf der Basis von Einzelbäumen vorgenommen. Grundlage für die Baumartenklassifizierung ist eine präzise Vorverarbeitung der Daten sowie die Überführung der ALS-Punktwolkendaten in eine bestimmte Rasterform. Die näheren Schritte dazu sind in den folgenden Kapiteln genauer beschrieben.

3.1 Datenvorverarbeitung

Im ersten Vorverarbeitungsschritt wurden die ALS-Punktwolken auf Ausreißer und Datenlücken überprüft und diese bereinigt. Anschließend folgte eine Einzelbaumdetektion (Lokale Maxima), bei der die Spitzen der einzelnen Bäume detektiert und mit den Eigenschaften Position und Höhe *gelabelt* wurden (MUSTAFIC et al. 2014; REITBERGER et al. 2010; WACK et al. 2003). Die auf diese Weise gewonnene Baumspitze stellt einen Baumkandidaten dar und wird im weiteren Verlauf als Mittelpunkt eines Baumes betrachtet.

In Rahmen dieser Untersuchung wurden (Deep Learning) Ansätze, die auf Rasterdaten basieren, gewählt. Ein Grund dafür ist einerseits die einfache Handhabung der Daten in einem regelmäßigen Gitter, andererseits wird die Rechenzeit dadurch stark reduziert und die Berechnung der ANNs auf einem Standard Computer möglich. In diesem Fall wurde die Information, nämlich die (x, y, z) – Koordinaten eines Baumes, in mehrere verschiedene Rasterbilder transformiert, die möglichst alle Baumes beinhalten sollen. wichtigen Eigenschaften eines Da bewusst auf eine Baumkronensegmentierung verzichtet wurde, um durch die Segmentierungsfehler nicht die Klassifizierung zu beeinflussen, wurden mehrere Radien (R=[1, 1.5, 2, 3] Meter) um eine Baumspitze festgelegt. Diese Radien sollen die Baumkronenumrisse darstellen, woraus dann das ANN selbstständig herausfinden soll, in welchem Bereich (Radien in Bezug auf die Baumspitze) sich die entscheidenden Informationen für die Separierung der Baumarten befinden. Auf diese Weise sind für jedes Radiussegment verschiedene Rasterbilder entstanden, die einen Baum bzw. Teile eines Baums repräsentieren. Beschreibende Merkmale in diesem Zusammenhang sind die im Folgenden näher beschriebenen Merkmale: Histogramme, Perzentile und Profile (Abb. 1).



Abb. 1: Vertikale Transformation der Baumstruktur/Punktverteilung in ein 1D-Rasterbild (links) und Profildarstellung einer Baumkrone in einem 2D-Farbbild (rechts). Die Farbe (im HSV Farbraum) wurde in Abhängigkeit von der Ausrichtung des Punktes in Bezug auf die Baumspitze vergeben. Dadurch kommen bestimmte Strukturen in der Baumkrone zum Vorschein.
Histogramme (1D) sind Bilder, die nur aus einer Pixelreihe (Dimension z. B. 1x20 Pixel) bestehen (Abb. 1 links). Diese sollen die vertikale Struktur eines Baumes, die in der Punktwolke enthalten ist, widerspiegeln. Dabei wird jeder Baum in n Höhenbereiche (z. B. in 20 Höhenbereiche) aufgeteilt und anschließend für jeden Bereich die Anzahl der darin enthaltenen Punkte (x, y, z) ermittelt.

Perzentile (1D) sind von der Struktur her ähnlich wie die Histogramme aufgebaut, aber mit dem Unterschied, dass für jeden Höhenbereich anstatt der Punkteanzahl das Perzentil berechnet wird.

Profile (2D) sollen gleichzeitig die vertikalen sowie die horizontalen Eigenschaften einer Baumkrone repräsentieren (Abb. 1 rechts). Unter dem Profilbild ist die Generierung von Profilsichten einer Baumkrone zu verstehen. Dabei werden die 3D-Punkte einer Baumkrone in einer Profilsicht durch Symbole (kleine Kreise) farblich dargestellt. Die Farbe (im HSV-Farbraum) wird in Abhängigkeit von der Ausrichtung des Punktes in Bezug auf die Baumspitze vergeben. Durch die Farbvergabe sollen bestimmte Verzweigungen und Verästelungsformen in der Baumkrone beschrieben werden (siehe Abb. 1 rechts), woraus das Modell bestimmte geometrische Merkmale für die Klassifizierung herausfinden soll. Die auf diese Weise aufbereitete Profilsicht eines Baumes wurde dann als ein Farb-Rasterbild mit der Auflösung von 100 x 100 Pixel abgespeichert.

Darüber hinaus wurden eine Reihe weiterer Transformationen (Standardabweichung, Vertikale Punktverteilung als 2D-Rasterbild abgebildet, Rauigkeit von Baumkrone, etc.) der 3D Baumpunktwolken in 2D-Rasterbilder getestet, allerdings konnte damit kein zufriedenstellendes Klassifizierungsergebnis erzielt werden. Von einer näheren Beschreibung dieser Transformationsverfahren wird daher abgesehen.

Die Erhebung der Referenzdaten (*ground truth*) wurde manuell am Bildschirm anhand von hoch aufgelösten Infrarot- und RGB Orthophotos mit einer Bodenauflösung von 50 cm bzw. 9 cm durchgeführt. In weiterer Folge wurde jeder Baumspitze die entsprechende Baumart zugewiesen. Damit konnte bereits bei der Ableitung der Merkmale (Generierung von Rasterbildern) jedem Radiussegment die Klasse zugewiesen werden.

3.1.1 Datensätze für Training und Validierung

Für das Trainieren des CNN-Modells und die Qualitätskontrolle wurden die erhobenen Referenzdaten im Vorfeld in Trainings- und Validierungsdaten aufgeteilt. Im Rahmen dieser Untersuchung wurde eine Klassifizierung in drei verschiedene Baumarten bzw. Baumartenklassen (Fichte, Kiefer und Laubbäume) durchgeführt, wobei sich die Klasse "Laubbäume" aus den Baumarten Buche, Birke, Eiche, Esche und Erle zusammensetzt.

Für das Trainieren der Modelle wurden für jede Klasse jeweils 500 möglichst repräsentative Bäume (Abb. 2 links) ausgewählt. Der Datensatz für die Validierung (Abb. 2 rechts) setzt sich aus drei Validierungsdatensätzen zusammen. Jeder dieser drei Datensätze hat leicht unterschiedliche Eigenschaften hinsichtlich der Altersstruktur, Baumdichte, Mischverhältnisse sowie Laubbaumkronenform.

Mit der Stratifizierung des Validierungsdatensatzes soll untersucht werden, ob bzw. wie stark der Klassifizierungserfolg von den vorliegenden Waldstrukturen beeinflusst wird und inwieweit sich die in einem Stratum generierten Modelle auf andere Waldstrukturen übertragen lassen. Der erste

Validierungsdatensatz besteht aus jeweils 70 Bäumen pro Klasse, die aus demselben Waldbereich stammen, wie die Trainingsdaten selbst. Diese wurden aber zum Trainieren nicht herangezogen. Der zweite Validierungsdatensatz besteht aus jeweils 72 Bäumen pro Klasse, die in einem benachbarten Waldbereich aufgenommen wurden, aber teilweise ähnliche Eigenschaften (Baumdichte und Mischverhältnis) wie die Trainingsdaten aufweisen. Der Unterschied zum ersten Validierungsdatensatz besteht darin, dass die Fichten- und Kieferbestände eine heterogenere Altersstruktur aufweisen. Der dritte Validierungsdatensatz besteht aus jeweils 128 Bäumen pro Klasse, die sich hinsichtlich Baumdichte, Mischverhältnisse und insbesondere der Laubbaumkronenform (hoher Birkenanteil) stark von den Trainingsdaten und den anderen beiden Validierungsdatensätzen unterscheiden.



Abb. 2: Trainingsdatensatz (links) und Validierungsdatensatz (rechts) für die Klassen Fichte, Kiefer und Laubbäume. Der Validierungsdatensatz setzt sich aus drei einzelnen Datensätzen zusammen.

Wie aus Abbildung 2 ersichtlich, wurden nur Bäume aus der mittleren- und oberen Waldschicht für die Klassifizierung und Validierung herangezogen, da die Bäume aus der unteren Schicht in den Daten nicht sichtbar sind und damit eine erfolgreiche Klassifizierung nicht möglich ist.

3.2 Klassifikation mit CNNs

Für die Klassifikation der Baumarten wurden Convolutional Neural Networks (CNNs) verwendet. Dabei wurden zwei verschiedene Ansätze verfolgt.

Im ersten Ansatz wurde ein Transfer Learning und bereits bestehende *pre-trainierte* Modelle, die auf ImageNet Database (DENG et al. 2009) trainiert wurden, verwendet und entsprechend der Aufgabenstellung nur die letzten Layer des CNNs *re-trainiert*. Dazu wurde InceptionV3 und VGG16 Netzarchitektur (SZEGEDY et al. 2015; SIMONYAN & ZISSERMAN 2015) mit ihren Gewichtungen (für Image Database) verwendet.

Im zweiten Ansatz wurden eigene CNNs (1D und 2D) erstellt und von Grund auf neu trainiert. Die Netzarchitektur ist in der Abbildung 3 dargestellt. Es wurde die gleiche Netzarchitektur für 1D-CNN und 2D-CNN verwendet, nur jeweils die Parameter (BENGIO 2012) entsprechend angepasst.



Abb. 3: Netzarchitektur von 1D- bzw. 2D CNN. Vom Aufbau her unterscheidet sich die Netzarchitektur zwischen den 1D- bzw. 2D CNN nicht, es wurden nur die Parameter für *Convolution* und *Max-Pooling* entsprechend angepasst.

Wie in der Abbildung 3 zu erkennen ist, kommen Mehrfach-Dropouts in der Netzarchitektur vor. Dadurch konnte verhindert werden, dass es in der Trainingsphase, in der nur ein relativ kleiner Datensatz (jeweils nur 500 Bäume pro Klasse) zum Trainieren zur Verfügung stand, zu einer Überbestimmung des Modells kommt (Abb. 4). Im Fall von Profilbildern war es möglich, durch *Image Augmentation* (WONG et al. 2016) mehrere Profilbilder pro Baum zu generieren, wodurch die Gesamtanzahl an Bildern für das Training signifikant erhöht werden konnte. Dabei wurde die Punktwolke eines Baumes mehrmals um 45° (in Bezug auf die Baumspitze) rotiert und in jeder Position ein Profilbild generiert.



Abb. 4: Überbestimmtes- (links), Optimal- (Mitte) und Unterbestimmtes- (rechts) Modell.

Nachdem die Punkte (x, y, z) einer Baumkrone in 2D-Rasterbilder (Histogramme, Profile) transformiert wurden (siehe Kapitel 3.1) und für jede Art von Transformation mehrere Radien (R=[1, 1.5, 2, 3] Meter) um die Baumspitze verwendet wurden, wurde anschließend für jeden Radius jeweils ein CNN Modell trainiert. Die Grundidee dabei ist, ein *Ensemble Learning* (DIETTERICH 2000) mit unterschiedlichen Modellen aufzubauen (Abb. 5), wobei die Klasse/Typ eines Baumes mehrfach durch die Radiussegmente repräsentiert und prädiziert wird. Danach werden die Wahrscheinlichkeiten, die aus den einzelnen Modellen resultieren, zusammengeführt und aus der Kombination der Einzelwahrscheinlichkeiten endgültig die Baumart prädiziert.

Dreiländertagung der DGPF, der OVG und der SGPF in Wien, Österreich – Publikationen der DGPF, Band 28, 2019



Abb. 5: Umwandlung einer 3D-Punktwolke einer Baumkrone in Rasterbilder und anschließende Baumartbestimmung durch Kombination von mehreren CNN-Modellen.

4 Ergebnisse und Diskussion

Die Ergebnisse, die mit den verschiedenen CNN-Modellen erreicht wurden, sind der Abbildung 6 zu entnehmen. Die Tests haben gezeigt, dass in den meisten Fällen durch das Zusammenführen der Einzelergebnisse verschiedener Inputbilder (Histogramme und Profile) ein leicht besseres Gesamtergebnis erzielt werden konnte. Die erreichte mittlere Klassifizierungsgenauigkeit, die sich aus allen drei Validierungsdatensätzen zusammensetzt, bewegte sich im Bereich zwischen 67 % und 74 %.

In Bezug auf die Baumart sticht die Fichte mit der höchsten erreichten Klassifizierungsgenauigkeit von bis zu 90 % besonders heraus. Die Kiefer und die Laubbäume konnten (im besten Fall) mit einer Genauigkeit von bis zu 78 % bzw. 67 % klassifiziert werden. Die deutlich geringere Trennbarkeit von Kiefern und Laubbäumen ist wahrscheinlich auf die sehr ähnliche (elliptische) Kronenform zurückzuführen, wohingegen die Kronenform einer Fichte (kegelförmig) sich deutlich von den anderen Baumarten unterscheidet.

Beim Performance-Vergleich der Inputdaten, die für die CNN-Klassifizierung verwendet wurden, fällt auf, dass die höchste Klassifizierungsgenauigkeit bei den Profilbildern erreicht werden konnte. Das lässt sich wahrscheinlich dadurch erklären, dass im Vergleich zu den Histogrammen nicht nur eine einfache vertikale, sondern auch die horizontale Verteilung der Punkte berücksichtigt wurde, sodass zu den einzelnen Baumhöhenschichten auch weitere Strukturmerkmale eingebunden wurden (z. B. einzelne Äste in der Baumkrone), die anscheinend eine wichtige Rolle bei der Trennung von Baumarten spielten.

Ein Vergleich der Genauigkeiten für alle drei Validierungsdatensätze (Abb. 6) zeigt, dass alle Modelle beim dritten Validierungsdatensatz schlechter abgeschnitten haben. Dieses Ergebnis wurde auch erwartet, da sich die Baumdichte sowie die Baumkronenform insbesondere bei der Klasse "Laubbäume" deutlich von den Trainingsdaten unterschieden haben. Beim Vergleich der Einzelergebnisse, die für die Klasse "Laubbäume" erzielt werden konnten, fällt auf, dass dieser



Effekt noch deutlicher zu erkennen ist. Hier ist die Genauigkeit bei der Klasse "Laubbäume" im dritten Validierungsdatensatz im Durchschnitt sogar um 50 % geringer.

Abb. 6: Klassifizierungsergebnisse der einzelnen Modelle für drei verschiedene Validierungsdatensätze sowie deren Mittelwerte.

Obwohl die 1D- und 2D-CNN Modelle verglichen mit den InceptionV3- und VGG16-Modellen mit einem relativ kleinen Datensatz trainiert wurden, konnte hier dennoch die höchste Klassifizierungsgenauigkeit erzielt werden. Das hängt mit der Tatsache zusammen, dass die eigenen Modelle optimal auf die Aufgabenstellung und den vorliegenden Datentyp angepasst sind. InceptionV3- und VGG16 Modelle wurden auf andere Datentypen hin trainiert und ein *re-training* der letzten Layer scheint nicht ausreichend zu sein.

Noch zu anmerken ist, dass die hier vorgestellte Netzarchitektur (1D- bzw. 2D CNN) mit einem einfachen Büro-PC von Grund auf mit nur 500 Datensätzen pro Klasse trainiert werden konnte. Die unerwartet hohe Klassifizierungsgenauigkeit ist damit zu begründen, dass die verwendeten Inputdaten bezüglich der Skalierung, des Blickwinkels und der Helligkeitsunterschiede homogen sind und dadurch ein erfolgreiches Trainieren bereits mit einem kleineren Trainingsdatensatz möglich ist.

Dreiländertagung der DGPF, der OVG und der SGPF in Wien, Österreich – Publikationen der DGPF, Band 28, 2019

5 Fazit & Ausblick

Im Rahmen dieser Studie wurde gezeigt, dass eine Baumartenklassifizierung anhand von ALS-Daten (und den nur darin enthaltenen Strukturmerkmalen) mit CNN erfolgreich durchgeführt werden kann. Durch die Kombination von mehreren CNN Modellen war es möglich, die Genauigkeit leicht zu verbessern. Darüber hinaus wurde gezeigt, dass bereits mit einer relativ einfachen CNN Architektur, die auf einem Büro-PC relativ schnell und einfach trainiert werden kann, zufriedenstellende Ergebnisse erzielt werden können. Die Untersuchung hat gezeigt, dass bei der Trennung der Baumarten bzw. Waldtypen Fichte, Kiefer und Laubbäume eine mittlere Klassifizierungsgenauigkeit von 74 % erzielt werden kann. Es hat sich weiterhin gezeigt, dass die erzielbaren Genauigkeiten sehr stark von der jeweiligen Baumart abhängen. So kann die Baumart Fichte mit der signifikant höchsten Genauigkeit klassifiziert werden.

Die erreichbaren Genauigkeiten für die Klasse "Laubwälder", wären wahrscheinlich bei ALS-Daten, die im Winter aufgenommen wurden, höher als bei den für diese Studie verwendeten Daten, die im Sommer in der belaubten Vegetationsphase aufgenommen wurden. Inwieweit durch die Verwendung von multi-temporalen Datensätzen (Sommer, leaf-on und Winter, leaf-off) der Klassifizierungserfolg erhöht werden kann, wird in weiteren Studien untersucht.

6 Literaturverzeichnis

- ABBASI, B., AREFI, H., BIGDELI, B., MOTAGH, M. & ROESSNER, S., 2015: Fusion of hyperspectral and lidar data based on dimension reduction and maximum likelihood. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 40(7/W3), 569-573.
- BUDDENBAUM, H., 2010: Charakterisierung von Forstbeständen mit Hilfe von Laserscanning und Reflexionsmodellierung. PhD Thesis, Universität Trier, 56-89.
- CHEN, Z., ZHANG, T. & OUYANG, C., 2018: End-to-End Airplane Detection Using Transfer Learning in Remote Sensing Images. Remote Sensing, **10**(1), Beitrag 139.
- DECHESNE, C., MALLET, C., LE BRIS, A., GOUET, V. & HERVIEU, A., 2016: Forest Stand Segmentation Using Airborne LIDAR Data and Very High Resolution Multispectral Imagery. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **41**(B3), 207-214.
- DENG, J., DONG, W., SOCHER, R., LI, L., LI, K. & FEI-FEI, L., 2009: ImageNet: A large-scale hierarchical image database. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops), Miami, FL, 248-255.
- DIETTERICH, T.G., 2000: Ensemble Methods in Machine Learning. Multiple Classifier Systems, Springer Berlin Heidelberg, Berlin, Heidelberg, 1-15.
- EYSN, L., HOLLAUS, M., LINDBERG, E., BERGER, F., MONNET, J.M., DALPONTE, M., KOBAL, M., PELLEGRINI, M., LINGUA, E., MONGUS, D. & PFEIFER, N., 2015: A Benchmark of Lidar-Based Single Tree Detection Methods Using Heterogeneous Forest Data from the Alpine Space. Forests, 6(12), 1721-1747.

- HEINZEL, J.N, WEINACKER, H. & KOCH, B., 2008: Full automatic detection of tree species based on delineated single tree crowns-a data fusion approach for airborne laser scanning data and aerial photographs. SilviLaser 2008, Sept. 17-19, Edinburgh, UK, 76-85.
- HUANG, Z., PAN, Z. & LEI, B., 2017: Transfer Learning with Deep Convolutional Neural Network for SAR Target Classification with Limited Labeled Data. Remote Sensing, **9**(9), Beitrag 907.
- LIU, H. & COCEA, M., 2018: Granular Computing Based Machine Learning: A Big Data Processing Approach. Cham: Springer International Publishing, **35**, 11-22.
- MUSTAFIC, S., KAINER, A. & SCHARDT, M., 2014: Einzelbaumdetektion anhand von Ebenenschnitten. Angewandte Geoinformatik 2014, Beiträge zum 26. AGIT-Symposium, Wichmann Verlag, 21-26.
- NILSSON, N.J., 1998: Introduction to Machine Learning: An Early Draft of a Proposed Textbook. eBook, 35-87.
- ØRKA, H.O., NÆSSET, E. & BOLLANDSÅS, O.M., 2007: Utilizing Airborne Laser Intensity for Tree Species Classification. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(3/W52), 300-304.
- ØRKA, H.O., NÆSSET, E. & BOLLANDSÅS, O.M. 2009: Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. Remote Sensing of Environment, **113**(6), 1163-1174.
- REITBERGER, J., 2010: 3D-Segmentierung von Einzelbäumen und Baumartenklassifikation aus Daten flugzeuggetragener Full Waveform Laserscanner. PhD Thesis, Technische Universität München, 111 S.
- SZEGEDY, C., VANHOUCKE, V., IOFFE, S., SHLENS, J. & WOJNA, Z., 2015: Rethinking the Inception Architecture for Computer Vision. Arxiv, http://arxiv.org/pdf/1512.00567v3.
- SIMONYAN, K. & ZISSERMAN, A., 2014: Very Deep Convolutional Networks for Large-Scale Image Recognition. CoRR, abs/1409.1556, Arxiv, https://arxiv.org/pdf/1409.1556.pdf.
- TORREY, L. & JUDE S., 2010: Transfer Learning. Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques. IGI Global, 242-264.
- WACK, R., SCHARDT, M., BARRUCHO, L., LOHR, U. & OLIVEIRA, T., 2003: Forest inventory for eucalyptus plantations based on airborne laserscanner data. Proceedings of the ISPRS working group III/3 workshop "3-D Reconstruction from Airborne Laserscanner and InSAR Data", 34, 3W13.
- WONG, C.S., GATT, A., STAMATESCU, V. & MCDONNELL, D.M., 2016: Understanding Data Augmentation for Classification: When to Warp? International Conference on Digital Image Computing: Techniques and Applications (DICTA), 1-6.
- YOSHUA, B., 2012: Practical recommendations for gradient-based training of deep architectures. Arxiv, https://arxiv.org/abs/1206.5533, 7-21.
- ZHAO, B., HUANG, B. & ZHONG, Y., 2017: Transfer Learning With Fully Pretrained Deep Convolution Networks for Land-Use Classification. IEEE Geosci. Remote Sensing Lett., 14(9), 1436-1440.

Evaluierung der Airborne LiDAR Bathymetrie in der Fließgewässer-Modellierung

KATHARINA BAUMGARTNER¹, ROBERT KLAR¹ & MARKUS AUFLEGER¹

Zusammenfassung: Die Airborne LiDAR Bathymetrie (ALB) verfügt über das immense Potential, selbst kleinräumige morphologische Strukturen flächig erfassen zu können. Eine durchaus offene Fragestellung betrifft die erreichbare Genauigkeit und die ALB Anwendbarkeit für die hydronumerische Simulation. Konsequenterweise befasst sich der vorliegende Beitrag mit einer umfassenden Evaluierung bestehend aus Datenkontrolle u.a. durch begleitende Feldmessungen und der Untersuchungen der Auswirkungen unterschiedlicher Refraktionierungskorrekturen, Netztopologien, -auflösung bzw. Rauheitszonenverteilungen. Als Anwendungsgebiet dient ein 1,5 km langer Gewässerabschnitt am Mareiterbach (Italien), der im Zuge des FHARMOR Projekts zweimal mit einem Riegl VQ-880-G Sensor vermessen wurde. Wesentliche Ergebnisse: (1) Die Genauigkeit der ALB Höhe liegt im Mittel innerhalb der GPS-Genauigkeit (<2,5 cm) und (2) bei hoher Netzauflösung lässt sich ein hervorragendes Kalibrierungsergebnis bereits mit wenigen Rauheitszonen erzielen.

1 Einleitung

Räumliche und zeitliche morphologische Veränderungen haben eine lebensraumprägende Bedeutung in der gesamtheitlichen Fließgewässerbetrachtung. Variierende Gewässerbettstrukturen bieten vielfältige Lebensräume und wirken sich positiv auf die aquatische Ökologie aus. Um solche kleinräumigen morphologische Strukturen mit hohem Detailgrad abbilden zu können, braucht es hochaufgelöste topografische Informationen der Gewässersohle. Die herkömmliche Methode der Querprofilvermessung, deren Qualität stark von subjektiven Entscheidungen der Vermesser abhängt und einen großen zeitlichen sowie personellen Aufwand darstellt, kann durch die topobathymetrische Laserscanvermessung (ALB - Airborne LiDAR Bathymetrie) ersetzt und darüber hinaus erheblich verbessert werden. Dabei operiert ein wasserdurchdringender Laser im grünen Wellenlängenbereich ($\lambda = 532$ nm) und erfasst die morphologischen Fließgewässerstrukturen mit einer Punktdichte von bis zu 50 Punkte/m². Dieses Verfahren kann selbst in unzugänglichen alpinen Regionen erfolgreich eingesetzt werden (AUFLEGER et al. 2015; KINZEL et al. 2013; MANDEL-BURGER et al. 2011; STEINBACHER et al. 2012).

Diese Arbeit entsteht im Rahmen des Forschungsprojektes "FHARMOR" - Fish Habitat in Alpine Rivers: Integrating Monitoring, Modelling and Remote Sensing (<u>www.fharmor.it/</u>). Es werden die Vorteile von periodischen Airborne bathymetrischen Lidar Aufnahmen anhand des Gebirgsbaches Mareiterbach untersucht. Mit einer wiederholten Untersuchung ist es möglich, morphologische Änderungen zu identifizieren. Das Hauptaugenmerk liegt auf der Validierung der Datenqualität,

¹ Universität Innsbruck, Institut für Infrastruktur, Arbeitsbereich für Wasserbau, Österreich.

der Überprüfung der refraktionierten wasserbenetzten Punkte, der Anwendbarkeit in der hydraulischen Modellierung und der Bedeutung der periodischen systematischen Flussbettaufnahmen für die Geschiebesimulation.

2 Methodik

2.1 Untersuchungsgebiet

Das Untersuchungsgebiet des Mareiterbaches (Abbildung 1) liegt in der Provinz Südtirol (Italien) und ist gekennzeichnet durch ein nivo-galziales Abflussregime mit starken täglichen Schwankungen bedingt durch die große Gletscherfläche im Einzugsgebiet. Die höchste Erhebung bildet der 3471 müA hohe Gipfel der Sonnklarspitze. Die zwei Zubringer Ratschingers- und Jaufentalbach münden im breiten Trogtal in den Mareiterbach. Im betrachteten Gewässerabschnitt ist eine Pegelmessstelle zur kontinuierlichen Wasserstandsaufzeichnung situiert und weist einen mittleren jährlichen Abfluss von 7 m³/s auf.



Abb. 1: Übersicht über Lage des Untersuchungsgebietes Mareiterbach in der Provinz Südtirol (Italien)

Der ausgewählte Flussabschnitt wurde im Jahr 2009 durch ein Revitalisierungsprojekt in ein dynamisches Flusssystem übergeführt, welches heute ein mittleres Gefälle von 1,3 % aufweist. Bachforellen (Salmo trutta fario) und Marmorierte Forellen (Salmo trutta marmoratus) dominieren die Fischgemeinschaft und eine flussaufwärts sowie flussabwärtsgerichtete Fischdurchgängigkeit ist wiedergegeben.

2.2 Airborne LiDAR Bathymetrie – Klassifizierung, Refraktion und Datenqualität

Der Mareiterbach, Nähe Sterzing in Südtirol, ist geprägt durch die hohe Abflussdynamik des teilweise vergletscherten Einzugsgebiets. Die dadurch hervorgerufene räumliche und zeitliche Variabilität der Flusssohle wird durch wiederholte hochaufgelöste topo-bathymetrische Laserscanvermessungen aus der Luft erfasst. Die erste Befliegung wurde am 15. Dezember 2016 unter Niederwasserbedingungen von der Firma AHM - AirborneHydroMapping GmbH (www.ahm.co.at)

durchgeführt. Es wurde der topo-bathymetrische Sensor VQ-880-G (RIEGL 2018) von der österreichischen Firma RIEGL Laser Measurement Systems GmbH eingesetzt. Die Befliegung umfasst 12 Scanstreifen, um einen konsistenten Datensatz zu erhalten. Die Befliegung wurde durch zeitgleiche Feldmessungen (Abflussmessungen mittels FlowTracker2, terrestrische Messungen mit Tachymeter Leica Viva TS15) begleitet. Die zweite Befliegung erfolgte rund zwei Jahre später am 18. November 2018 unter ähnlichen Verhältnissen. Die Bearbeitungsschritte der Datenverarbeitung beider Befliegungsdatensätze inkludiert die Datenkontrolle, gefolgt vom Streifenabgleich im PreProcessing. Für die Transferierung der ALB Punktwolke in das gebräuchliche UTM Koordinatensystem mit dem Referenzsystem 1989 (ETRS89) wurden Passflächen im Feld aufgenommen. Im anschließenden PreProcessing wurde die Software HydroVISH, welche die Vorteile des Datenformates HDF5 (HDF-GROUP 2013) nutzt, verwendet. Am morphologisch hoch strukturierten Mareiterbach wurde an die automatische Klassifizierung eine manuelle Klassifizierung nachgeschaltet, um die Komplexität der lokalen Geometrie besser zu berücksichtigen. Die Datenbearbeitung erfolgte in enger Zusammenarbeit mit der Firma AHM. In Abbildung 2 ist das Ergebnis der klassifizierten Punktwolke illustriert.



Abb. 2: Klassifizierte Punktwolke: schwarze und graue Punkte zeigen die Topographie und Bathymetrie, blau zeigt den extrahierten Wasserspiegel

Eine Datenkontrolle zur Qualitätssicherung ist unerlässlich (MCKEAN et al. 2014). Durch die zeitgleiche terrestrische Vermessung war es möglich, die Differenz zwischen Höheninformationen aus der klassifizierten Punktwolke (z_{ALB}) und der terrestrischen Punkte (z_{terr}) zu bilden. Die Ergebnisse werden durch die mittlere Abweichung (Median), dem RMSE (root-mean-squarequare-error) und der Standardabweichung (std) beschrieben. Der RMSE wurde berechnet mit

$$RMSE = \sqrt{\frac{\Sigma(z_{ALB} - z_{terr})^2}{n}}$$
(1)

wobei z_{terr} die terrestrische Höhe, z_{ALB} die abgeleitete Höhe aus der klassifizierten LiDAR-Punktwolke am selben Ort und *n* die Anzahl der terrestrisch vermessenen Punkte ist. Beim Medienübergang Luft zu Wasser kommt es zu einer Änderung der Ausbreitungsrichtung des

Laserstahls. Der Lichtstrahl wird gebrochen (Refraktion) und gleichzeitig verlangsamt sich die Ausbreitungsgeschwindigkeit unter Wasser. Daher ist eine sogenannte Refraktionskorrektur notwendig (WESSEL et al. 2016). Abbildung 3 zeigt eine schematische Übersicht der Refraktion des ALB Signals im grünen Wellenlängenbereich. Für die Korrektur wird ein Modell der Wasseroberfläche benötigt. Mit diesem Wasseroberflächenmodell und den Trajektorien kann dann eine Längen- und Winkelkorrektur der Bathymetrie-Punktwolke erfolgen. Eine interessante Fragestellung ergibt sich nun aus dem Vergleich möglicher Wasseroberflächenmodelle. Das erste Vergleichsmodell ist eine Triangulation der vermessenen Wasserspiegelpunkte der ALB Punktwolke. Als zweites Modell dient eine Triangulation aus simulierten Wasserspiegelpunkten, gewonnen aus einer Fließgewässersimulation mit rasterförmigem Berechnungsnetz der Auflösung 0,5 m und dem Abfluss zur Befliegungsstunde als hydraulische Randbedingung.



Abb. 3: Schematische Darstellung der Ablenkung des ALB Signals (grüner Wellenlängenbereich)

2.3 Hydraulische Simulation

Auf der hervorragenden Basis der zwei Befliegungen wurden verschiedene Berechnungsnetze für das 1,5 km lange Untersuchungsgebiet erstellt. Für die numerische Simulation wurde die kommerzielle Software HYDRO AS-2D (HYDROTEC 2015) verwendet. Die Berechnung basiert auf einer 2D-tiefengemittelten Strömungsgleichung mit der räumlichen Diskretisierung nach der Finiten-Volumen Methode. Im Rahmen dieser Untersuchung wurden Berechnungsnetze mit unterschiedlichen praxisorientierten Methoden erstellt. Die Punktdezimierung der klassifizierten Punktwolke (140 Punkte/m²) wurde mit einer statistischen Ausdünnung in einem regelmäßigen Raster durchgeführt. Die Rastergröße reicht von 0,25 x 0,25 m bis 1,00 x 1,00 m und wurden im nächsten Schritt zu einem Dreiecksnetz verarbeitet. Eine weitere Möglichkeit der Netzdezimierung ist die Reduktion mit dem "Visualization Tool Kit (VTK)". Es entsteht ein unstrukturiertes unregelmäßiges Dreiecksnetz. Für die einfache rasche numerische Lösung sind einfache Formen notwendig, daher wurde in einer dritten Variante ein optimiertes Netz erstellt, welches sich an der Gewässerform orientiert und aus Viereckelementen mit einem Seitenverhältnis von 1:2 besteht (Kurzbezeichnung: E4Q). Um die Bedeutung der hochaufgelösten Geländedaten aus der Airborne LiDAR Befliegung hervorzuheben, wurde ein Netz auf Basis von Querprofilen (QP) erzeugt. Die Punktinformationen wurden aus der ALB Punktwolke abgeleitet. Alle angeführten Berechnungsnetze bringen Vor- und Nachteile mit sich und wurden im Detail analysiert. Eine zentrale Frage ist, wie sich die Netztopologie und -auflösung auf die Kalibrierung (den Kalibrierungsprozess) auswirken. Für die Kalibrierung des hydraulischen Simulationsmodells wird der flächenhafte Wasserspiegel herangezogen, welcher bei der Airborne LiDAR Befliegung aufgenommen wurde. Als obere Randbedingung des numerischen Modells wird ein Abfluss von 1,70 m³/s angesetzt. Dieser wurde am Befliegungstag mittels FlowTracker2 gemessen. Für eine anschauliche Ergebnisinterpretation werden über das Untersuchungsgebiet Auswertungs-Querprofile im Abstand von ca. 10 m verteilt. An diesen werden die Differenzen zwischen den gemittelten Wasserspiegeln aus der ALB Vermessung und den simulierten Wasserspiegeln gegenübergestellt.

3 Ergebnisse und Diskussion

Zur Beurteilung der Datenqualität werden die Unterschiede zwischen der terrestrischen Punkt-Vermessung und der ALB-Höhe an der entsprechenden Position gegenübergestellt. Die wesentlichen statistischen Parameter (Median, RMSE, Standardabweichung) der Bodenpunkt- und Wasserspiegelgenauigkeit sind in Abbildung 4 zusammen mit den Histogrammen dargestellt. Beide Histogramme zeigen nahezu eine Normalverteilung. Die Mediane liegen bei -0,014 m und - 0,024 m. Das negative Bias liegt innerhalb der GPS-Genauigkeit der terrestrischen Punktvermessung. Die Abweichung ist somit nicht signifikant. Dieses positive Ergebnis bescheinigt der ALB-Technologie eine große Höhengenauigkeit. Die Fluss-Bathymetrie kann somit mit geringen geometrischen Unsicherheiten in numerischen Netzen abgebildet werden. Zudem ist der ALB-Wasserspiegel hervorragend für die Kalibrierung geeignet.



Abb. 4: Datenqualität: Genauigkeit ALB-Höhe (z_{ALB}) versus terrestrisch vermessener Punkthöhe (z_{terr}), links: Genauigkeit der Bodenpunkte; rechts: Wasserspiegelgenauigkeit

Zur Beurteilung der Auswirkungen unterschiedlicher Wasseroberflächenmodelle auf die Refraktionskorrektur und somit auf die Lagen- und Höhengenauigkeit von unter der Wasseroberfläche liegenden Bathymetrie-Bodenpunkten werden zwei Methoden verglichen. Zum einen wird eine Triangulation jener ALB-Punkte als Wasseroberflächenmodell verwendet, die als Wasseroberflächenpunkte klassifiziert wurden. Vor allem in Randbereichen mit geringen Wassertiefen ist eine solche Klassifizierung schwierig bzw. manchmal sogar unmöglich. Daher hat diese Methode Anwendungsgrenzen. Die zweite Methode bedient sich einer hydraulischen Simulation und verwendet den simulierten Wasserspiegel als Wasseroberflächenmodell. Die Bathymetrie mit Refraktionskorrektur gemäß Methode Eins dient hier als Basis für ein feines Berechnungsnetz mit einer Rasterauflösung von 0,5 m. Eine stationäre Berechnung des Abflusses zur Befliegungsstunde ergibt die Wassertiefen in jedem Knotenpunkt. Die Rauheit wird flächig kalibriert, sodass der Mittelwert des berechneten Wasserspiegels über alle Netzknoten jenem des gemessene Wasserspiegels (klassifizierte ALB-Punktwolke) gleicht. Das Wasseroberflächenmodell aus Methode Zwei (Sim_{Boden}) enthält Bereiche mit geringen Wassertiefen und zeigt lokale Wasserspiegelgefälle sehr detailreich, speziell in ufernahen Bereichen.

Die Refraktionskorrektur ergibt zwei unterschiedliche resultierende Bathymetrien: ALB_{Boden} gemäß Methode Eins und Sim_{Boden} gemäß Methode Zwei. Abbildung 5 zeigt mittels Boxplot (links) und Differenzenkarte (rechts) das Ergebnis einer Gegenüberstellung. In der Karte (rechts) ist die Differenz zwischen ALB_{Boden} und Sim_{Boden} als Differenzenplot dargestellt (ALB-Sim). Es zeigen sich räumlich kleine Unterschiede. Der Mittelwert aller Abweichungen liegt bei 0,001 m und die Standardabweichung bei 0,013. Die markantesten Abweichungen treten bei steilen Ufern bzw. abrupt abfallenden Stellen wie großen Steinen auf. Große Differenzen entstehen dort, wo die Simulation keine Wasserspiegel aufweist. Hier ist z. B. der nicht benetzte Seitenarm in der Simulation zu nennen. Diese Fehler sind aber für die praktische Anwendung von untergeordneter Bedeutung. Zusammengefasst lässt sich aus den Ergebnissen ableiten, dass ein Wasseroberflächenmodell gemäß Methode Eins aus der klassifizierten ALB-Punktwolke gut für die Refraktionskorrektur geeignet ist.



Abb. 5: Ergebnis der Refraktionskorrektur; links: Boxplot der Differenz zwischen den Bathymetrien nach Methoden Eins und Zwei; rechts: Differenzenplot, flächige Darstellung lokaler Abweichungen

Die hydronumerische Berechnung der flächigen Fließgeschwindigkeiten und –tiefen mit dem gemessenen Abfluss zur Befliegungsstunde ermöglicht eine Analyse der Auswirkungen verschiedener Netztopologien und -auflösungen. Die in Kapitel 2.3 beschriebenen Berechnungsnetze wurden mit unterschiedlichen Reduktionsmethoden direkt aus der ALB-Punktwolke abgeleitet. Tabelle 1 bietet einen Überblick der resultierenden Netze mit ihren wichtigsten Eigenschaften wie der Anzahl der Knoten, der Punktdichte und der Art der Netzelemente.

Bezeichnung	Anzahl Knoten	Punktdichte [Pt/m ²]	Netzelemente
R25	~ 2.141.000	~ 15 Pt/m²	Dreieckselemente
R50	~ 547.000	~ 4 Pt/m²	Dreieckselemente
R100	~ 138.000	~1 Pt/m²	Dreieckselemente
VTK	~ 646.000	~ 5 Pt/m²	Dreieckselemente
E4Q 0,5x1,0m	~ 194.000	~ 1,5 Pt/m²	Viereckelemente
QP 1,0x1,5m	~ 77.000	~ 0,5 Pt/m²	Viereckelemente

Tab. 1: Übersicht über die verschiedenen Berechnungsnetze

In der hydraulischen numerischen Modellierung wird der Rauheitsbeiwert nach Manning-Strickler (k_{st}) als Stellschraube für die Kalibrierung verwendet. Mit kalibrierten Rauheitsbeiwerten können kleinräumige komplexe Strukturen wie Felsen und steile Abstürze selbst bei grober Netzauflösungen zumindest näherungsweise modelliert werden. Es werden insgesamt drei räumliche Verteilungen für die Rauheitsbeiwerte untersucht. Die einfachste Rauheitszonenverteilung ist eine Belegung mit einem einheitlichen Rauheitsbeiwert, welcher im Zuge der Kalibrierung zwischen k_{st}=17 und k_{st}=30 variiert wird. Der Flussabschnitt am Mareiterbach ist jedoch räumlich stark strukturierte und morphologisch vielfältig ausgeprägt. Daher werden zwei weitere differenzierte Zonenverteilungen untersucht, die durch eine großräumige Einteilung "large-scale" bzw. eine kleinräumige Einteilung "small-scale" charakterisiert sind. In Abbildung 6 sind diese beiden differenzierten Rauheitszonenverteilungen dargestellt.



Abb. 6: Gebiet Mareiterbach, Übersicht über die Rauheitszonenverteilung, obere Abbildung "large-scale" und untere "small-scale" Rauheitszonenverteilung

Abbildung 7 stellt die Ergebnisse der einzelnen Berechnungsmodelle einander gegenüber, indem die jeweilige Übereinstimmung der simulierten Wasserspiegel mit der gemessenen Wasserspiegelage aus der klassifizierten ALB-Punktwolke ausgewertet wird. Die Boxplot Übersicht gibt die Differenz zwischen simulierten Wasserspiegel und dem Wasserspiegel aus der ALB Befliegung, gemittelt über die Querprofile, welche im Abstand von ca. 10 m verteilt über die Längsachse liegen, wieder. Neben dem besten Ergebnis mit einer Rauheitszone (schwarze Boxplots) sind die Ergebnisse mit der Rauheitszonenverteilung "large scale" (blau eingefärbt) und "small-scale" (grüner Boxplot) ersichtlich. Tabelle 1 zeigt die zugehörigen statistischen Kennwerte. Ein wesentliches Ergebnis ist der Zusammenhang zwischen Modellauflösung und Ausreißern, welche über Whisker (5% ~ 95%) dargestellt werden. Die Auswirkungen unterschiedlicher Netztopologien kann durch die Wasserspiegelabweichungen quantifiziert werden. Diese zeigen, dass bei hoch aufgelösten Berechnungsnetzen wie z. B. beim VTK oder Raster50 (R50) Netz das Streuungsmaß kleiner ist.



Abb. 7: Boxplot der Differenz aus simulierten und ALB Wasserspiegel Quantilwerten für unterschiedliche Berechnungsnetztopologien. Schwarzer Boxplot bestes Ergebnis mit einem Rauheitsbeiwert, Blauer mit der "large-scale" Rauheitszonenverteilung und grüne Farbe ist das Ergebnis mit der detaillierten "small-scale" Einteilung

Die Ergebnisse zeigen ferner, dass mit einer detaillierteren räumlichen Einteilung der Rauheitszonen die statistischen Ergebnisse verbessert und gleichzeitig die Abweichungen minimiert werden. Dieser Effekt tritt bei allen Berechnungsnetztopologien auf. Von größerer Bedeutung ist dies jedoch erst bei Berechnungsnetzen mit niedriger Auflösung wie z. B. bei Raster100 (R100) oder E4Q 0,5x1,0 m, da die fehlende geometrische Strukturgenauigkeit im Netz dann mit der detaillierteren Rauheitsanpassung kompensiert werden kann. Dieser Vergleich führt deutlich vor Augen, dass die Oberflächen- und Formrauheit speziell bei niedrig aufgelösten Berechnungsnetzen, in Form einer detaillierten Kalibrierung mehrerer Rauheitszonen berücksichtigt werden muss. Dreiländertagung der DGPF, der OVG und der SGPF in Wien, Österreich – Publikationen der DGPF, Band 28, 2019

Bezeichnung Mittelwert [m]		RMSE [m]	Standartabweichung [m]		
R25 kst 30	-0,004	0,030	0,030		
R50 kst 25	0,004	0,036	0,035		
R50 kst large	0,000	0,026	0,026		
R50 kst small	0,002	0,022	0,022		
R100 kst 25	0,000	0,040	0,035		
R100 kst large	-0,003	0,030	0,026		
R100 kst small	-0,002	0,024	0,022		
VTK kst 30	-0,002	0,030	0,030		
VTK kst <i>large</i>	0,011	0,026	0,024		
VTk kst small	0,012	0.025	0,022		
E4Q 0,5 kst 25	-0,003	0,039	0,039		
E4Q 0,5 kst large	-0,005	0,030	0,025		
E4Q 0,5 kst small	-0,004	0,023	0,023		

Tab. 1: Übersicht über die verschiedenen Berechnungsnetztopologien

Bis noch vor wenigen Jahren wurde die Bathymetrie durchwatbarer Gewässerabschnitte standardmäßig mittels terrestrischer Querprofilvermessung erhoben. Um das Verbesserungspotential einer flächigen ALB-Vermessung gegenüber der klassischen Querprofilvermessung quantitativ hervorzuheben, wird als weitere Netztopologie ein Berechnungsnetz aus fiktiven Querprofilen untersucht. Die Querprofilspuren werden entlang markanter Geländesprünge gelegt, ganz so wie es versierte Vermesser im Feld tun würden. Der Querprofilabstand variiert somit über das Untersuchungsgebiet und liegt zwischen 5 und 40 m. Abbildung 8 und Tabelle 2 zeigen die Ergebnisse für diese Querprofilberechnungsnetztopologie jeweils für die drei Rauheitszonenverteilungen, welche dem Raster25 (R25) gegenübergestellt wird.



Abb. 8: Boxplot der Differenz aus simulierten und ALB Wasserspiegel Quantilwerten für die Querprofilberechnungsnetze. Schwarzer Boxplot zeigt bestes Ergebnis mit einem Rauheitsbeiwert, Blauer mit der "large-scale" Einteilung und grüne Farbe ist das Ergebnis mit der detaillierten "small-scale" Einteilung

Die Abweichung zwischen gemessenen ALB Wasserspiegel und simuliertem Ergebnis liegt im Mittel unter -0,004 m. Die Abweichungen über den RMSE sind stets größer als 0,080 m. Die erhöht auftretenden Ausreißer führen zu größeren Streuungswerten, welche auf die nicht erfasste Geometrie z. B. fehlende Abstürze zurückzuführen sind. Mit diesem Ergebnis kann klar gezeigt werden, dass die Anwendung der Airborne LiDAR Bathymetrie ein sinnvolles Potential beinhaltet, um hydraulische Bedingungen im Detail abzubilden.

Bezeichnung	Mittelwert [m]	RMSE [m]	Standartabweichung [m]
R25 kst 30	-0,004	0,030	0,030
QP kst 25	-0,003	0,091	0,091
QP kst <i>large</i>	-0,004	0,084	0,084
QP kst small	-0,002	0,081	0,081

Tab. 2: Übersicht über das Berechnungsnetz QP (Querprofilnetz)

4 Fazit & Ausblick

Die Erfassung der 'Unterwassertopographie' mit ALB hat gezeigt, dass ein hohes Potential in der kleinmaßstäblichen Aufnahme der morphologischen Strukturen steckt. Die Auswertungen der numerischen Simulationen verdeutlichen, dass der flächige Wasserspiegel aus der ALB Vermessung zur Kalibrierung geeignet ist. Die Kontrolle der Refraktionierung der wasserbenetzten Punkte zeigt eine gute Übereinstimmung. Die Vorgehensweise mit dem Wasseroberflächenmodell wird somit als geeignet bestätigt. Es zeigt sich, dass die Auswahl der Auflösung einen Einfluss auf den Kalibrieraufwand hat. Mit einem Rasternetz von 0,50x0,50 m (R50) kann mit kleinem Kalibrieraufwand und nur wenigen Rauheitszonen eine sehr gute Übereinstimmung der Wasserspiegellagen erzielt werden. Aus der Untersuchung geht weiters hervor, dass die Netzerstellung aus Querprofildaten große lokale Abweichungen der simulierten Wasserspiegellagen mit sich bringt, welche auf die fehlenden morphologischen Strukturen, die das Untersuchungsgebiet prägen, zurückzuführen sind.

Um das Potential der topo-bathymetrischen Aufnahme auch für morphologische Untersuchungen voll nutzen zu können, bedarf es einer wiederholenden Datenerhebung. Dies stellt ein gutes Instrument zur Erfassung der hydro-morphologischen Veränderungen und einhergehenden Lebensräume. Eine zweite ALB Befliegung wurde am 18. November 2018 erfolgreich durchgeführt. Mit dieser Datengrundlage ist es möglich, ein Volumenbilanzierung im Detail zu erstellen, welche für die Weiterentwicklung der Genauigkeit der Geschiebetransportmodellierung dienen soll.

Dreiländertagung der DGPF, der OVG und der SGPF in Wien, Österreich – Publikationen der DGPF, Band 28, 2019

5 Literaturverzeichnis

- AUFLEGER, M., STEINBACHER, F., BARAN, R. & VALERIE, N., 2015: Airborne Hydomapping -Hochaufgelöste Gewässervermessung "wie im Flug". Dresdner Wasserbauliche Mitteilungen 53, Universität Dresden, 415-424.
- DOBLER, W., STEINBACHER, F., BARAN, R., RITTER, M. & AUFLEGER, M., 2014: High Resolution Bathymetric Lidar Data to Hydraulic - Modelling a Mountain Stream by Bathymetric Lidar Data. Proceedings of the 11th International Conference on Hydroinformatics (HIC 2014).
- HDF-Group, o.J.: The HDF5 Home Page. http://www.hdfgroup.org (11.12.2018).
- MANDELBURGER, G., PFENNINGBAUER, M., STEINBACHER, F. & PFEIFER, N., 2011: Airborne Hydrografic LiDAR Mapping - Potential of a new technique for capturing shallow water bodies. 19th International Congress on Modelling and Simulation, Perth, Australia, 2416-2422.
- MCKEAN, J., TONINA, D., BOH, C. & WRIGHT, C.W., 2014: Effects of bathymetric lidar errors on flow properties predicted with a multi-dimensional hydraulic mode. Journal of geophysical research: Earth Surface, **119**(3), 664-664.
- KINZEL, P.J., LEGLEITER, C.J. & NELSON, J.M, 2013: Mapping River Bathymetry with a Small Footprint Green LiDAR: Applications and Challenges. JAWRA Journal of the American Water Resources Association, 49, 183-204.
- RIEGL, 2018: VQ-880-G Datenblatt 2018, http://www.riegl.com/uploads/tx_pxpriegldownloads/ RIEGL_VQ-880-G_DataSheet_2018-07-04.pdf (28.08.2018).
- STEINBACHER, F., PFENNINGBAUER, M., AUFLEGER, M. & ULLRICH, A., 2012: High Resolution Airborne Shallow Water Mapping. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 39(B1), 55-60.
- WESSEL, M., ANSELMETTI, F., BARAN, R., HILBE, M., GESSLER, S. & WINTERSTELLER, P., 2016: Schlussbericht: Tiefenschärfe - Hochauflösende Vermessung Bodensee, Ein Interreg IV-Forschungsprojekt von 2012 – 2015, Internationale Gewässerschutzkommission für den Bodensee (IGKB), Blaue Reihe, Bericht Nr. 61(Hrsg.).

Voxelbasierte Änderungsdetektion des Gewässerbodens in ALB Daten unter Verwendung von semantischen Referenzdaten

RICHARD BOERNER¹, LUDWIG HOEGNER¹ & UWE STILLA¹

Zusammenfassung: Dieser Beitrag befasst sich mit der Änderungsdetektion in bathymetrischen Punktwolken, die mit einem Airborne Laserscanner (ALS) aufgenommen worden sind. Ziel ist es, Änderungen der Wasseroberfläche und des Gewässerbodens automatisch zu erkennen. Hierfür ist neben der Schätzung der Wasseroberfläche aus der Punktwolke auch eine Refraktionskorrektur für Punkte unter der Wasseroberfläche nötig. Für den Referenzdatensatz liegen dabei aus einer vorangegangenen Klassifikation bereits auf Punktebene Label vor, um zwischen statischen, für die Koregistrierung geeigneten, und dynamischen Bereichen mit potentiellen Änderungen zu unterscheiden. In einem ersten Schritt werden die vorhandenen Label des Referenzdatensatzes über einen diskreten Voxelraum auf die neue Punktwolke übertragen. Falsch zugeordnete Label im dynamischen Bereich werden basierend auf einer Histogrammanalyse identifiziert und entsprechend geändert. Aus den so klassifizierten Punkten wird die Wasseroberfläche geschätzt und für die darunterliegenden Punkte eine Refraktionskorrektur durchgeführt. Die Belegungsanalyse im Voxelraum liefert eine Aussage darüber, welche dynamischen Bereiche der neuen Punktwolke mit der Referenzpunktwolke verglichen werden können, in welchen Bereichen eine Refraktionskorrektur anzubringen ist aber fehlende Wasserechos die Korrektur erschweren und welche Bereiche zwar korrigiert werden konnten aber hier keine Daten in der Referenz vorliegen.

1 Einleitung

Die Grundvoraussetzung für eine Änderungsdetektion des Gewässergrundes in der Airborne Laserbathymetrie (ALB) ist die Klassifikation von Wasseroberfläche und Gewässerboden. Die Schätzung der Wasseroberfläche ist notwendig, da der Laserstrahl beim Übergang vom Medium Luft ins Medium Wasser gebrochen wird und die Nichtberücksichtigung dieser Brechung zu fehlerhaften 3D Koordinaten der Bodenpunkte führt. Im besonderen Fall der Fließgewässervermessung können diese beiden Klassen als zeitlich veränderbar angenommen werden. So ändert sich sowohl der Wasserstand in seiner Höhe als auch der Gewässerboden in seiner Geometrie aufgrund von morphodynamischen Einflüssen des Flusses. Eine Überwachung der morphodynamischen Einflüsse erfordert eine multitemporale Erfassung des Untersuchungsgebietes und dementsprechend multitemporale gelabelte Daten.

Dieser Beitrag befasst sich mit der Übertragung von klassifizierten Referenzdaten auf zusätzliche Befliegungsdaten. Hierbei sollen sowohl die Möglichkeit der Kombination von Daten verschiedener Sensoren als auch die dynamische Natur der Klassen berücksichtigt werden. Die Übertragung der Labels des Referenzdatensatzes wird dazu verwendet, um eine Vorklassifikation von Änderungen im Gewässerboden vorzunehmen. Diese Klassifikation der Änderung beinhaltet die Klas-

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

sen: "Vergleichbar", "im Wasserbereich keine Refraktionskorrektur anbringbar" und "ohne Referenzinformation". Dadurch lässt sich der Suchbereich für einen Signifikanztest auf geometrische Änderungen eingrenzen und die Genauigkeit je nach vorhandener Refraktionskorrektur oder Referenzdaten angleichen.

Das vorgestellte Verfahren baut auf der Diskretisierung von 3D Punktwolken in einem Voxelgitter auf. Ein solches Voxelgitter wurde bereits zur Bodendetektion in BOERNER et al. (2017) verwendet. Jeder Punkt des neuen Datensatzes kann hierbei einem Voxel im Referenzdatensatz zugeordnet werden und demzufolge das Klassenlabel des Voxels übernehmen. Die Labelübertragung im geometrisch veränderlichen Fall ist aber schwieriger als im statischen Fall, da die sich bewegenden Punkte (Wasser, Flussboden) in nicht gelabelte Voxel projiziert, bzw. ein Bodenpunkt in ein als Wasser klassifizierten Voxel projiziert werden könnte.

Aus diesem Grund wird im Nachgang zur Übertragung der Labels geprüft, ob Falschzuweisungen der Wasserlabels erkennbar sind. Für diese Probe wird ein Höhenhistogramm bestimmt und den Voxeln korrespondierend zur geringsten Höhe die Klasse Boden zugeordnet. Dieses Verfahren beruht auf der Annahme, dass hauptsächlich Bodenechos aufgezeichnet werden. Eine Belegungsanalyse der korrigierten Punkte liefert dann die Klassen "vergleichbar" oder "in Referenz nicht belegt". Die Klasse "nicht korrigiert" ergibt sich, wenn aufgrund einer Lücke im Wassermodell kein Schnitt mit der Wasseroberfläche bestimmbar ist.

Das vorgestellte Verfahren wird an einem Befliegungsdatensatz vom Jahre 2012 vom Mangfallgebiet in Gmund getestet und evaluiert. Von diesem Gebiet liegen neuere Befliegungsdaten vom Jahr 2017 vor, welche mit einem neueren Sensormodell aufgenommen wurden und eine erhöhte Punktdichte aufweisen.

Der Vorteil des vorgestellten Verfahrens besteht darin, dass es auch auf multisensoriellen Daten anwendbar ist und so für eine Übertragung von semantischer Information und eine Qualitätsbeurteilung von zusätzlichen UAV Daten verwendet werden kann. Vorarbeiten zur Registrierung von Referenz- und UAV-Daten wurde in BOERNER et al. (2018) vorgestellt.

2 Vorarbeiten

In den letzten Jahren sind ALB Sensoren immer weiterentwickelt worden und stellen nun eine Möglichkeit zur genaue Vermessung von großflächigen Flachwasserbereichen dar (MANDLBUR-GER et al. 2011; PFENNIGBAUER et al. 2010; STEINBACHER et al. 2012). Besonders Interessant im Bereich der Gewässervermessung ist die Änderungsdetektion des Flussbettes, womit sich z.B. Aussagen über die Lebensqualität von Fischen treffen lassen (WHEATON et al. 2010a).

Für die Änderungsdetektion finden sich je nach Anwendung und verwendeten Sensoren unterschiedliche Herangehensweisen und Lösungsansätze. Im Kontext der Gewässervermessung und im Bereich der Laserbathymetrie erfolgt eine Änderungsdetektion meistens zur Erstellung von Änderungskarten des Gewässerbodens. Hierbei werden die Daten zunächst in Boden- und Wasserpunkte klassifiziert und anschließend ein Geländemodell auf Grundlage der korrigierten Bodenpunkte berechnet. Dies geschieht für den Referenz- und den Änderungsdatensatz und der Vergleich beider Datensätze erfolgt durch eine Bildung der Differenz der Oberflächenmodelle. Die Beträge der Differenz werden dann einem statistischen Test unterzogen und auf Signifikanz getestet, um eine tatsächliche geometrische Änderung von einer auf Messrauschen basierten Änderung zu unterscheiden (WHEATON et al. 2010b). Hierbei ist die Klassifikation in Wasser- und Bodenpunkte ausschlaggebend für die Erzeugung des Geländemodells. Zum einen sollen nur Bodenpunkte zur Berechnung des Modells benutzt werden und zum anderen wird die Wasseroberfläche benötigt, um die Punkte am Gewässergrund aufgrund der Refraktion des Laserstrahls am Medienübergang von Luft zu Wasser zu korrigieren. Eine Klassifikation in Wasser und Boden erfolgt hierbei entweder semiautomatisch (MANDLBURGER et al. 2015) oder auf Grundlage der Fullwaveform-Auswertung des Laserstrahls. In der Praxis ergeben sich aber hier Probleme, wenn nicht genügend Wasserechos vorliegen oder der Laserstrahl das Wassermedium nicht bis zum Boden durchdringen kann.

Mit der Annahme, dass semantische Referenzdaten mit räumlich zugeordneten Klassenlabeln vorliegen, ergibt sich nun die Frage, ob diese zur Klassifikation neuer Daten verwendet werden können. Hierbei soll auch eine Klassifikation von Änderungen erfolgen, welche Bereiche maskiert, die sich geometrisch geändert haben, Bereiche die in der Referenz nicht aufgenommen wurden oder zu denen keine Aussage gemacht werden kann, da die Refraktionskorrektur aufgrund fehlender Echos nicht vorgenommen werden konnte.



Abb. 1: Prinzipskizze des vorgestellten Verfahrens

3 Methode

Das vorgestellte Verfahren ist in Abbildung 1 schematisch veranschaulicht. Der erste Schritt besteht in der Übertragung der semantischen Label der Referenzpunktwolke auf die neu aufgenommene Punktwolke. Dies geschieht unter Verwendung eines Voxelgitters. Das Voxelgitter wird in die Referenzpunktwolke gelegt und jeder Voxel mit einer räumlichen ID versehen. Bei gleicher Unterteilung des Voxelgitters in der anderen Punktwolke lässt sich so die Belegung jedes semantischen Referenzvoxels im neuen Datensatz prüfen. Wird ein entsprechend belegtes Voxel in den Referenzdaten gefunden, wird hier der Median der enthaltenen Klassenlabels der Punkte der Referenzpunktwolke auf die Punkte der neuen Punktwolke im entsprechenden Voxel übertragen. Neue Voxel ohne entsprechende Korrespondenz werden hierbei als geändert gelabelt. Dieses Verfahren funktioniert im Bereich von statischen Objekten ohne Erweiterung. Im Bereich des dynamischen Flussgebietes werden hier zusätzlich kleine Änderungen eingefügt.

Der Bereich des Flusses wird aus den Referenzdaten als 2.5D Wassermodel übernommen. Die Klassen "Änderung", "Boden" und "Wasser" im Flussbereich werden von den entsprechenden Klassenlabels des statischen Bereiches durch eine andere Klassen ID abgegrenzt. Diese Abgrenzung ist nötig, um im folgenden Schritt zwischen statischen und dynamischen Bereichen zu unterscheiden. Anders als im statischen Bereich werden im dynamischen Bereich die neuen Label nicht durch den Median im Referenzvoxel bestimmt. Stattdessen werden als Wasserpunkte diejenigen Punkte markiert, welche dem Wassermodel am nächsten liegen. Hierfür werden in den Referenzdaten die Mittelwerte der Höhen der Boden- und Wasserpunkte berechnet. Zur Bestimmung des Klassenlabels in den neuen Daten wird nun das nearest neighbour Verfahren zu den Mittelwerten benutzt. Lässt sich kein nearest neighbour finden, weil die Lage der Mittelwerte einen Schwellwert überschreitet wird die Klasse wie im statischen Fall bestimmt.



Abb. 2: Histogrammbasierte Bestimmung der Konfidenzbereiche von Wasser und Boden Punkten.

Aufgrund der Dynamik im Flussgebiet, muss nun getestet werden, ob Wasserlabels im Flussbereich zu Bodenlabels geändert werden sollten. Dies kann passieren, wenn sich die Lage des Bodens zur Wasseroberfläche hin verändert hat oder nur Bodenechos aufgenommen wurden (diese würden dann dem niedrigsten Signifikanzbereich zugeordnet werden). Hierfür wird im Flussbereich ein Höhenhistogramm der Boden- und Wasserpunkte erzeugt. Unter der Prämisse, dass Bodenechos häufiger aufgenommen werden als Wasserechos, wird der unterste Signifikanzbereich der Bodenklasse zugeordnet. Findet sich also ein Wasserpunkt innerhalb dieses Bereiches, wird er neu der Bodenklasse zugeordnet. Die Signifikanzniveaus sind an einem Beispielhistogramm in Abbildung 2 veranschaulicht.



Abb. 3: Entscheidungsbaum zur Klassifikation der Änderungen

Im letzten Schritt werden die Änderungslabels für die Klassen im Flussbereich aufgrund eines Entscheidungsbaumes vergeben (Abbildung 3). Diese so übertragenen Label auf den neuen Datensatz werden dann verwendet, um das Wassermodel der neuen Daten zu berechnen. Lässt sich der Laserstrahl zum konkreten Bodenpunkt mit dem neuen Wassermodell schneiden, wird die Refraktionskorrektur angebracht, andernfalls wird das Label für "keine Korrektur möglich" vergeben. Die korrigierten Bodenpunkte werden dann erneut mit den korrigierten Bodenpunkten der Referenzdaten auf Belegung überprüft. Finden sich Bodenpunkte in der Referenz, wird das Label für einen vergleichbaren Bereich vergeben, andernfalls für "keine Referenzdaten vorhanden". Diese Änderungslabels können dann verwendet werden, um die Änderung der Bodengeometrie auf Signifikanz zu prüfen und mit einem absoluten Wert der Änderung zu versehen. Der Vorteil dieser Methode besteht nun darin, dass Bereiche, welche unterschiedliche Sichtbarkeiten und unterschiedliche Anzahl von Wasserechos aufweisen, mit einem entsprechenden Label versehen werden, um ebenfalls die Qualität der Daten zu beurteilen.

4 Daten

Als Testgebiet dient das Gebiet der Mangfall vom Tegernsee bis zur Autobahn A8 im Südosten Bayerns, Deutschland. Hier liegen zwei Befliegungskampagnen vor, welche das Gebiet mit einem bathymetrischen LiDAR system aufgenommen haben. Die erste Befliegung fand 2012 statt, wobei ein Riegl VQ 860 verwendet wurde und die zweite Befliegung fand 2017 mit einem Riegl VQ 880 statt. Die Auflösung des 1. Datensatzes ist mit ca. 5 dm sowohl zwischen als auch entlang einer Scanzeile etwas geringer als die des zweiten Datensatzes, welcher 1dm entlang einer Scanzeile hat. Des Weiteren besitzt das Gebiet stellenweise sehr flache Wasserstände als auch wenige Wasseroberflächenechos in der 2. Befliegung. Für einen ausgewählten Bereich um eine Ortschaft am Tegernsee (Gmund) wurden semantische Referenzdaten manuell erzeugt. Die Rohdaten der jeweiligen Befliegung sind in Abbildung 4 dargestellt, die semantische Referenz in Abbildung 5.



Abb. 4: Rohdaten. Links Datensatz von 2012, rechts: Datensatz von 2017



Abb. 5: Referenzlabel des Datensatzes aus 2012

5 Ergebnisse

Die gezeigten Testdaten wurden zur visuellen Inspektion des vorgestellten Verfahrens verwendet. Da für die Änderungsanalyse keine Referenzdaten vorliegen, muss in diesem Beitrag auf eine nummerische Evaluierung verzichtet werden.

Als Sampling Parameter werden je nach Teil des Verfahrens unterschiedliche Werte verwendet. Zum einen führen kleinere Voxel zu einer genaueren Abgrenzung der einzelnen Klassen, erhöhen aber zeitgleich den nicht belegten Raum. So werden weniger Klassenlabels übertragen, wenn die Voxel zu klein gewählt werden. Auf der anderen Seite führen zu große Voxel zu mehr falschen Klassenzuweisungen, besonders am Grenzbereich der Klassen. Des Weiteren werden im Bereich des Flusses, verformte Voxel entlang der Z-Achse verwendet, um eine zylindrische Nachbarschaft zu verwenden. Da solch eine Nachbarschaft mehr Punkte entlang der Höhe beinhaltet sind die Höhenhistogramme besser gefüllt und es kann nach einer Trennlinie zwischen Boden und Wasser gesucht werden.



Abb. 6: Übersicht der Ergebnisse

Die Ausdehnungen der Voxel werden somit an die Auflösung des Scanners und die erwartete Durchdringungstiefe der Wassersäule gekoppelt. Bei einer Auflösung der Referenzdaten von ca. 0.5m wird für die erste Übertragung der Label eine quadratische Voxelauflösung von maximal 1m gewählt. Zum Testen der Label im Wasserbereich wird eine zylindrische Auflösung von maximal $3 \times 3 \times 12$ m gewählt um sowohl in der Fläche als auch in der Höhe mehr Punkte im Voxel zu erhalten. Bei der Klassifizierung der Änderungen wurde eine Auflösung von $2 \times 2 \times 6$ m verwendet.

In Abbildung 6 sind die Ergebnisse der jeweiligen Schritte dargestellt. Die linke Seite zeigt einen Blick von oben über den Datensatz und auf der rechten Seite sind die markierten Profile zur genaueren Begutachtung dargestellt. Die Ergebnisse zeigen, dass im statischen Bereich die Label erfolgreich übertragen werden konnten. Hier sind auch einige Änderungen sichtbar, welche aus Anbauten der Gebäude bestehen oder aus einer geänderten Bebauung. Profil 2 zeigt die Notwendigkeit der Probe auf falsch zugeordnete Wasserpunkte. Hier befindet sich auf der linken Seite ein Kanal, welcher zwischen den Aufnahmen erhöht wurde. Das führt dazu, dass sich der Boden im zweiten Datensatz auf der Höhe des alten Wasserstandes befindet, was wiederum zu einer falschen Zuordnung führt. Das histogrammbasierte Verfahren zur Neuzuordnung der wasserpunkte funktioniert hier sichtlich gut. Es basiert allerdings auf der Annahme, dass eher Bodenechos aufgezeichnet werden, als Wasserechos. Diese Annahme passt zum neuen Datensatz, aber nicht unbedingt zum alten. Somit müsste die Probe auf falsche Zuordnungen auf ein präferiertes Verhalten angepasst werden.

Als Endergebnis ist besonders in den Profilen ersichtlich, dass die Änderungsklassen eine gute Abschätzung zur Qualität und Vergleichbarkeit der Daten liefern. Die wichtigen Klassen "vergleichbar" und "nicht in Referenz" können weiterhin verwendet werden, um die vergleichbaren Bereichen mit einer quantitativen Aussage der Änderung zu versehen.

6 Zusammenfassung

Dieser Beitrag zeigt ein Verfahren zur geometrischen Übertragung von semantischen Labels einer Referenzpunktwolke auf einen neuen Datensatz. Hierbei wird davon ausgegangen, dass die Punktwolken koregistriert sind. Die Übernahme der Referenzlabel basiert hierbei im Kern auf einer Belegungsanalyse zweier Voxelstrukturen. Hierbei wird zwischen statischen Bereichen und dynamischen Bereichen im Flussgebiet unterschieden. Die wichtigen Klassen zur Gewässervermessung, Boden und Wasser, werden in einem Folgeschritt auf Falschzuordnungen geprüft. Hier muss eine präferierte Klasse definiert werden und basierend auf einem Höhenhistogramm wird für die jeweils andere Klasse entschieden, ob eine falsche Zuordnung vorliegt oder nicht. Aufgrund der wenigen Wasserechos des neuen Datensatzes und des vollständigen Bodenlayers, wurde in diesem Beitrag der Boden als präferierte Klasse gewählt. Zur weiteren Analyse erfolgt im letzten Schritt des vorgestellten Verfahrens eine Unterteilung des dynamischen Bereiches in Punkte, welche eine Refraktionskorrektur erhalten haben und mit Punkten in der Referenz verglichen werden können, in Bodenpunkte innerhalb des Referenz Wasserbereiches aber ohne Wasserecho (sprich ohne Refraktionskorrektur in den neuen Daten) und in Punkte welche eine Korrektur erhalten haben, aber keine passenden Bodenpunkte in der Referenz auffindbar sind. Dreiländertagung der DGPF, der OVG und der SGPF in Wien, Österreich – Publikationen der DGPF, Band 28, 2019

Diese sogenannten Änderungsklassen lassen sich zukünftig verwenden um den Bereich der Signifikanzanalyse abzugrenzen und innerhalb der Signifikanzanalyse unterschiedliche Genauigkeiten für nicht korrigierte Bodenpunkte oder Bodenpunkte innerhalb von Bereichen mit keinen Beobachtungen zu bestimmen.

7 Danksagung

Diese Arbeit ist gefördert von der Bayerischen Forschungsstiftung, Projekt "Schritthaltende 3D-Rekonstruktion und –Analyse (AZ-1184-15)", Sub-Projekt "Änderungsdetektion in Punktwolken". Die Luftbefliegung wurde von der Firma "SteinbacherConsult" durchgeführt.

8 Literatur

- BOERNER, R., HOEGNER, L. & STILLA, U., 2017: Voxel based segmentation of large airborne topobathymetric LiDAR data. The International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 42, 107-114.
- BOERNER, R., XU, Y., HOEGNER, L. & STILLA, U., 2018: Registration of UAV data and ALS data using point to DEM distances for bathymetric change detection. The International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, **42**(1), 51-58.
- MANDLBURGER, G., PFENNIGBAUER, M., STEINBACHER, F. & PFEIFER, N., 2011: Airborne Hydrographic LiDAR Mapping–Potential of a new technique for capturing shallow water bodies. In Proceedings of the 19th International Congress on Modelling and Simulation, Perth, Australia, 12-16.
- MANDLBURGER, G., HAUER, C., WIESER, M. & PFEIFER, N., 2015: Topo-bathymetric LiDAR for monitoring river morphodynamics and instream habitats - A case study at the Pielach River. Remote Sensing, 7(5), 6160-6195.
- PFENNIGBAUER, M., STEINBACHER, F., ULLRICH, A. & AUFLEGER, M., 2010: A novel approach to laser-based hydrographic data acquisition. In European LiDAR Mapping Forum, Salzburg, Austria.
- STEINBACHER, F., PFENNIGBAUER, M., AUFLEGER, M. & ULLRICH, A., 2012: High resolution airborne shallow water mapping. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 39(B1), 55-60.
- WHEATON, J. M., BRASINGTON, J., DARBY, S. E., MERZ, J., PASTERNACK, G. B., SEAR, D. & VER-ICAT, D., 2010a: Linking geomorphic changes to salmonid habitat at a scale relevant to fish. River research and applications, **26**(4), 469-486.
- WHEATON, J. M., BRASINGTON, J., DARBY, S. E. & SEAR, D. A., 2010b: Accounting for uncertainty in DEMs from repeat topographic surveys: improved sediment budgets. Earth surface processes and landforms: the Journal of the British Geomorphological Research Group, 35(2), 136-156.

Entwicklung von Indizes zur Ableitung der Eisenkonzentration verockerter Fließgewässer mittels hyperspektraler VNIR- und SWIR-Daten

CHRISTOPH ULRICH¹, LUTZ BANNEHR¹, MICHAEL HUPFER² & MARTINA SCHULZE³

Zusammenfassung: Durch Anhebung des Grundwassers im Lausitzer Braunkohlerevier gelangen Eisen(II), Sulfat und Säure in die umliegenden Oberflächengewässer. Das Eisen fällt in Gewässern mit neutralem pH-Wert flockenartig aus und bildet bei hohen Konzentrationen eine rotbraune Verfärbung. Zum Monitoring von Gewässern bieten sich u.a. Fernerkundungsverfahren an, um flächendeckend Informationen über den Ist-Zustand des Gewässers in Hinblick auf den Verockerungszustand zu erhalten. Zu diesem Zweck wurden zeitgleich zur Befliegung der Spree Gewässerproben erhoben. Die limnischen Parameter wurden mit den Spektraldaten korreliert. Zur quantitativen Bestimmung des Gesamteisengehalts zeigten die Ratios $R_{RS}(580) / R_{RS}(455)$ und $R_{RS}(650) / R_{RS}(455)$ die stärksten kausalen Zusammenhänge $R^2 > 0,4$ auf. Korrelationen zwischen den Eisenparametern und dem spektralen SWIR-Bereich waren nicht vorhanden. Zukünftig werden weitere komplexe Modellierungsalgorithmen zur Eisenableitung herangezogen.

1 Einleitung

Durch den intensiven Abbau von Rohstoffen im Bergbau erfolgt ein starker Eingriff in die Umwelt. Dabei werden neben der Förderung von Rohstoffen großen Mengen von Abraum und Abwässer erzeugt. Beim Tagebau, wie in der Untersuchungsregion Lausitzer Braunkohlerevier, wird zusätzlich der Grundwasserspiegel gesenkt (HÜTTL 1998). Bei diesem Prozess reagieren die im Boden befindliche Eisendisulfidminerale, Pyrit oder Markasit, mit Luftsauerstoff und Niederschlagswasser. Dies führt zum Zerfall der Minerale in Sulfat, Eisen(II)-Ionen und Wasserstoff (SÁNCHEZ 2008; SCHULTZE et al. 2010). Die Pyritverwitterung ist in Gl. 1 in Tab. 1 beschrieben (EDWARDS et al. 1999; MOSES et al. 1987; SCHULTZE et al. 2010). Mit der Verminderung der Tagebautätigkeiten im Lausitzer Braunkohlerevier begannen Anfang 1990 die Renaturierungsmaßnahmen, wodurch der Grundwasserspiegel in einigen Regionen wieder angehoben wurde. Die freigesetzten Stoffe in der Hydrosphäre gelangen über das Grundwasser in die umliegenden Oberflächengewässer (BILEK et al. 2016). Dort oxidieren die eigentragenden Eisen(II)-Ionen mit dem im Wasser befindlichen Sauerstoff zu Eisen(III)-Oxid (s. Gl. 2) (LUTHER 1987; SCHULTZE et al. 2010; SINGER & STUMM 1970). In Gewässern mit neutralem pH-Wert hydrolisiert es zu Eisen(III)-Hydroxid und flockt aus (s. Gl. 3), (GLEISNER & HERBERT 2002; SCHULTZE et al. 2010; SINGER & STUMM 1970). Bei einer niedrigen Eisen(III)-Konzentration ist dies kaum sichtbar. Ab einer Konzentration von ca. 2 mg/L führt die Ausfällung zu einer gelben bis rotbraunen Verfärbung des Gewässers. Dieser Prozess wird als Verockerung bezeichnet.

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

² Leibniz-Institut für Gewässerökologie und Binnenfischerei, Abteilung Chemische Analytik und Biogeochemie, Müggelseedamm 301, D-12587 Berlin, E-Mail: hupfer@igb-berlin.de

³ Hochschule Anhalt, Angewandte Biowissenschaften und Prozesstechnik, Bernburger Straße 55 D-06366 Köthen, E-Mail: martina.schulze@hs-anhalt.de

$FeS_2 + \frac{7}{2}O_2 + H_2O \rightarrow Fe^{2+} + 2SO_4^{2-} + 2H^+$	Gl. 1
$Fe^{2+} + \frac{1}{4}O_2 + H^+ \rightarrow Fe^{3+} + \frac{1}{2}H_2O$	Gl. 2
$Fe^{3+} + 3H_2O \rightarrow Fe(OH)_3 \downarrow + 3H^+$	GL 3

Tab. 1: Reaktionsgleichungen der Pyritverwitterung.

Durch den Verockerungsprozess kommt es zur verstärkten Trübung des Gewässers. Die Menge des eindringenden Sonnenlichtes verringert sich und die Lebensbedingungen für sedimentbewohnende Organismen nehmen durch die Eisenablagerungen ab. Als Folgeerscheinung reduziert sich die Biodiversität im Gewässer.

Aufgrund der räumlichen Ausbildung der Verockerung bieten sich Methoden der Fernerkundung zur flächenhaften Ableitung der Eisenkonzentration an. Zu diesem Zweck wurden Anfang September 2018 hyperspektrale Fernerkundungsdaten mit dem Gyrocopter der Hochschule Anhalt (HSA) und zeitgleich limnische In-situ Daten in Zusammenarbeit mit dem Leibniz Institut für Gewässerökologie und Binnenfischerei (IGB) erhoben. Wegen ihres großen Einflusses auf die Gewässerfärbung wurden die Gewässerproben auf folgende limnische Parameter untersucht: Gesamteisen, Eisen(II), Eisen(III), gelöster organischer Kohlenstoff (engl. Dissolved organic Carbon (DOC)), Sulfat und Chlorophyll a (Chl-a)

2 Untersuchungsgebiet

Das Untersuchungsgebiet befindet sich südlich von Cottbus und erstreckt sich entlang des Flusses Spree von Spreewitz bis Neuhausen (siehe Abb. 1). In diesem Bereich weist die Spree eine starke Verockerung auf.

Der erhöhte Stoffeintrag in Folge des Grundwasserwiederanstieges führt zum:

- Anstieg der Sulfatkonzentration,
- Anstieg der Eisenkonzentration (Eisen(II), Eisen(III)),
- Senkung des pH-Wertes und
- Trübung des Gewässers durch Schlammbildung.

Nach UHLMANN et al. (2012) erfolgt ein hoher Eisen(II)-Eintrag durch das Grundwasser bei Ruhlmühle. Das Eisen(II)-Oxid fällt durch die Reaktion mit dem sauerstoffreichen Wasser und unter neutralen pH-Bedingungen als Eisen(III)-Hydroxid aus. Dies resultiert in einer ockerartigen Färbung der Spree. Durch die natürliche Selbstreinigung von Flüssen fällt das Eisen(III)-Hydroxid bei langsamen Fließgeschwindigkeiten aus und lagert sich an der Flusssole bzw. am Boden der Talsperre Spremberg ab. Dadurch ist flussabwärts der Talsperre Spremberg eine natürliche Färbung der Spree zu beobachten.



Abb. 1 Darstellung der Untersuchungsgebietes entlang der Spree von Neustadt bis Neuhausen. Eigene Darstellung; Datenquellen: BKG, OSM; Hintergrundkarte: OSM

3 Datenerhebung

Zur Erfassung der hyperspektralen Fernerkundungsdaten wurde der Gyrocopter der HSA genutzt. Für die Befliegung wurden als optischen Sensoren eine NIKON D800e, ein Hyperspektralsensor HySpex VNIR 1600 sowie ein HySpex SWIR 384 eingesetzt. Ein Trägheitsnavigationssystem (INS) diente zur hochgenauen Erfassung der Lage- und Positionsdaten. Die Sensorik wird automatisch über die Flugmanagementsoftware (FMS) Aviatrix ausgelöst. Aus den RGB-Bildern wurde ein Near True-Orthophoto mit 6 cm Bodenpixelauflösung (engl. ground sampling distance (GSD)) mittels Structure from Motion-Algorithmen berechnet. Dieses diente als Basis für die Bestimmung der Boresight-Winkel für die Prozessierung der hyperspektralen Daten. Die Hyperspektraldaten wurden mit den Softwares PARGE und ATCOR rektifiziert, georeferenziert und atmosphärenkorrigiert. Um möglichst kleine aber signifikante spektrale Unterschiede zu detektieren wurde die spektrale Auflösung des HySpex VNIR auf 7,4 nm Bandbreite gesetzt. Aus dieser Bandbreite resultiert, dass im Spektralbereich von 400 -1.000 nm 80 Spektralkanäle aufgezeichnet werden können. Bei einer Flughöhe von 700 m beträgt die GSD 20 cm. Die Daten des HySpex SWIR besitzen eine spektrale Auslösung von 5,45 nm bei 288 Kanälen im Wellenlängenbereich von 1.000 - 2.500 nm. Die GSD des HySpex SWIR ist geringer als die vom HySpex VNIR 1600 und liegt bei 60 cm für eine Flughöhe von 700 m. Die hohe räumliche Auflösung bietet große Vorteile gegenüber räumlich gering aufgelösten Satellitendaten bei der Ableitung von Wasserinhaltsstoffen über Binnengewässer.

Der zweite Teil bestand aus der Erhebung sowie Analyse der Gewässerparameter nach den Analyseverfahren in Tab. 2. Die Gewässerproben wurden mittels eines 30 cm bzw. 50 cm langen LIMNOS Schöpfer erhoben. Dadurch besteht jede Probe aus einer Mischung der oberen Wassersäule. Im Labor der HSA wurden die Gewässerparameter: Gesamteisengehalt, Eisen(II), Eisen(III), Anionen, TOC und DOC bestimmt. Die Chl-a-Analyse erfolgte im Labor des IGB. Die Bestimmung des Gesamteisengehalts wurde nach DIN 38406-1:1983-05 durchgeführt. Dazu wurden 200 ml der Mischprobe unfiltriert mit 2 ml 20% Schwefelsäure vor Ort stabilisiert und im Labor analysiert. Die Analyse zur Eisen(II)-Bestimmung erfolgte ebenfalls nach DIN 38406-1:1983-05. Dazu wurden 280 ml der Mischprobe mit 2,8 ml 20% Schwefelsäure vor Ort stabilisiert und luftblasenfrei in einer Glasflasche verschlossen. Die Bestimmung der Eisen(III)-Konzentration ergibt sich durch die Differenz aus dem Gesamteisengehalt und der Eisen(III)-Konzentration. Zur Bestimmung der Anionen nach DIN EN ISO 10304-1 und DOC nach DIN ISO 1484 wurde die Mischprobe vor Ort mittels eines Spritzenfilters mit 0,45 µm Porenweite filtriert und im Labor analysiert. Die Bestimmung von TOC erfolgte mittels unfiltrierten Proben gemäß DIN ISO 1484. Die Gewässerproben zur Bestimmung der Chl-a– Konzentration wurden vor Ort unfiltriert in 1L-Gefäßen dunkel und kühl aufbewahrt. Im Labor erfolgte die Analyse mittels HPLC gemäß MEHNERT et al. (2012) und SHATWELL et al. (2012).

Tab. 2: Analyseverfahren zur Bestimmung der chemischen Gewässerparameter mit einem großen Einfluss auf die Gewässerfärbung.

Beschreibung	Verfahren
Bestimmung der Gesamteisen- und Eisen(II)-Konzentration	DIN 38406-1:1983-05
Bestimmung der Eisen(III)-Konzentration aus der Differenz aus den Konzentrationen von Gesamteisen und Eisen(II)	Differenzberechnung
Bestimmung der Anionen	DIN EN ISO 10304-1
Bestimmung von DOC und TOC als nicht ausblasbarer organischer Kohlenstoff (NPOC)	DIN ISO 1484
Bestimmung der Chlorophyll a - Konzentration	HPLC-Verfahren nach (MEHNERT et al. 2012; SHATWELL et al. 2012)

4 Aktive Gewässerinhaltsstoffe in der Fernerkundung

Nach HEEGE (2000) wird die Farbe eines Gewässers durch die optischen Eigenschaften der gelösten sowie partikulären Bestandteile bestimmt. Folglich beschreiben alle aktiven Gewässerinhaltsstoffe in Summe die Reflektanz (RRS) des Gewässers durch das spezifische Absorptions- sowie Rückstreuungsverhalten je Wellenlänge. Der relevante Spektralbereich zur Ableitung der aktiven Inhaltsstoffe liegt zwischen 400 - 1000 nm. Zur Gruppe der optisch aktiven Gewässerinhaltsstoffe gehören: Wasser, Gelbstoffe (engl. colored dissolved organic matter (CDOM)), Phytoplankton und Schwebstoffe (engl. total suspended matter (TSM)) (KIRK 2010; OGASHAWARA et al. 2017). Wasser absorbiert einen Großteil des eindringenden Sonnenlichtes. Das Rückstreuverhalten von Wasser ist gering und nur im Wellenlängenbereich von 400 - 500 nm erhöht. Zur Gruppe des CDOM gehören die gelösten Substanzen im Gewässer. Diese Absorbieren das Licht hauptsächlich im Wellenlängenbereich kleiner 500 nm (ASMALA et al. 2012; GITELSON 1992; KOPAČKOVÁ & HLADÍKOVÁ 2014). Häufig werden dem CDOM Huminstoffe und DOC zugeordnet. Die meisten Modelle zur Ableitung von CDOM basieren auf Wellenlängen zwischen 400 und 500 nm (KOPAČKOVÁ & HLADÍKOVÁ 2014). WEYHENMEYER et al. (2014) stellten fest, dass gelöstes Eisen wie DOC ebenfalls in Korrelation mit der Absorption bei 420 nm steht. Nach ASMALA et al. (2012) ist der Einfluss gelösten Eisens auf das Absorptionsverhalten des Gewässers ab 520 nm deutlich sichtbar und bei kürzeren Wellenlängen

vernachlässigbar. Phytoplankton gehört in der Gewässerfernerkundung zu den lebenden partikulären Materialen. Diese Gruppe besteht aus zahlreichen verschiedenen freischwebenden Algengruppen. Chl-a ist ein Schlüsselmolekül der photosynthetisch aktiven Organismen. Daher wird es oft als Indikator der Biomasseproduktion herangezogen. Zu den spektralen Merkmalen von Chl-a gehören einerseits ein geringer Reflektanzbereich zwischen 400 und 500 nm (GITELSON 1992) und andererseits ein Reflektanzmaximum bei 580 nm (DEKKER 1993; FRAUENDORF 2002; GITELSON 1992). Die nicht lebenden partikulären Materialen im Gewässer sind dem TSM zuordenbar. Dieser Gruppe gehören beispielweise abgestorbenes partikuläres organisches Material, kleine Sandpartikel und ausgefallende Eisenflocken an (FRAUENDORF 2002). Gewässer mit hoher TSM-Konzentration besitzen ein erhöhtes Rückstreuungsverhalten, wodurch die zumessende Reflektanz größer ist als bei natürlichen Gewässern. Dies haben ULRICH et al. (2018) anhand von Simulationen mittels des WATER COLOR SIMULATER (WASI) im Wellenlängenbereich von 400 - 850 nm nachgewiesen. KNAEPS et al. (2012) haben anhand des Multispektrometers Ocean and Land Color Instrument (OLCI) am Bord des Sentinel 3 festgestellt, dass bei steigender TSM-Konzentration bis 402 mg/L zusätzliche Informationen des SWIR-Bereichs zu genauerer Modellierung beitragen. Gemäß HEEGE (2000) ist die Absorption elektromagnetischer Energie von Wasser auch bei hohen TSM-Konzentrationen ab 1.300 nm dominierend.

5 Eisen-Indizes

Zur Entwicklung von Eisen-Indizes wurden Ratios aus der Literatur herangezogen sowie eigene anhand von Spektralanalysen abgeleitet. ULRICH et al. (2016) konnten mittels Spektralindizes basierend auf Landsat 7 TM Eisen(III) im Rio Doce nach dem Bergbauunglück im November 2015 nachweisen. Zur Ableitung wurden die Kanäle TM₃ (660 nm) / TM₁ (485 nm) nach DURNING et al. (1998) genutzt. Aufgrund der hohen Trübung der Spree im Untersuchungsgebiet wurden zur Ableitung der Eisenkonzentration Indizes basierend auf sehr trüben Gewässern herangezogen. DOXARAN et al. (2002b) entwickelten empirische Modelle zur TSM-Bestimmung sehr trüber Gewässer mittels SPOT-HRV. Sie stellten fest, dass der Kanal XS₃ (840 nm) sowie die Ratios aus XS₃/XS₁ (545 nm) und XS₃/XS₂ (645 nm) eine sehr hohe Korrelation mit der TSM-Konzentration in der Mündung des Gironde aufweisen. KNAEPS et al. (2015) führten ebenfalls Spektralanalysen zur Ableitung von TSM im Gironde durch. Hierbei nutzten sie die Spektralbereiche SWIR I (1.000 – 1.200 nm) sowie SWIR II (1.200 – 1.300 nm) zur Ableitung der Indizes. Für die Analysen der Spree wurden die Spektralkanäle zu entsprechenden Mittelwerten zusammengefasst.

Des Weiteren wurden die Spektren im Wellenlängenbereich von 415 - 1.315 nm mit den Daten an den Messstellen verglichen. Von 400 - 500 nm ist die Reflektanz aufgrund von Absorptionen durch CDOM und Chl-a gering (GITELSON 1992). Im Spektralbereich von 550 - 700 nm sind bei den Messstellen mit einer Gesamteisenkonzentration über 2 mg/L die meisten Maxima von über 6 % zu finden. Die Absorption durch Wasser ist zwischen 700 und 1.000 nm deutlich erkennbar. Jedoch steigt bei den Messstellen mit mindestens 2 mg/L Gesamteisen die Reflektanz von 1.000 - 1.315 nm stetig an. In Abb. 2 sind beispielhaft fünf Spektren mit angegebener Gesamteisenkonzentration im Spektralbereich von 415 - 1.315 nm dargestellt. Der Sprung bei 1.000 nm ergibt sich aus den Ungenauigkeiten der Kalibration der beiden Hyperspektralsensoren HySpex VNIR und HySpex SWIR. Im Wellenlängenbereich um 1.000 nm überlappen sich



Abb. 2: Darstellung von fünf Spektren der Spree im Wellenlängenbereich von 415 bis 1.315 nm mit unterschiedlicher Gesamteisenkonzentration.

diese. Bei einem kalibrierten System können Unterschiede von bis zu 4 % vorkommen. Zu den Ratios der Literatur wurden basierend auf den Spektralverläufen weitere Ratios zur

Ableitung von Eisen hergeleitet. Aufgrund der starken Steigung bieten sich die Ratios zur Ableitung von Eisen hergeleitet. Aufgrund der starken Steigung bieten sich die Ratios von R_{RS}(580) / R_{RS}(455) sowie R_{RS}(650) / R_{RS}(455) an. Weiterhin wurde ein Ratio von DOXARAN et al. (2002a) auf den Spektralverlauf angepasst R_{RS}(790) / R_{RS}(650). Aufgrund der deutlichen Ausprägung wurde die Steigung zwischen R_{RS}(455) \rightarrow R_{RS}(580) sowie von R_{RS}(650) \rightarrow R_{RS}(790) berechnet. Des Weiteren wurde untersucht, ob bestimmte Wellenlängen mit den Gewässerinhaltstoffen der Spree im kausalen Zusammenhang stehen. Zu den gewählten Wellenlängen zählen: 417 nm, 455 nm, 520 nm, 580 nm, 650 nm, 790 nm, 1.100 nm und 1.250 nm.

6 Ergebnisse

In Tab. 3 ist ein Überblick über die erhobenen Gewässerparameter dargestellt. Im Mittel betrug die Konzentration des Gesamteisen 6,13 mg/L, von Eisen(II) 2,6 mg/L und von Eisen(III) 3,53 mg/L. Die höchsten Eisenkonzentrationen wurden bei der Mündung der Kleinen Spree in die Spree bei Spreewitz bestimmt. Die niedrigsten Eisenkonzentrationen befanden sich bei Neuhausen. Die Sulfatkonzentration liegt im Mittel bei 549 mg/L. Die Maximalkonzentration wurde in Höhe des Industriegebietes Schwarze Pumpe gemessen. Die niedrigsten Sulfatkonzentration wurden bei Spreewitz bestimmt. Die Gewässerproben bei Spreewitz zeigen die höchsten Konzentrationen von DOC mit 5 mg/L sowie TOC mit 5,90 mg/L auf. Die geringste gemessene DOC-Konzentration mit 2 mg/L befindet sich mittig zwischen Schwarze Pumpe und Trattendorf. Flussabwärts von Spremberg wurde die niedrigste TOC-Konzentration mit 3,60 mg/L bestimmt. Im Mittel betragen die Konzentrationen von DOC 3,13 mg/L und von TOC 4,50 mg/L im Untersuchsuchungsgebiet zum Tag der Beprobung. Die Chl-a-Konzentration nimmt im Untersuchungsgebiet von Spreewitz bis zur Talsperre stetig ab. Bei Spreewitz beträgt die Chl-a-Konzentration 3,53 µg/L und beim Vorstaubecken der Talsperre Spremberg

noch 1,61 μ g/L. Erst nach der Talsperre Spremberg bei Neuhausen stieg die Chl-a-Konzentration auf 18,66 μ g/L an.

	Gesamteisen [mg/L]	Eisen(II) [mg/L]	Eisen(III) [mg/L]	Sulfat [mg/L]	DOC [mg/L]	TOC [mg/L]	Chl-a [µg/L]
Mittelwert	6,13	2,60	3,53	549,25	3,13	4,50	3,90
Standardabweichung	2,94	2,01	1,09	50,58	0,74	0,67	4,53
Max	10,90	5,74	5,18	606,00	5,00	5,90	18,66
Min	0,45	0,01	0,43	471,00	2,00	3,60	1,61

Tab.	3: Übersicht über	die erhobenen	Gewässerparame	ter im	Untersuchund	saebiet

Die analysierten Gewässerproben zeigen deutlich, dass südlich der Talsperre Spremberg eine hohe Eisen- und Sulfatkonzentration vorliegt. Erst nördlich der Talsperre Spremberg beträgt die Gesamteisenkonzentration weniger als 1 mg/L, wodurch diese eine wichtige Schlüsselposition für den Rückhalt der Eisenfrachten besitzt (BILEK & KOCH 2012). Weiterhin steigt die Chl-a -Konzentration erst nach der Talsperre Spremberg deutlich an.

Zur genauen Analyse wurden die erhobenen limnischen Daten innerhalb der 2,5fachen Standardabweichung herangezogen. Dadurch wird der Einfluss von Extremwerten auf das Bestimmtheitsmaß (R²) ausgeschlossen. Folglich erfolgt eine Reduzierung der Probenanzahl auf n = 15. Der kausale Zusammenhang von Gesamteisen zu Eisen(II) beträgt R² = 0,98 und zu Eisen(III) R² = 0,84. Zwischen Eisen(II) und Eisen(III) beträgt die Korrelation R² = 0,71. Aufgrund des chemischen Zusammenhangs der drei Eisenkonzentrationen ist auch die Korrelation sehr stark. Die kausalen Zusammenhänge der TOC-Konzentration zum Gesamteisen mit R² = 0,63 und zum Eisen(II) mit R² = 0,7 sind stark ausgeprägt. Dagegen ist die Korrelationen von TOC zu Eisen(III) (R² = 0,37) und zu Sulfat (R² = 0,34) geringer. Die Sulfatkonzentration besitzt ebenfalls eine ausgeprägte Korrelation zum Gesamteisen (R² = 0,43) sowie zu Eisen(II) (R² = 0,71). Zwischen Sulfat und Eisen(III) beträgt die Korrelation R² = 0,32 sowie zu TOC R² = 0,34. Korrelationen von Chl-a zum Gesamteisen (R² = 0,5), zu Eisen(II) (R² = 0,54) sowie zu Eisen(III) (R² = 0,56) sind stark ausgeprägt.

Kausale Zusammenhänge zwischen den einzelnen Wellenlängen und den drei Eisentypen konnten nur bei $R_{RS}(650)$ und $R_{RS}(790)$ erkannt werden. Die anderen Gewässerparameter wiesen keine Korrelation $R^2 \ge 0.25$ auf. Zur Wellenlänge $R_{RS}(650)$ besitzt Eisen(II) die stärkste



Abb. 3: Darstellung der Korrelation zwischen der Wellenlänge R_{RS}(650) zu den Konzentrationen des Gesamteisen(links), Eisen(II) (mittig) und Eisen(III) (rechts)

Korrelation mit $R^2 = 0.43$. Eine logarithmische Korrelation weisen Gesamteisen mit $R^2 = 0.31$ und Eisen(III) mit $R^2 = 0.15$ zu R_{RS}(650) auf (s. Abb. 3)

In Abb. 4 sind die kausalen Zusammenhänge von $R_{RS}(790)$ zum Gesamteisen ($R^2 = 0,33$), zum Eisen(II) ($R^2 = 0,37$) und zur Eisen(III)-Konzentration ($R^2 = 0,19$) dargestellt.



Abb. 4 Darstellung der Korrelation zwischen der Wellenlänge R_{RS}(650) zu den Konzentrationen des Gesamteisen(links), Eisen(II) (mittig) und Eisen(III) (rechts)

Die Eisen(II)-Konzentration weist weiterhin eine starke exponentielle Korrelation mit $R^2 = 0,43$ zur Steigung $R_{RS}(455) \rightarrow R_{RS}(580)$ auf. Auch der Gesamteisengehalt zeigt einen starken kausalen Zusammenhang mit $R^2 = 0,33$ auf. Eisen(III) besitzt keine Korrelation mit $R^2 = 0,13$ zur Steigung $R_{RS}(455) \rightarrow R_{RS}(580)$ (s. Abb. 5).



Abb. 5: Darstellung der Korrelation zwischen der Steigung R_{RS}(455) → R_{RS}(580) zu den Konzentrationen des Gesamteisen(links), Eisen(II) (mittig) und Eisen(III) (rechts)

Zur Steigung $R_{RS}(650) \rightarrow R_{RS}(790)$ besitzt Eisen(II) die stärkste Korrelation mit $R^2 = 0,5$. Der exponentielle Zusammenhang zur Gesamteisenkonzentration beträgt $R^2 = 0,43$ und zur Chl-a – Konzentration $R^2 = 0,35$. Zur Eisen(III)-Konzentration liegt eine niedrige lineare Kor-



Abb. 6: Darstellung der Korrelation zwischen der Steigung R_{RS}(650) → R_{RS}(790) zu den Konzentrationen des Gesamteisen, Eisen(II), Eisen(III) und Chl-a

relation mit $R^2 = 0.2$ vor (s. Abb. 6).

Weiterhin konnten starke Korrelationen zwischen einzelne Ratios im VNIR und den drei Eisenkonzentrationen, TOC und Chl-a identifiziert werden.

In Abb. 7 sind die Korrelationen zwischen dem Ratio R_{RS}(580) / R_{RS}(455) nach Gl. 4 und den Konzentrationen zum Gesamteisen, Eisen(II), Eisen(III) und TOC dargestellt.
$$Index = \frac{R_{RS}(580) - R_{RS}(455)}{R_{RS}(580) + R_{RS}(455)}$$
Gl. 4

Die stärkste lineare Korrelation mit $R^2 = 0,44$ tritt beim Eisen(II) auf. Die Konzentrationen vom Gesamteisen ($R^2 = 0,4$) und TOC ($R^2 = 0,4$) besitzen ebenfalls einen starken kausalen Zusammenhang zum Ratio $R_{RS}(580) / R_{RS}(455)$. Die Korrelation zur Eisen(III)-Konzentration ($R^2 = 0,25$) sowie zu den übrigen Gewässerparametern ist gering ausgeprägt.



Abb. 7: Darstellung der Korrelation zwischen dem Ratio R_{RS}(580) / R_{RS}(455) zu den Konzentrationen des Gesamteisen, Eisen(II), Eisen(III) und TOC

Gemäß Gl. 5 wurde der Ratio R_{RS}(650) / R_{RS}(455) erstellt.

$$Index = \frac{R_{RS}(650) - R_{RS}(455)}{R_{RS}(650) + R_{RS}(455)}$$
Gl. 5

Mit $R^2 = 0,43$ zeigt Eisen(II) den stärksten linearen Zusammenhang zum Ratio $R_{RS}(650) / R_{RS}(455)$ gemäß Gl. 5 auf (s. Abb. 8). Die Korrelation zum Gesamteisen zeigt mit $R^2 = 0,4$ eine starke Ausprägung. Im Vergleich zur Gl. 4 ist der lineare Zusammenhang zwischen TOC und dem Ratio aus Gl. 5 geringer mit $R^2 = 0,36$. Dagegen steigt die Korrelation des Eisen(III) auf $R^2 = 0,27$ an.



Abb. 8: Darstellung der Korrelation zwischen dem Ratio R_{RS}(650) / R_{RS}(455) zu den Konzentrationen des Gesamteisen, Eisen(II), Eisen(III) und TOC

Weiterhin zeigen die die Konzentration von Gesamteisen ($R^2 = 0,37$), Eisen(II) ($R^2 = 0,40$) und TOC ($R^2 = 0,40$) große lineare Zusammenhänge zum Ratio $R_{RS}(650) / R_{RS}(480)$ gemäß Gl. 6 auf. Die Korrelation vom Eisen(III) zum Ratio ist gering mit $R^2 = 0,21$ ausgeprägt (s. Abb. 9).

$$Index = \frac{R_{RS}(650) - R_{RS}(480)}{R_{RS}(650) + R_{RS}(480)}$$
Gl. 6



Abb. 9: Darstellung der Korrelation zwischen dem Ratio R_{RS}(650) / R_{RS}(480) zu den Konzentrationen des Gesamteisen, Eisen(II), Eisen(III) und TOC

Außerdem weist die Eisen(II)-Konzentration eine Korrelation ($R^2 = 0,27$) zum Ratio $R_{RS}(850) / R_{RS}(550)$ in Anlehnung an DOXARAN et al. (2002a) in Gl. 7 auf.

$$Index = \frac{R_{RS}(850) - R_{RS}(550)}{R_{RS}(850) + R_{RS}(550)}$$
Gl. 7

Stärker ist der exponentielle Zusammenhang zum Chl-a mit $R^2 = 0,36$ ausgeprägt. Zum Gesamteisen ($R^2 = 0,23$) besteht eine geringe und zum Eisen(III) ($R^2 = 0,12$) keine Korrelation.



Abb. 10: Darstellung der Korrelation zwischen dem Ratio R_{RS}(850) / R_{RS}(550) zu den Konzentrationen des Gesamteisen, Eisen(II), Eisen(III) und TOC.

Aufgrund der starken Korrelationen wurden die Ratios nach Gl. 4 und 5 sowie die Steigung $R_{RS}(650) \rightarrow R_{RS}(790)$ für die Mündung der Kleinen Spree bei Spreewitz berechnet. In diesem Bereich wurden die höchsten Eisenkonzentrationen gemessen. In Abb. 11 ist in (a) die Mündungsstelle durch das True-Orthophoto dargestellt. In (b) wurde der Index nach Gl. 4 und in (c) der Index nach Gl. 5 berechnet. Entsprechend der Korrelationen liegt der Wertebereich der Ratios in Abb. 11 (b) und (c) zwischen 0,4 (blau) und 0,6 (rot), wobei 0,4 auf eine niedrige und 0.6 auf eine hohe Eisenkonzentration hinweist. Weiterhin ist die Steigung $R_{RS}(650) \rightarrow R_{RS}(790)$ in Abb. 11 (d) dargestellt. In Analogie zu den Ratios liegt der verwendete Wertebereich zwischen -2 (blau) und -9 (rot), wobei -2 für eine niedrige und -9 für eine hohe Konzentration steht.

Bei allen drei berechneten Ableitung ist deutlich die Strömungsänderung durch die Kleine Spree zu erkennen. Dabei kommt es zur Durchmüschung beider Flüsse. Flussabwärts der Mündung befindet sich ein kleines Gefälle, welches zur weiteren Durchmischung sowie zur Beflüftung der Spree beiträgt. Bei diesem Vorgang entsteht unteranderem auch schaum, welcher bei den berechneten Ratios durch einen sehr hohen Index gekennzeichnet wird. Bei der berechneten Steigungs ist dieser kaum zu erkennen. Weiterhin weisen die Indizes eine starke Beeiflussung durch die Reflexionen von Vegetation auf. Nahe der Beumkronen sind hohe Eisenkonzentration dargestellt, welche bei der Steigungsberechnung nicht erkennbar sind. Weitere Analysen in einem Teilgebiet der Untersuchungsregion mit n = 11 zeigten, dass die gemessene Sichttiefe [cm] mitteles Secchi-Scheibe stark mit dem Index nach Gl. 4 $(R^2 = 0,54)$ und Gl. 5 ($R^2 = 0,49$) korreliert. Weiterhin besteht zwischen der Sichttiefe und dem Gesamteisen ($R^2 = 0,66$) sowie zum Eisen(III) ($R^2 = 0,53$) ein starker linearer Zusammenhang. Die gemessene Sichttiefe weist keine Korrelation zur Steigung R_{RS}(650) \rightarrow R_{RS}(790) auf.



Abb. 11: Berechnung dreier Eisen-Ableitungen an der Mündung der Kleinen Spree (Spreewitz). In (a) ist das RGB-True-Orthophoto dargestellt. Weiterhin sind die Indizes nach GI. 4 in (b) und nach GI. 5 in (c) abgebildet. Die Steigung R_{RS}(650) → R_{RS}(790) ist in (d) dargestellt

In Abb. 12 (a) ist die Spree anhand des Hyperspekatrlbildes bei Spremberg dargestellt. Bei Spremberg betragen die Konzntrationen von Gesamteisen 3,52 mg/L, von Eisen(II) 0,08 mg/L ,von Eisen(III) 3,44 mg/L, von TOC 3,7 mg/L und von Chl-a 2 µg/L. Im Vergleich zur Mündung der Kleinen Spree haben alle Konzentration bis auf die Sulfatkonzentration abgenommen. Die Sulfatkonzentration stieg auf 564 mg/L an. Der niedrigere Gesamteisengehalt ist anhand der Indizes nach Gl. 4 (b) und 5 (c) deutlich zu erkennen. Nur die Steigung $R_{RS}(650) \rightarrow R_{RS}(790)$ in (d) zeigt eine Erhöhung an. Dies könnte an einem veränderten Verhältnis des Gesamteisen zum Chl-a resultieren, da auch Chl-a eine starke Korrelation zur Steigung $R_{RS}(650) \rightarrow R_{RS}(790)$ mit $R^2 = 0,25$ besitzt.



Abb. 12: Berechnung dreier Eisen-Ableitungen bei Spremberg. In (a) ist das Hyperspektralbild (RGB) dargestellt. Weiterhin sind die Indizes nach GI. 4 in (b) und nach GI. 5 in (c) abgebildet. Die Steigung R_{RS}(650) → R_{RS}(790) ist in (d) dargestellt.

7 Fazit und Ausblick

Die Folgen des offenen Tagebaus sind am Beispiel der Spree unmittelbar ersichtlich. In Folge der Grundwasseranhebung kam es zum erhöhten Eiseneintrag, wodurch es zur Verockerung im Bereich zwischen Spreewitz und der Talsperre Spremberg kam. Zur Ableitung der Konzentrationen des Gesamteisens, Eisen(II) und Eisen(III) wurden die erhobenen limnischen Daten mit den Spektraldaten im VNIR- und SWIR-Wellenlängenbereich korreliert. Dazu wurden für die Spektralbereich von 400 – 1.300 nm Ratios, Steigungen und spezifische Wellenlängen der Literatur entnommen sowie anhand des Spektralverlaufes abgeleitet. Zur Ableitung der Eisenkonzentrationen wurden die Ratios $R_{RS}(580) / R_{RS}(455)$ (Gl. 4) und $R_{RS}(650) / R_{RS}(455)$ (Gl. 5) sowie die Steigung $R_{RS}(650) \rightarrow R_{RS}(790)$ herangezogen. Zu allen drei Ableitungen zeigten die Konzentrationen des Gesamteisen und Eisen(II) Korrelationen $R^2 > 0,4$ auf. Der kausale Zusammenhang zur Eisen(III)-Konzentration betrag $0,2 < R^2 < 0,3$.

Weiterhin beträgt die gemessene Sichttiefe mittels Secchi-Scheibe starke Korrelationen zu den Ratios $R_{RS}(580) / R_{RS}(455)$ nach Gl. 4 ($R^2 = 0,54$) und $R_{RS}(650) / R_{RS}(455)$ nach Gl. 5 ($R^2 = 0,49$) sowie zum Gesamteisengehalt ($R^2 = 0,66$) und Eisen(III)-Konzentration ($R^2 = 0,53$). Folglich ist mittels der Indizes der Ratios nach Gl. 4 und 5 die Veränderung bezüglich des Gesamteisengehalts erkennbar. Aufgrund des großen Einflusses von Chl-a auf die Steigung $R_{RS}(650) \rightarrow R_{RS}(790)$, war bei Spremberg keine Reduzierung der Eisenkonzentration erkennbar. Eine deutliche Ableitung von Eisen(II) und Eisen(III) konnte nicht ermittelt werden. Da es sich bei der Untersuchung um die Ableitung von Indizes handelt können keine Aussagen über die Konzentration der Parameter getroffen werden. Basierend auf den erlangten Erkenntnissen werden zukünftig komplexe Modellierungsalgorithmen zur Ableitung von Konzentrationsangabe anhand von Fernerkundungsdaten anstelle von Indizes herangezogen.

8 Literatur

- ASMALA, E. STEDMON, C. A. & THOMAS, D. N., 2012: Linking CDOM spectral absorption to dissolved organic carbon concentrations and loadings in boreal estuaries. Estuarine, Coastal and Shelf Science, **111**,107-117.
- BILEK, F. & KOCH, C., 2012: Eisenretention in der Talsperre Spremberg. Dresden.
- BILEK, F., MORITZ, F. & ALBINUS, S., 2016: Iron-Hydroxide-Removal from Mining Affected Rivers. Mining Meets Water – Conflicts and Solutions, Leipzig, 151-158.
- DEKKER, A. G., 1993: Detection of optical water quality parameters for eutrophic waters by high resolution remote sensing. Amsterdam.
- DOXARAN, D. FROIDEFOND, J. M. & CASTAING, P., 2002a: A reflectance band ratio used to estimate suspended matter concentrations in sediment-dominated coastal waters. International Journal of Remote Sensing, **23**(23), 5079-5085.
- DOXARAN, D., FROIDEFOND, J. M., LAVENDER, S. & CASTAING, P., 2002b: Spectral signature of highly turbid waters: Application with SPOT data to quantify suspended particulate matter concentrations. Remote Sensing of Environment, **81**(1), 149-161.
- DURNING, W. P., POLIS, S. R., FROST, E. G. & KAISER, J. V., 1998: Integrated Use of Remote Sensing and GIS for Mineral Exploration Final Report, 25.
- EDWARDS, K. J., GOEBEL, B. M., RODGERS, T. M., SCHRENK, M. O., GIHRING, T. M., CARDONA, M. M., HU, B., MCGUIRE, M. M., HAMERS, R. J. & PACE, N. R., 1999: Geomicrobiology of Pyrite (FeS2) Dissolution: Case Study at Iron Mountain, California.

Geomicrobiology Journal, 16(2),155-179.

- FRAUENDORF, J., 2002: Entwicklung und Anwendung von Fernerkundungsmethoden zur Ableitung von Wasserqualitätsparametern verschiedener Restseen des Braunkohlentagebaus in Mitteldeutschland. Dissertation, Mathematisch-Naturwissenschaftlich-Technischen Fakultät der Universität Halle-Wittenberg, Halle (Saale). https://sundoc.bibliothek.uni-halle.de/diss-online/02/02H226/of index.htm
- GITELSON, A., 1992: The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. International Journal of Remote Sensing, **13**(17), 3367-3373.
- GLEISNER, M. & HERBERT, R. B., 2002: Sulfide mineral oxidation in freshly processed tailings: batch experiments. Journal of Geochemical Exploration, **76**(3), 139-153.
- HEEGE, T., 2000: Flugzeuggestützte Fernerkundung von Wasserinhaltsstoffen am Bodensee.
- HÜTTL, R. F., 1998: Ecology of post strip-mining landscapes in Lusatia, Germany. Environmental Science & Policy, 1(2), 129-135.
- KIRK, J. T. O., 2010: Light and Photosynthesis in Aquatic Ecosystems. Canberra.
- KNAEPS, E., DOGLIOTTI, A. I., RAYMAEKERS, D., RUDDICK, K. & STERCKX, S., 2012: In situ evidence of non-zero reflectance in the OLCI 1020nm band for a turbid estuary. Remote Sensing of Environment, **120**, 133-144.
- KNAEPS, E., RUDDICK, K. G., DOXARAN, D., DOGLIOTTI, A. I., NECHAD, B., RAYMAEKERS, D.
 & STERCKX, S., 2015: A SWIR based algorithm to retrieve total suspended matter in extremely turbid waters. Remote Sensing of Environment, 168, 66-79.
- KOPAČKOVÁ, V. & HLADÍKOVÁ, L., 2014: Applying Spectral Unmixing to Determine Surface Water Parameters in a Mining Environment. Remote Sensing, **6**(11), 11204-11224.
- LUTHER, G. W. ,1987: Pyrite oxidation and reduction: Molecular orbital theory considerations. Geochimica et Cosmochimica Acta, **51**(12), 3193-3199.
- MEHNERT, G., RÜCKER, J., NICKLISCH, A., LEUNERT, F. & WIEDNER, C., 2012: Effects of thermal acclimation and photoacclimation on lipo-philic pigments in an invasive and a native cyanobacterium of temperate regions. European Journal of Phycology, 47(2),182-192.
- MOSES, C. O., NORDSTROM, D. K., HERMAN, J. S. & MILLS, A. L., 1987: Aqueous pyrite oxidation by dissolved oxygen and by ferric iron. Geochimica et Cosmochimica Acta **51**(6),1561-1571.
- OGASHAWARA, I., MISHRA, D. R. & GITELSON, A. A., 2017: Remote Sensing of Inland Waters: Background and Current State-of-the-Art. In: MISHRA, D. R., I. OGASHAWARA und A. A. GITELSON (Hrsg.), Bio-optical Modeling and Remote Sensing of Inland Waters, 1-20.
- 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. In:Conferencia invitada: Sánchez España, Macla 10, 34-43.
- SCHULTZE, M., POKRANDT, K. H. & HILLE, W., 2010: Pit lakes of the Central German lignite mining district: Creation, morphometry and water quality aspects. Limnologica 40(2), 148-155.
- SHATWELL, T., NICKLISCH, A. & KÖHLER, J., 2012: Temperature and photoperiod effects on phytoplankton growing under simulated mixed layer light fluctuations. Limnology and Oceanography, **57**(2), 541-553.
- SINGER, P. C. & STUMM, W., 1970: Acidic Mine Drainage: The Rate-Determining Step. Science, 167(3921), 1121-1123.

- UHLMANN, W., THEISS, W., NESTLER, S., ZIMMERMANN, K. & CLAUS, T., 2012: Weiterführende Untersuchungen zu den hydrochemischen und ökologischen Auswirkungen der Exfiltration von eisenhaltigem, saurem Grundwasser in die Kleine Spree und in die Spree. Projektphase 2: Präzisierung der Ursachen und Quellstärken für die hohe Eisenbelastung. Dresden.
- ULRICH, C., BANNEHR, L. & LAUSCH, A., 2016: Ableitung von Eisen (II, III) oxid in Fließgewässern mittels Multispektraldaten. 36. Wissenschaftlich-Technische Jahrestagung der DGPF, Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 25, 34-43.
- ULRICH, C., SCHULZE, M. & BANNEHR, L., 2018: Spektrale In -Situ Messungen als Basis für ein flächendeckendes Monitoring verockerter Gewässer. Forum für Hydrologie und Wasserbewirtschaftung, Dresden, 39(18), 27-37.
- WEYHENMEYER, G. A., PRAIRIE, Y. T. & TRANVIK, L. J., 2014: Browning of boreal freshwaters coupled to carbon-iron interactions along the aquatic continuum. PLoS ONE, 9(2).

Dokumentation des Gletscherrückgangs am Gössnitzkees für den Zeitraum 1982-2018 – eine Gletschergeschichte mit Ablaufdatum

VIKTOR KAUFMANN¹ & WOLFGANG SULZER²

Zusammenfassung: Gletscher dienen in der Hochgebirgsforschung als untrügliche Klimaindikatoren, da sie auf geänderte Umweltbedingungen relativ schnell mit Längen-, Flächenund Volumenänderungen reagieren. 1982 wurde das Gössnitzkees (Schobergruppe, Hohe Tauern, Kärnten) erstmals in das Gletschermessnetz des Österreichischen Alpenvereins (ÖAV) eingebunden. Dies war gleichzeitig auch der Ausgangspunkt für weitere Detailuntersuchungen, welche insbesondere die Erweiterung des Methodenspektrums der geometrischen Gletscherbeobachtung zum Ziel hatten. In diesem Beitrag werden nicht nur die im Laufe der Zeit eingesetzten Messmethoden dargestellt, sondern auch deren Ergebnisse aus glaziologischer Sicht vergleichend präsentiert. Seit 1850, dem neuzeitlichen Gletscherhochstand, hat sich das Gössnitzkees ohne nennenswerte Unterbrechungen kontinuierlich in seiner flächenmäßigen Ausdehnung verkleinert. Aus derzeitiger Sicht ist davon auszugehen, dass das Gössnitzkees in den kommenden Jahren fast völlig abschmelzen wird, wiewohl seine Schuttbedeckung den Abschmelzprozess noch etwas verlangsamen wird können. Das 2018 mit einer UAV-gestützten Luftbildbefliegung festgestellte Einsinken/Abrutschen des Eiskörpers im zentralen Teil des Gletschers lässt vermuten, dass der Gletscher in seinem Endstadium in mehrere Lappen (Toteiskörper) zerfallen wird.

1 Einleitung

Seit dem Ende des 19. Jahrhunderts ist ein Ansteigen der mittleren jährlichen Lufttemperatur im Alpenraum von rund 2°C zu beobachten (ZAMG 2018). Diese geänderten Klimaverhältnisse führen nicht nur zum Abschmelzen der Gletscher, sondern sind auch verantwortlich für das Tauen des Permafrosts, das höhenzonale Verschieben von Vegetationsgrenzen und das Einwandern von wärmeliebenden Tieren in höhere Regionen. Die Gletscheränderungen (Länge, Fläche, Volumen/Masse) spiegeln im hohen Maße das regionale Klima wider und sind daher aus klimatologischer Sicht besonders interessant. Aus diesem Grunde werden Gletscher im Speziellen schon seit über 150 Jahren beobachtet und auch vermessen (MAYER 2010).

Das Gössnitzkees ist mit 49.83 ha (Stand 2009, FISCHER et al. 2015b) der größte Gletscher in der Schobergruppe, Hohe Tauern, Zentralalpen, Österreich und befindet sich in der Kernzone des Nationalparks Hohe Tauern, Kärnten (Abb. 1). Der Gletscher ist ein typischer Kargletscher und wird hauptsächlich von Lawinen aus den umgebenden Couloirs ernährt. Ein größeres Nährgebiet für Schnee und Eis fehlt. Charakteristisch für den Gletscher und auch für die gesamte Schober-

¹ Technische Universität Graz, Institut für Geodäsie, Arbeitsgruppe für Fernerkundung und Photogrammetrie, Steyrergasse 30, A-8010 Graz, E-Mail: viktor.kaufmann@tugraz.at

² Karl-Franzens-Universität Graz, Institut für Geographie und Raumforschung, Arbeitsgruppe für Geospatial Technologies, Heinrichstraße 36, A-8010 Graz, E-Mail: wolfgang.sulzer@uni-graz.at



Abb. 1: Lage des Gössnitzkeeses

gruppe ist die starke Schuttbedeckung (LANG & LIEB 1993). Abbildung 2 zeigt den zentralen Bereich des Gletschers zu unterschiedlichen Epochen (1863, 1982 und 2018).

1982 haben Gerhard K. Lieb und Viktor Kaufmann am Gössnitzkees im Rahmen des Gletschermessdienstes des ÖAV (FI-SCHER et al. 2018) mit einfachen Gletscherlängenänderungsmessungen begonnen, um die Veränderung des Gletschers zu dokumentieren. Diese Messungen wurden vermessungs-technisch zwischenzeitlich (1989-2002) durch Wolfgang Sulzer durchgeführt. Derzeit liegt die Verantwortung bei Michael Krobath (LIEB & KELLERER-PIRKLBAUER 2018). Unabhängig davon liefen bzw. laufen

noch weiterführende Detailstudien im universitären Bereich mit dem Fokus auf geometrisches Gletschermonitoring, um mit einer höheren raumzeitlichen Auflösung den offensichtlichen Gletscherrückgang zu erfassen. Schwerpunktmäßig werden nun in diesem Beitrag nicht nur die Forschungsergebnisse der letzten Jahre, sondern auch besonders aktuelle aus dem Jahre 2018 präsentiert.



Abb. 2: Photographische Aufnahmen des Gössnitzkeeses. Links: 1863, Photo Gustav Jägermayer, Reproduktion © Albertina, Wien, <u>https://www.albertina.at/</u>; Mitte: 15.9.1982, Photo V. Kaufmann; rechts: 18.8.2018, Photo V. Kaufmann



Abb. 3: Flächenänderung der Vergletscherung 1850-2012

2 Methoden

Das Methodenspektrum in der Gletschervermessung ist mannigfaltig (KÄÄB 2005). Am Beispiel des Gössnitzkeeses werden von den Autoren aus praktischen Gründen, wie z.B. Kosten, Machbarkeit und Genauigkeitsanforderungen, die im Folgenden angeführten fünf Methoden favorisiert und in den Mittelpunkt der Betrachtungen gestellt. Andere Verfahren, wie z.B. terrestrisches Laserscanning (PAAR et al. 2012), optische Satellitenbildaufnahme (GANDER et al. 2003) oder satellitengestützte SAR-Interferometrie (KAUFMANN et al. 2009) wurden ebenfalls am Beispiel des Gössnitzkeeses untersucht, werden aber in diesem Beitrag bewusst ausgeblendet.

2.1 Geodätische Vermessung

Zwischen 1982 und 2002 wurde die Lage der Gletscherstirn alljährlich durch Trilateration (anfangs Maßbandmessungen, später Distometermessungen) in einem lokalen Koordinatensystem (Ausgangspunkte A82, B82; siehe Abb. 4) eingemessen. Durch den starken Gletscherrückgang

musste jedoch der Bezugsrahmen in Richtung des rückschreitenden Gletschertors nachgeführt werden. Der seit 1987/88 vorhandene Gletscherrandsee erschwerte die Vermessungsarbeiten erheblich. Hilfsdreiecke mussten alljährlich neu vermessen werden, um den Gletscherrand im Stirnbereich durch Einzelpunktmessung erfassen zu können. 1996 erfolgte die Erweiterung/Umstellung auf Tachymetrie mit einer Totalstation. Das lokale Festpunktfeld wurde laufend um weitere Punkte (A84, B84, A00, B00 und C00; siehe Abb. 4) ergänzt und gleichzeitig in das Koordinatensystem der Landesvermessung übergeführt. Die alljährliche Tachymetrie erfolgte durch Polarpunktaufnahme mit bekanntem Anschluss vom Standpunkt H1. Eingemessen wurden der Gletscherrand im Stirnbereich, zwei Gletscherlängsprofile (Profile 1 und 2), acht Sonderpunkte (farbmarkierte Felsecken am Gletscher) und die Uferlinie des Gletscherrandsees (Abb. 4). Aus den Profilmessungen kann die höhenzonale Eisdickenänderung und über die Sonderpunkte die horizontale Fließgeschwindigkeit und ebenfalls auch die Eisdickenänderung abgeleitet werden. Seit 2014 erfolgt die alljährliche Vermessung mittels RTK-GNSS-Technologie/RTK-Vermessung (Real Time Kinematic). Als Referenzpunkt (Basispunkt) dient der Festpunkt B84.

2.2 Klassische Luftbildvermessung

Die flächenhafte Erfassung von Gletschern ist nur bildgestützt, vorzugsweise durch Luftbildvermessung, ökonomisch durchführbar. Auswerteergebnisse sind digitale Geländemodelle (DGM) und Gletscherflächen, welche in einem GIS nach glaziologischen Gesichtspunkten analysiert werden können. Die Zeitintervalle der Luftbildaufnahmen sind im Allgemeinen mehrjährig: Für den österreichischen Alpenbogen wurden bis jetzt insgesamt drei vollständige Gletscherinventuren (1969, 1989, 2006) durchgeführt (LAMBRECHT & KUHN 2007; FISCHER et al. 2015a). Bei der jüngsten Inventur wurden auch ALS-Daten unterstützend verwendet (ABER-MANN et al. 2010). Die Gletscherinventare stehen als GIS-Datensätze für weiterführende Analysen zur Verfügung (FISCHER et al. 2015b).

Im Rahmen der eigenen Gletscherforschung in der Schobergruppe wurde in einem vom Nationalpark Hohe Tauern geförderten Projekt die Gletschergeschichte des Gössnitzkeeses und des benachbarten Hornkeeses für den Zeitraum 1850 (max. neuzeitliche Vergletscherung) bis 1997 erfasst. Die drei ältesten Gletscherstände wurden aus Geländebefunden (Seiten- und Endmoränen) und vorhandenem Kartenmaterial rekonstruiert (KAUFMANN & LIEB 2002). Exakte Luftbildauswertungen begannen mit der Epoche 1954 (KAUFMANN & PLÖSCH 2000). Die Datenerfassung erfolgte mit einem analytischen Plotter (DSR-1 von Kern). Im Zuge von weiteren Teilprojekten (KAUFMANN & LADSTÄDTER 2008a) bzw. auch von studentischen Arbeiten (MODER 2008; KLINGHAN 2017) wurden weitere Epochen des Zeitraumes 2002-2012, nunmehr digitalphotogrammetrisch (ImageStation von Z/I Imaging), ausgewertet. Bildmaterial aus dem Jahr 2015 ist noch nicht ausgewertet. Von allen genannten Epochen liegen in einem GIS georeferenzierte Gletscherflächen (-masken) und digitale Geländemodelle (2.5m-Raster) vor. Schwierigkeiten bei der Luftbildauswertung ergaben sich teilweise bei der Oberflächenerfassung überstrahlter, texturloser Flächen bzw. bei der Abgrenzung der Gletscherflächen im stark schuttbedeckten Bereich. Mit zunehmendem Gletscherrückgang und einhergehender vermehrter Schuttbedeckung ist die Oberflächenerfassung nicht mehr kritisch, wohingegen die Gletscherabgrenzung immer schwieriger wird. Für das Gössnitzkees wurde folgende Vorgangsweise gewählt: Bei fortwährendem Gletscherrückgang kann der Gletscherrand älterer Epochen durch Verschnitt des jeweiligen DGM mit einem zeitlich jüngeren Geländemodell rekonstruiert werden. Für die interaktive Korrektur der morphologisch unklaren Bereiche eignen sich besonders farbige Überlagerungen der Höhenschichtlinienbilder. Der Gletscherstand 2012 kann also erst mit der Auswertung 2015 gesichert dokumentiert werden (vgl. dazu Abschnitt 3.4). Eine weitere Methode den vorerst visuell nach morphologischen Gesichtspunkten erfassten Gletscherrand zu kontrollieren bzw. zu verbessern, besteht darin, die Fließgeschwindigkeit (Oberflächenbewegung) für die Abgrenzung der Gletscherfläche heranzuziehen. Für die Berechnung von 2D-Bewegungsvektoren wurde ein Matlab-Tool entwickelt, welches multi-temporale Orthophotos miteinander vergleicht. Der Algorithmus von Richard Ladstädter (KAUFMANN & LADSTÄDTER 2004) funktioniert mit Stereoorthophotos und ist in der Lage, aus sogenannten Quasi-Orthophotos nicht nur exakte 3D-Verschiebungsvektoren zu berechnen, sondern auch die multi-temporalen DGM zu verbessern. Im Rahmen der selbstständigen Arbeiten wurde nur Datenmaterial des Luftbildarchivs des BEV (analog, digital) verwendet. Für ausgewählte Epochen waren die photogrammetrischen Orientierungselemente bekannt. Das BEV ist bemüht, das Bundesgebiet Österreich in einem Zeitintervall von drei Jahren vollständig neu zu erfassen, was auch in den vergangenen Jahren zumindest für den Bereich des Gössnitzkeeses gelang. Für die dazwischen liegenden Jahre wird von den Autoren die Luftbildaufnahme mittels kostengünstiger UAV (Unmanned Aerial Vehicles) vorgeschlagen. Vielversprechende Pilotstudien wurden am nahegelegenen Tschadinhorn Blockgletscher, einem Kriechphänomen des diskontinuierlichen Hochgebirgspermafrosts (KAUFMANN et al. 2018), bzw. am Pasterzengletscher in der Glocknergruppe (SEIER et al. 2017) durchgeführt.

2.3 UAV-gestützte Luftbildvermessung

Im Sommer 2018 erfolgte im Rahmen der alljährlichen Gletschervermessung eine UAVgestützte Luftbildaufnahme mit einer DJI Pantom 4 (DJI 2018). An notwendigen Genehmigungen für den diesbezüglichen Bildflug lagen (1) das Einverständnis der Grundeigentümer, (2) die behördliche Bewilligung durch die AustroControl und (3) die Ausnahmegenehmigung zur Durchführung eines Drohnenfluges in der Kernzone des Nationalparks Hohe Tauern vor. Die Bilddatengewinnung erfolgte über zwei Wegpunktflüge (nominelle Flughöhe über Grund: 100 m, Längs- und Querüberdeckung: 82% bzw. 63%). Das Aufnahmegebiet umfasst den zentralen Teil des Gletschers oberhalb der Gletscherstirn und den vorgelagerten Gletscherrandsee (Abb. 4). Zu jeder der 136 Flugaufnahmen wurden die GPS-Koordinaten aufgezeichnet. Für die genaue Georeferenzierung des Bildverbandes wurden vor der Befliegung elf große und für die Validierung acht weitere, kleinere Passpunktsignale gleichmäßig im Befliegungsgebiet ausgelegt und mittels RTK-GNSS eingemessen. Die Datenauswertung erfolgte mit der kommerziellen Software Agisoft PhotoScan (AGISOFT 2018). Die verwendete Kamera wurde mit den Parametern f, cx, cy, K1, K2, P1, P2, B1 und B2 im Zuge einer Simultankalibrierung modelliert. Mit Hilfe der kleinen Passpunktsignale und der teilweise luftsichtbaren sonstigen geodätischen Punkte wurde ein unabhängiger RMS-Fehler für die Genauigkeitsbewertung berechnet. Detailuntersuchungen bezogen sich u.a. auf die GPS-gestützte Bildtriangulation ohne bzw. mit nur einem Passpunkt (vgl. Tab. 1). Weiters wurde die Höhengenauigkeit des ermittelten DGM entlang der Profillinie 1 durch Vergleich der interpolierten Höhenwerte mit den diesbezüglichen geodätischen Messungen bestimmt. Eine weitere Validierungsmöglichkeit ergab sich durch Differenzbildung des ermittelten DGM (nun im 1m-Raster) mit den ALS-basierten DGM aus 2012. Als Ergebnis der Datenauswertung wurde ein DGM mit 10 cm Rasterweite und ein Orthophotomosaik mit 5 cm Bodenauflösung berechnet (vgl. Abb. 4). Weitere glaziologisch relevante Informationen, wie z.B. der sichtbare Gletscherrand, morphologische Strukturlinien (Abrisse) und die Uferlinie des Gletscherrandsees, wurden interaktiv im Orthophotomosaik bzw. im 3D-Modell erfasst.

Passpunkte	RMS _X (m)	RMS _Y (m)	RMS _z (m)	Rückprojektions- fehler (Pixel)
11	0.031	0.022	0.022	0.47
unabh. Kontrolle	0.030	0.022	0.036	0.40
keine	0.769	1.142	0.456	0.44
1 (Pkt. 101)	0.773	1.100	0.666	0.44
2 (101, 111)	0.188	0.141	0.500	0.44

Tab. 1: Genauigkeitsangaben zur Bildtriangulation



Abb. 4: Orthophotomosaik 2018

2.4 Airborne Laserscanning (ALS)

Vom Amt der Kärntner Landesregierung (KAGIS) wurde für das gegenständliche Projekt hochauflösende Höhendaten als DGM bzw. ein digitales Oberflächenmodell (DOM) im 1m-Raster zur Verfügung gestellt. Der Unterschied zwischen DOM und DGM ist visuell kaum zu erkennen. Für alle weiterführenden Analysen wurde stets das DGM verwendet. Insbesondere dient das ALS-Modell als Höhenbezug, um etwaige systematische Offsets in den Höhen anderer Vergleichsdaten (Epochen/Methoden) erkennen und kompensieren zu können. Aus den statistischen Daten können Genauigkeitsaussagen abgeleitet werden. Aus praktischen Gründen wird von einem quasi fehlerfreien ALS-Modell ausgegangen.

2.5 Terrestrisch-photogrammetrische Vermessung

Der Einsatz der terrestrischen Photogrammetrie hat in der Gebirgskartographie und insbesondere in der Gletschervermessung eine lange Tradition (RINNER & BURKHARDT 1972). Am Beispiel des Gössnitzkeeses wurde versucht an diese Tradition anzuschließen. Es wurden im Laufe der Zeit unterschiedliche analoge und digitale Kameras verwendet (KAUFMANN 2010). Aufgrund der topographischen Situation ist eine flächendeckende terrestrisch-photogrammetrische Aufnahme des Gössnitzkeeses nicht möglich. Der zentrale Teil des Gletschers kann jedoch vom Gegenhang aus (oberhalb der 1850er-Moräne, am Wanderweg zum Roten Knopf gelegen) relativ gut eingesehen werden. 1988 wurde hier erstmals eine photogrammetrische Standlinie mit einem Phototheodoliten Zeiss TAL (Terrestrischen Ausrüstung Leicht) aufgenommen. Im Gelände sind derzeit vier terrestrisch-photogrammetrische Standpunkte (Abb. 3: L, R, M1 und M2), an denen vorzugsweise die Aufnahmen getätigt wurden, markiert. Eine optimale Stereokonfiguration ergibt sich für die Basislinie M2-R. Stand gerätetechnisch bedingt vorerst die Zweibildauswertung im Mittelpunkt, ist jetzt die Mehrbildauswertung Stand der Technik. Die Auswertung aller Epochen 1988-2015 erfolgte digital-photogrammetrisch (KAUFMANN & LADSTÄDTER 2008b; KAUFMANN 2010; KAUFMANN & SEIER 2016). Die Auswertung auf einer klassischen digitalphotogrammetrischen Arbeitsstation (ImageStation von Z/I Imaging), welche primär für den Luftbildfall entwickelt wurde, ist aufwändig und sicher nur dem Spezialisten vorbehalten. Eine Automatisierung des photogrammetrischen Auswerteprozesses ist/war in der gegebenen Aufnahmedisposition nicht möglich. Die Georeferenzierung der terrestrisch-photogrammetrischen Modelle erfolgt über 55 stabile natürliche Passpunkte, welche im absolut orientierten Modell 2003 (Rolleiflex 6006) gemessen wurden. Neue Möglichkeiten der photogrammetrischen Auswertung eröffneten Softwaresysteme auf Basis von Structure-from-Motion (SfM) Technologie. Im Rahmen des gegenständlichen Projekts wurden die kommerziellen Softwaresysteme Photo-Modeler Scanner, Agisoft PhotoScan und Pix4Dmapper Pro verwendet. Es wurden unterschiedliche Konstellationen, wie z.B. klassischer Stereofall M2-R und Mehrbildfall (Aufnahmen aller vier Standpunkte), untersucht. Die monoskopische Messung der natürlichen Passpunkte stellt eine gewisse Fehlerquelle dar. Eine Signalisierung wäre prinzipiell möglich gewesen, wurde aber aus Zeitgründen nicht realisiert. Die terrestrisch-photogrammetrischen Aufnahmen erfolgten stets zeitgleich mit der alljährlichen geodätischen Vermessung. Dies eröffnete die Möglichkeit, das terrestrisch-photogrammetrische Auswerteergebnis in seiner Genauigkeit zu validieren. Der Vergleich der photogrammetrisch ermittelten Höhenprofile (Profil 1) mit den geodätischen Referenzmessungen zeigt in allen Fällen eine positive Systematik, insofern, dass das photogrammetrische Höhenprofil stets höher liegt (bis zu 0.53 m) als das geodätische. Dies liegt wahrscheinlich darin begründet, dass die geodätische Messung punktuell zwischen den Felsblöcken des aufliegenden Blockschutts getätigt wird und die photogrammetrische Messung eher die "Oberfläche"

erfasst. Aus den multitemporalen Aufnahmen lassen sich durch Epochenvergleich Längenänderung, Eisdickenänderung und Fließgeschwindigkeit des Gletschers ableiten. Von besonderem Interesse sind Zeitrafferdarstellungen zur Visualisierung der Oberflächenveränderung (Fließbewegung, Eisdickenänderung und sonstige Massenbewegungen). Diesbezügliche Computeranimationen siehe KAUFMANN (2018). Ergänzend zu den bereits bestehenden Auswertungen wird in diesem Beitrag die Gletscherveränderung zwischen 2015 (terrestrisch-photogrammetrische Aufnahme mit einer Nikon D300) und 2018 (UAV-gestützte Datenerfassung) präsentiert. In der Differenzbildung der beiden DGM erfolgte eine Zentrierung der diesbezüglichen Höhenwerte auf die jeweils zeitgleich durchgeführten geodätischen Messungen.

3 Auswerteergebnisse

3.1 Längenänderung

Das Gössnitzkees ist zwischen 1982 und 2018 kontinuierlich kleiner geworden. Die Lageänderung im Bereich der Gletscherstirn ist durch die alljährliche Vermessung (Trilateration, 1982-2002; Tachymetrie, zeitlich übergreifend 1996-2013; RTK-GNSS, seit 2014) gut dokumentiert. Die in Profilrichtung 1 gemessene Längenänderung beträgt insgesamt 286.5 m. Die mittlere Fließgeschwindigkeit liegt aktuell (2017-2018) bei 0.26 m/Jahr.

3.2 Eisdickenänderung

Die Eisdickenänderung wurde aus den geodätischen Profilmessungen (jährlich) und den erstellten DGM (jährlich bzw. mehrjährig) abgeleitet. Die Abbildungen 5 und 6 zeigen die Eisdickenänderung in Profilrichtung 1. Für das Profil 1 ergibt sich eine mittlere jährliche Eisdickenänderung von -1.47 m/Jahr für den Beobachtungszeitraum 1998-2018. 1850 hatte das Gössnitzkees im Bereich des Gletscherrandsees (Punkt H1) eine Eismächtigkeit von ca. 81 m. Das aus den UAV-Aufnahmen abgeleitete Höhenprofil 1 wurde mit den geodätischen Messungen überprüft und stimmt nach Abspaltung eines systematischen Anteils von 0.064 m auf ± 0.104 m mit diesem überein.

Die Abbildungen 7 und 8 zeigen für den zentralen Bereich oberhalb der Gletscherstirn flächenhaft die mittlere jährliche Eisdickenänderung für die Zeiträume 2012-2018 bzw. 2015-2018.

3.3 Flächenänderung

Im Jahr 2012 betrug die Flächenausdehnung des Gössnitzkeeses 56.56 ha. Die Fläche des Gössnitzkeeses hat sich somit von 1850 bis 2012 um rund 64% verringert (Tab. 1). Der photogrammetrisch ermittelte Gletscherrand von 2012 ist jedoch geringfügig in den schuttbedeckten Bereichen unsicher und muss erst über die Auswertung einer jüngeren Epoche (etwa 2015) bestätigt werden.

3.4 Exakter Gletscherrand 2012

Durch die UAV-gestützte Luftbildaufnahme von 2018 steht eine aktuelle Geländeaufnahme zur Korrektur des vorläufigen Gletscherrandes 2012 im zentralen Bereich des Gletschers zur Verfügung. Die Abbildung 7 zeigt den korrigierten Verlauf des Gletscherrandes 2012 und somit die diesbezügliche Gletscherfläche. Die Analyse der Differenzhöhen in den stabilen, eisfreien Bereichen ergibt einen systematischen Offset von 0.126 m. Die Standardabweichung (1 σ) beträgt ±0.134 m. Das Signifikanzniveau in der Abb. 7 wurde mit 2 σ gewählt.

3.5 Volumenänderung

In der Tab. 1 sind die berechneten Volumenänderungen für den Beobachtungszeitraum 2006-2012 angeführt. Als Referenz dient der rekonstruierte Gletscherstand von 1850.

Epoche	Fläche (ha)	ne (ha) Flächenänderung Volumenänderung kumulativ (%) kumulativ (Mill. m³)		mittlere jährliche Eisdi- ckenänderung (m/Jahr)	
1850	155.56				
2006	62.92	-59.55	-84.52	-0.50 (1850-2006)	
2009	60.92	-60.84	-85.95	-0.82 (2006-2009)	
2012	56.56	-63.64	-87.89	-1.11 (2009-2012)	

Tab. 1: Gletscherveränderung 1850-2012



Abb. 5: Eisdickenänderung 1996-2018 entlang der Profillinie 1



Abb. 6: Eisdickenänderung 1850-2018 entlang der Profillinie 1





Abb. 7: Eisdickenänderung 2012-2018 im zentralen Bereich

4 Diskussion

Das Gössnitzkees zeigt sich als stark schuttbedeckter Gletscher ohne nennenswerte Nährgebiete für Schnee und Eis. Sämtliche Gletscherteilflächen liegen faktisch unterhalb der Gleichgewichtslinie. Bei den derzeit herrschenden regionalklimatischen Bedingungen ist ein nahezu vollständiges Abschmelzen des Gletschers vorhersehbar. Restflächen (mit Toteis) könnten sich in Gunstlagen noch über längere Zeit halten. Der Zerfall der Gletscherfläche ist nun auch im zentralen Teil des Gletschers offensichtlich. Das Eisfreiwerden von größeren Flächen, insbesondere im Fußbereich der Klammerköpfe und auch im orographisch linken Teil des Gletschers, ist durch unterschiedliche Messverfahren belegt.



Abb. 8: Eisdickenänderung 2015-2018 im zentralen Bereich

Abschließend wird nun versucht, die eingesetzten Messverfahren kritisch zu beleuchten (Tab. 2). Die Methodenauswahl zur kleinräumigen Gletscherbeobachtung wird von der spezifischen Aufgabenstellung, den vorhandenen Ressourcen (Zeit, Geld, Personal, Wissen, Verfügbarkeit von Messgeräten) und den Bedingungen vor Ort (Zugänglichkeit, topographische Verhältnisse, Infrastruktur) abhängen. Bei der Gewinnung von flächenhafter Information ist eine Kombination aus klassischer Luftbildauswertung mit Daten vom BEV (durchschnittlich alle drei Jahre) und einer optionalen UAV-gestützten Luftbildaufnahme bzw. terrestrisch-photogrammetrischen Aufnahme im Jahresrhythmus für ausgewählte Teilbereiche des Gletschers empfehlenswert. Geodätische Messungen würden weiterhin alljährlich mittels RTK-GNSS durchzuführen sein. Eine wesentliche Erleichterung bei den Geländearbeiten im Rahmen von UAV-Einsätzen würde der Wegfall der bisher für die indirekte Georeferenzierung notwendigen Passpunktmessung ergeben. Moderne UAV-Systeme mit der Möglichkeit einer zentimetergenauen direkten Georeferenzierung mittels RTK- bzw. PPK (Post-Processed Kinematic)-Technologie sind bereits im operationellen Einsatz und könnten daher für die Gletschervermessung im Allgemeinen und für das Gössnitzkees im Speziellen von Interesse sein (PIX4D 2018).

Messmethode	Vorteil	Nachteil	Genauigkeit
Totalstation	etabliertes Verfahren	schwere Ausrüstung, perso- nalintensiv, nur Einzelpunkt- messung, Begehung notwen- dig	±1-3 cm
RTK-GNSS	etabliertes Verfahren, wetter- unabhängig, leichte Ausrüs- tung	nur Einzelpunktmessung, Begehung notwendig, Ab- schattung, Multipath	±1-3 cm
klassische Luftbildaufnahme	etabliertes Verfahren, groß- räumig, flächendeckend, gute Bodenauflösung, Bildorientie- rung durch AT meist bekannt	Aufnahmezyklus von 3 Jah- ren (BEV), kostenintensive photogrammetrische Work- station, Probleme bei Schnee-/Schattenflächen	homogene Genauigkeit, ±10-15 cm oder besser
UAV-gestützte Luftbildauf- nahme	freie Wahl von Aufnahme- zeitpunkt und Bodenauflö- sung, kosteneffizientes Ver- fahren, Möglichkeit der direk- ten Georeferenzierung (RTK/PPK), optimal für De- tailstudien (z.B. im Zungen- bereich), hohe Genauigkeit und geometrische Auflösung möglich, flächendeckend	administrative und rechtliche Hürden, wetterabhängig (Temperatur, Wind, Nieder- schlag), Einschränkungen im hochalpinen Bereich mit Abschattung, beschränkte Flächenleistung bei geringer Batterieleistung, Visual Line- of-Sight (VLOS), ohne RTK/PPK aufwändige Pass- punktsignalisierung und - messung, Modelldeformatio- nen, Begehung notwendig	homogene Genauigkeit, ±3-5 cm
Airborne Laserscanning (ALS)	flächendeckend, gute Auflö- sung, DOM, DGM, keine weitere Verarbeitung not- wendig, gute Referenz für die Höhe	meist nur Daten einer Epoche vorhanden, Neuaufnahmen im Gebirge eher selten, Aufnahmeintervalle i. Allg. mehr als 10 Jahre	1m-Raster
terrestrisch- photogrammetrische Auf- nahme	billig (Consumer-Kamera, Software), Automatisierung der Auswertung, flächenhafte Information, hohe Auflösung (Phototextur), 3D-Auswertung in Punkten, Linien und Flä- chen, demokratisiertes Ver- fahren (kein Spezialwissen notwendig), leichte Ausrüs- tung, Einmannaufnahme möglich, Zeitrafferaufnahmen	Georeferenzierung über Passpunkte, wetterabhängig, abhängig von der topographi- schen Situation, Gegenhang notwendig, entfernungsab- hängige Tiefenmessgenauig- keit, Modelldeformation	inhomogene Genauigkeit, in der Bildebene ±1 Pixel bis Subpixel, in Aufnahmerich- tung quadratisches Fehlerge- setz, Genauigkeit auch ab- hängig von der Mehrfach- überdeckung

Tab. 2: Bewertung der eingesetzten Messverfahren

5 Resümee

Bezugnehmend auf das Beispiel Gössnitzkees bietet sich für das Langzeitmonitoring von kleinen Gletschern (in Österreich) bei einer zeitlichen Auflösung von einem Jahr folgende Vorgangsweise an: (1) Gesamtauswertung, etwa alle drei Jahre, aus Luftbildern des BEV und (2) jährliche Vermessung der Gletscherstirn und mindestens eines Längsprofils mit RTK-GNSS und gegebenenfalls (3) ergänzende Auswertungen aus UAV-Befliegungen bzw. terrestrischphotogrammetrische Aufnahmen (optional TLS-Aufnahmen) bei (4) gleichzeitiger Integration vorhandener ALS-Daten. Die Messungen am Gössnitzkees werden bis auf weiteres durchgeführt. Nach dem völligen Abschmelzen des Gletschers werden Blockwerk und Schutt im Karbereich als stumme Zeugen der einstigen Vergletscherung zurückbleiben.

6 Dank

Besonderer Dank gilt allen Personen, die im Laufe der nahezu vier Jahrzehnte tatkräftig an dem Monitoringprojekt Gössnitzkees mitgewirkt haben. Die Sektionen Barmen & Wuppertal des Deutschen Alpenvereins haben das gegenständliche Gletschermonitoring stets wohlwollend (freies Quartier auf der Elberfelderhütte) unterstützt. Finanzielle Unterstützung erfolgte u.a. durch die Nationalparkverwaltung Hohe Tauern, Kärnten. Die UAV-Befliegung fand im Rahmen der NAWI Graz Forschungskooperation zwischen den beteiligten Universitäten statt.

7 Literaturverzeichnis

ABERMANN, J., FISCHER, A., LAMBRECHT, A. & GEIST, T., 2010: On the potential of very highresolution repeat DEMs in glacial and periglacial environments. The Cryosphere, 4, 53-65.

AGISOFT, 2018: <u>http://www.agisoft.com/</u>, letzter Zugriff 21.12.2018.

DJI, 2018: https://www.dji.com/de/phantom-4, letzter Zugriff 21.12.2018.

- FISCHER, A., PATZELT, G., ACHRAINER, M., GROß, G., LIEB, G.K., KELLERER-PIRKLBAUER, A. & BENDLER, G., 2018: Gletscher im Wandel – 125 Jahre Gletschermessdienst des Alpenvereins. Springer Spektrum, Springer-Verlag, 139 S.
- FISCHER, A., SEISER, B., STOCKER-WALDHUBER, M., MITTERER, C. & ABERMANN, J., 2015a: Tracing glacier changes in Austria from the Little Ice Age to the present using a lidarbased high-resolution glacier inventory in Austria. The Cryosphere, **9**, 753-766.
- FISCHER, A., SEISER, B., STOCKER-WALDHUBER, M., MITTERER, C. & ABERMANN, J., 2015b: The Austrian Glacier Inventories GI 1 (1969), GI 2 (1998), GI 3 (2006), and GI LIA in ArcGIS (shapefile) format. PANGAEA, <u>https://doi.pangaea.de/10.1594/PANGAEA.844988</u>, letzter Zugriff 21.12.2018, supplement to FISCHER et al. 2015a.
- GANDER, D., HOLLINGER, G., LANZENDÖRFER, H. & STRAUSS, C., 2003: Virtueller Überflug der Schobergruppe (Nationalpark Hohe Tauern, Kärnten). Unpublizierte Bakkalaureatsarbeit, Institut für Geodäsie, Technische Universität Graz, 70 S.
- MODER, T., 2008: Ermittlung der Volumen- und Flächenänderung von Gössnitz- und Hornkees aus Luftbildzeitreihen. Unpublizierte Bakkalaureatsarbeit, Institut für Geodäsie, Technische Universität Graz, 52 S.
- KAUFMANN, V., 2010: Dokumentation des Gletscherrückgangs am Gössnitzkees (Schobergruppe, Hohe Tauern) für den Zeitraum 1988-2009 mittels terrestrischer Photogrammetrie. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 19, Kohlhofer, G. & Franzen, M. (Hrsg.), 30. Wissenschaftlich-Technische Jahrestagung der DGPF, 1.-3. Juli 2010 an der TU Wien, Tagungsband, 513-522.
- KAUFMANN, V., 2018: <u>https://www.staff.tugraz.at/viktor.kaufmann/animations.html</u>, letzter Zugriff 21.12.2018.

- KAUFMANN, V. & LADSTÄDTER, R., 2004: Documentation of the movement of the Hinteres Langtalkar rock glacier. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 35(B7), 893-898.
- KAUFMANN, V. & LADSTÄDTER, R., 2008a: Documentation of the retreat of Gössnitzkees and Hornkees Glaciers (Hohe Tauern Range, Austria) for the time period 1997-2006 by means of aerial photogrammetry. Proceedings of the 6th ICA Mountain Cartography Workshop, 11-15 February 2008, Lenk, Switzerland, Institute of Cartography and Geoinformation, ETH Zurich, Switzerland, 115-123.
- KAUFMANN, V. & LADSTÄDTER, R., 2008b: Application of terrestrial photogrammetry for glacier monitoring in alpine environments. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37(B8), 813-818.
- KAUFMANN, V. & LADSTÄDTER, R., 2004: Documentation of the retreat of a small debris-covered cirque glacier (Gössnitzkees, Austrian Alps) by means of terrestrial photogrammetry. Proceedings of the 4th ICA Mountain Cartography Workshop, 30 September – 2 October 2004, Vall de Núria, Spain, Monografies tècniques, 8, Institut Cartogràfic de Catalunya, Barcelona, 65-76.
- KAUFMANN, V. & LIEB, G.K., 2002: Investigations on the retreat of two small circue glaciers (Gössnitzkees and Hornkees) in the Austrian Alps, Europe. High-Mountain Remote Sensing Cartography 1998, Proceedings of the 5th International Symposium on the Use of Remote Sensing Data in Mountain Cartography, Karstad University Studies 2002:27, 75-82.
- KAUFMANN, V. & PLÖSCH, R., 2000: Mapping and visualization of the retreat of two circue glaciers in the Austrian Hohe Tauern National Park. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 33(B4), 446-453.
- KAUFMANN, V. & SEIER, G., 2016: Long-term monitoring of glacier change at Gössnitzkees (Austria) using terrestrial photogrammetry. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 41(B8), 495-502.
- KAUFMANN, V., KELLERER-PIRKLBAUER, A. & KENYI, L.W., 2009: Gletscherbewegungsmessung mittels satellitengestützter Radar-Interferometrie: Die Pasterze (Glocknergruppe, Hohe Tauern, Kärnten. Zeitschrift für Gletscherkunde und Glazialgeologie, **42**(1), 85-104.
- KAUFMANN, V., SEIER, G., SULZER, W., WECHT, M., LIU, Q., LAUK, G. & MAURER, M., 2018: Rock glacier monitoring using aerial photographs: Conventional vs. UAV-based mapping

 a comparative study. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(1), 239-246.
- KÄÄB, A., 2005: Remote Sensing of Mountain Glaciers and Permafrost Creep. Schriftenreihe Physische Geographie: Glaziologie und Geomorphodynamik, 48, Geographisches Institut der Universität Zürich, 264 S.
- KLINGHAN, T., 2017: Quantifizierung und Visualisierung des Gletscherrückgangs am Gössnitzkees und Hornkees (Schobergruppe, Kärnten). Unpublizierte Bakkalaureatsarbeit, Institut für Geodäsie, Technische Universität Graz, 43 S.
- LAMBRECHT, A. & KUHN, M., 2007: Glacier changes in the Austrian Alps during the last three decades, derived from the new Austrian glacier inventory. Annals of Glaciology, 46, 177-184.
- LANG, H. & LIEB, G.K., 1993: Die Gletscher Kärntens. Verlag des Naturwissenschaftlichen Vereins für Kärnten, Klagenfurt, 184 S.

- LIEB, G.K. & KELLERER-PIRKLBAUER, A., 2018: Gletscherbericht 2016/2017. Bergauf 2/2018, 73(143), 20-29.
- MAYER, C., 2010: The early history of remote sensing of glaciers. Remote Sensing of Glaciers: Techniques for Topographic, Spatial and Thematic Mapping of Glaciers, P. Pellikka & W.G. Rees (Hrsg.), CRC Press, Taylor and Francis Group, 67-80.
- MODER, T., 2008: Ermittlung der Volumen- und Flächenänderung von Gössnitz- und Hornkees aus Luftbildzeitreihen. Unpublizierte Bakkalaureatsarbeit, Institut für Geodäsie, Technische Universität Graz, 52 S.
- PAAR, G., HUBER, N.B., BAUER, A., AVIAN, M. & REITERER, A., 2012: Vision-Based Terrestrial Surface Monitoring. Terrigenous Mass Movements: Detection, Modelling, Early Warning and Mitigation Using Geoinformation Technology, Pradhan, B. & Buchroithner, M. (Hrsg.), Springer, Berlin, Heidelberg, 283-348.
- PIX4D, 2018: <u>https://www.pix4d.com/blog/rtk-ppk-drones-gcp-comparison</u>, letzter Zugriff 21.12.2018.
- RINNER, K. & BURKHARDT, R. (Hrsg.), 1972: Gletscherphotogrammetrie. Handbuch der Vermessungskunde (Jordan/Eggert/Kneissl), Band III a/2, Photogrammetrie, J.B. Metzlersche Verlagsbuchhandlung, Stuttgart, § 121, 1428-1470.
- SEIER, G., KELLERER-PIRKLBAUER, A., WECHT, M., HIRSCHMANN, S., KAUFMANN, V., LIEB, G. K. & SULZER, W., 2017: UAS-Based Change Detection of the Glacial and Proglacial Transition Zone at Pasterze Glacier, Austria. Remote Sensing 2017, 9(6), 549; <u>doi:10.3390/rs9060549</u>.
- ZAMG, 2018: <u>https://www.zamg.ac.at/cms/de/klima/informationsportal-klimawandel/stand-punkt/klimavergangenheit/neoklima/lufttemperatur</u>, letzter Zugriff 21.12.2018.

Erfassung von Veränderungen der Sedimentkörper eines Wildbaches (Krumeggerbach, Wölzer Tauern) mittels unbemanntem Luftfahrzeug

GERNOT SEIER^{1,2}, MATTHIAS WECHT² & WOLFGANG SULZER²

Zusammenfassung: Unbemannte Luftfahrzeuge (uLFZ) werden derzeit in vielerlei Zusammenhang thematisiert. Ja es scheint zuweilen so, als ermöglichten diese Geräte für viele Themenbereiche eine detailliertere Sichtweise oder aber einen gänzlich neuen Lösungsweg. In diesem Beitrag wird u. a. der Frage nachgegangen, wie sich uLFZ zur Erfassung und Quantifizierung von Veränderungen der Erdoberfläche einsetzen lassen; und zwar ergänzend oder als Alternative zu bestehenden Herangehensweisen. Am Beispiel des Krumeggerbaches wird aufgezeigt, dass uLFZ für die Quantifizierung der Veränderungen eines Wildbaches geeignet sind. Darüber hinaus werden einige Details der Qualitätsbeurteilung diskutiert und dies verdeutlicht, dass dies auch die zur Interpretation herangezogenen Ergebnisse beeinflussen kann.

1 Einleitung

Im Juli 2011 kam es als Folge von Starkniederschlägen mit räumlichem Schwerpunkt über den Wölzer Tauern zu einer Hochwassersituation im Bereich des Schöttlbaches und des Wölzertales wodurch u. a. in direkter Umgebung der Stadt Oberwölz (47°12' N, 14°17' E) starke Schäden verursacht wurden. Immer wieder ist dieses Gebiet von Unwettern und Hochwasserschäden betroffen, zuletzt etwa im September 2018 (WOCHE STEIERMARK 2018) oder auch Anfang August 2017 (ORF 2017, DIE PRESSE 2017). Abseits des Starkniederschlagsereignisses sind die leicht erodierbaren Sedimentkörper der Wildbäche Schöttl- und Krumeggerbach als ein weiterer Faktor für diese Situation zu betrachten. Zur Quantifizierung der Veränderungen der Sedimentkörper wurden neben einer Reihe unterschiedlicher Verfahren (i. W. Sedimenttransport-, Bewegungs-, Niederschlags- und Abflussmessungen, vgl. SASS et al. 2015) Aufnahmen mithilfe eines unbemannten Luftfahrzeuges (uLFZ) durchgeführt. Die wesentlichen Sedimentquellgebiete wurden von STANGL et al. (2016) dokumentiert und wurden als postglaziale Sedimentkörper identifiziert, die in beiden Fällen (Schöttlbach, Krumeggerbach) unmittelbar an das Gerinne angrenzen. Durch Befliegungen mithilfe unbemannter Luftfahrzeuge konnten räumlich hochauflösende Datengrundlagen wie digitale Höhenmodelle (DHM) und Orthophotos für mehrere Zeitpunkte (November 2014 und Juni 2015) erfasst werden (SEIER et al. 2017a,b). Der vorliegende Beitrag präsentiert Ergebnisse unter Berücksichtigung einer weiteren Befliegung zum Zeitpunkt Mai 2017.

¹ Universität für Bodenkultur Wien, Institut für Landschaftsentwicklung, Erholungs- und Naturschutzplanung, Peter-Jordan-Straße 65, A-1180 Wien, E-Mail: gernot.seier@boku.ac.at

² Karl-Franzens-Universität Graz, Institut für Geographie und Raumforschung, Heinrichstraße 36, A-8010 Graz, E-Mail: [Gernot.Seier, Matthias.Wecht, Wolfgang.Sulzer]@uni-graz.at

Der Großteil der Abschnitte entlang der ufernahen Bereiche wurde in bisherigen Studien allerdings als entweder Transportgebiet für Sedimentfracht oder aber als nicht-signifikante Änderungen ausweisendes Areal charakterisiert. Diese Charakterisierung ist z. T. jedoch fraglich und der vorliegende Beitrag zielt daher unter Fokussierung auf das Teileinzugsgebiet Krumeggerbach darauf ab, zu zeigen, ob und wie (a) die Auswahl der zur Qualitätsbeurteilung der DHM herangezogenen Kontrollpunkte relevant ist, (b) wie eine Qualitätsbeurteilung des Lagefehlers direkt anhand der Orthophotos erfolgen kann und (c) wie sich eine Änderung des Schwellenwertes, der markante von nicht markanten vertikalen Änderungen unterscheidet, auf die Ergebnisse auswirkt.

2 Das Untersuchungsgebiet

Das Untersuchungsgebiet liegt im Einzugsgebiet des Schöttlbaches und befindet sich in den Wölzer Tauern als Teil der Niederen Tauern (Abb. 1). In räumlicher Hinsicht fokussiert wird dabei auf einen Abschnitt des Mittellaufes des Krumeggerbaches, der als Nebengewässer des Schöttlbaches zum Sedimenteintrag beiträgt. Die Sedimentverfrachtung im Krumeggerbach ist insbesondere dadurch von Interesse, da Sedimenttransport hier in größerem Ausmaß stattfindet als im Bereich des Schöttlbaches (SASS et al. 2015). Das beobachtete Mittelwasser bei Oberwölz (am Auslass des Einzugsgebietes) lag von 2012-2015 bei 1,5 m³/s, das 30-jährliche Hochwasser bei 80 m³/s und das 100-jährliche Hochwasser bei 128 m³/s (SCHNEIDER et al. 2016). Beim Hochwasserereignis 2011 stieg der Abfluss binnen kurzer Zeit auf etwa 100 m³/s an und verursachte massiven Geschiebetransport und u. a. dadurch kam es zu Überflutungen (SCHNEIDER et al. 2016). In Abb. 2 sind die zentral gelegenen Bereiche im Untersuchungsgebiet dargestellt, die ein tief eingeschnittenes Bachbett und daran anschließende und z. T. durch Erosion geprägte Hangbereiche umfassen. Das Befliegungsgebiet ist etwa 40 ha groß, wobei die Größe des zentralen und im Mittelpunkt der eigentlichen Auswertung stehenden Bereiches etwa 3,1 ha beträgt.



Abb. 1: Das Untersuchungsgebiet ist ein Teilabschnitt im Mittellauf des Krumeggerbaches (rot umrandet) und befindet sich unweit von Oberwölz in der Obersteiermark.



Abb. 2: Der zentrale Bereich des Untersuchungsgebietes Krumeggerbach, Blickrichtung Süd-Ost (aufgenommen am 17.05.2017). Die mittels uLFZ-Befliegungen dokumentierbaren Änderungen umfassen i. W. die Areale entlang des Gerinnes und eines seitlichen Erosionshanges.

3 Methoden

3.1 Befliegung mittels unbemanntem Luftfahrzeug

Im Bereich des Mittellaufes des Krumeggerbaches wurden mithilfe eines als Nurflügler konzipierten uLFZ (Quest UAV, Spannweite von knapp unter 2 m, Abfluggewicht knapp unter 5 kg) Aufnahmen des Untersuchungsgebietes unter Verwendung einer handelsüblichen Systemkamera (Sony α 6000, Objektiv Sony E 16 mm F2.8) bei einer Flughöhe von etwa 140 m ü. Gr. akquiriert (Fluggeschwindigkeit etwa 65 km/h, Flugdauer etwa 10 min, Längsüberlappung etwa 80 % bzw. Querüberlappung etwa 60 %, Basis-/Höhenverhältnis etwa 1:2,2, Bildmaßstab etwa 1:8750). In Abb. 3 ist das Aufnahmesetting (Start- und Landeareal, Flugrouten, Pass- und Kontrollpunkte) dargestellt.



Abb. 3: Die zurückgelegten Flugrouten der drei uLFZ-Befliegungen sowie Kontrollpunkte bzw. Passpunkte (1-5 und vier zusätzliche), letztgenannte umgrenzen das eigentliche Interessensgebiet, den anhand der Schummerung dunkel erscheinenden und deutlich abzugrenzenden Erosionshang und daran anschließenden Bachbereich.

3.2 Auswertung der Luftbildaufnahmen

Die mittels uLFZ erfassten Luftbildaufnahmen (65 Aufnahmen am 29.11.2014, 98 Aufnahmen am 25.06.2015, 61 Aufnahmen am 17.05.2017) wurden photogrammetrisch (unter Anwendung des Structure-from-Motion (SfM)-Ansatzes) zu einem DHM und Orthophoto mit einer Bodenauflösung von 0,04 m verarbeitet. Das SfM-Verfahren basiert auf photographischen Aufnahmen aus unterschiedlichen Blickwinkeln mit der Besonderheit, dass die Bilder unter Verwendung von Matching-Verfahren automatisch verknüpft und Punktwolken generiert werden, wobei simultan auch die äußere und z. T. auch innere Kameraorientierung berechnet bzw. abgeschätzt werden können (ULLMAN 1979, BROWN & LOWE 2005, SNAVELY et al. 2008, WESTOBY et al. 2012). Im vorliegenden Fall wurde das mithilfe der Software Agisoft PhotoScan Professional (v. 1.2.6 2834) umgesetzt. Zur indirekten Georeferenzierung der Luftbildaufnahmen

wurden fünf (November 2014, Juni 2015) bzw. vier zusätzliche (Mai 2017) Passpunkte unter Verwendung eines Differentiellen Globalen Positionierungssystems (DGPS) mittels Echtzeit-Korrekturverfahren (EPOSA 2017) geodätisch vermessen. Die vertikalen Änderungen der Erdoberfläche wurden durch direkten Vergleich der DHM, d.h. Differenz der Höhenmodelle, berechnet. Durch den Abgleich der Orthophotos mittels normierter Kreuzkorrelation (Normalized Cross Correlation, NCC) können oberflächenbedingte Änderungen quantifiziert werden und dies basiert auf den Grauwerten der Bilder und ausreichendem Bildkontrast (KÄÄB & VOLLMER 2000, KAUFMANN & LADSTÄDTER 2002). Zur Umsetzung wurde die Software CIAS (CIAS 2013) verwendet, die es wie in NCC-Verfahren i. A. erlaubt, eine Mustermatrix in einer Suchmatrix unter definiertem Punkabstand abzugleichen.

4 Ergebnisse

4.1 Qualitätsbeurteilung

Die Qualitätsbeurteilung der DHM erfolgte unter Verwendung von unabhängigen, d. h. zur Georeferenzierung nicht verwendeten DGPS-Kontrollpunkten (n = 89) durch direkten Vergleich der geodätisch bzw. photogrammetrisch ermittelten Höhenwerte. Die Standardabweichung (SD) dieser Höhenwertdifferenzen liegt in allen Fällen zwischen $\pm 0,17$ m und $\pm 0,23$ m bei maximalen Differenzen die zumeist in einem Wertebereich von etwa $\pm 0,5$ m bzw. z. T. auch deutlich darunter liegen (im Mittel 0,01 m bis 0,08 m, vgl. Abb. 4a-c). Werden allerdings von den 89 unabhängigen Kontrollpunkten nur jene 14 herangezogen, die im zentralen Bereich des Untersuchungs- und damit tatsächlichen Interessensgebietes liegen, dann können große Höhenwertdifferenzen die mangels Bildüberlappung am Rande des photogrammetrisch basierten Modells liegen als modellbedingt und nicht die Qualität im untersuchten Zentralbereich beschreibend angesehen werden. Die auf diese ausgewählten Kontrollpunkte bezogenen maximalen Differenzen sind deutlich geringer, während die mittleren Differenzen in ähnlicher Größenordnung liegen und die Standardabweichung zwischen $\pm 0,04$ m und $\pm 0,12$ m liegt (Abb. 4d-f).

Die relative Lagegenauigkeit der Orthophotos wird im vorliegenden Beispiel anhand NCC beurteilt. Gegenüber üblichen Anwendungen dieses Verfahrens sollen hier nicht jene Differenzen der Bildinhalte, die durch Änderungen der Erdoberfläche entstanden sind, sondern jene, die durch modellbedingte Fehler im Zuge der photogrammetrischen Auswertung entstanden sind, ermittelt werden. Um diese aus unterschiedlichen Gründen (Anordnung und Anzahl der Passpunkte, Aufnahmekonfiguration und -geometrie, Qualität der Aufnahmen etc.) auftretenden Fehler zu quantifizieren wurde das NCC-Verfahren an Teilflächen im unmittelbaren ufernahen Bereich, für die durch visuellen Vergleich der Orthophotos unveränderte Oberflächensituationen anzunehmen waren, angewendet. Die ermittelten Lagedifferenzen der drei Orthophotos liegen dabei im Mittel zwischen 0,06 m und 0,10 m bei einer Standardabweichung zwischen $\pm 0,02$ m bis $\pm 0,04$ m (Abb. 5).



Abb. 4: Höhenwertdifferenzen (m) der DGPS-Messungen (n = 89) vom 26.11.2016: (a) DGPS vs. uLFZ-DHM November 2014, (b) DGPS vs. uLFZ-DHM Juni 2015, (c) DGPS vs. uLFZ-DHM Mai 2017, (d)-(f): in derselben Reihenfolge angeführte Höhenwertdifferenzen mit reduzierter Anzahl von Kontrollpunkten (n = 14), die im zentralen Bereich liegen.



Abb. 5: Lagefehler der Orthophotos an zu erwartenden lagestabilen Punkten (n = 718) des zentralen Bachbereiches, berechnet mittels NCC: (a) uLFZ-Orthophotos November 2014 vs. Juni 2015, (b) uLFZ-Orthophotos November 2014 vs. Mai 2017, (c) uLFZ-Orthophotos Juni 2015 vs. Mai 2017.

4.2 Höhendifferenzen

Das in Abb. 6 dargestellte Querprofil A-B zeigt für den zentralen Untersuchungsbereich (BMN M31-Koordinaten des Punktes A: Rechtswert 523746, Hochwert 234692,6; Koordinaten des Punktes B: Rechtswert 523687,7, Hochwert 234666,1), dass im Zeitraum zwischen November

2014 und Juni 2015 kaum Unterschiede festzustellen sind. Anders verhält es sich im längeren Zeitabschnitt zwischen Juni 2015 und Mai 2017: die steilsten Bereiche des seitlich an den Bach angrenzenden Hanges wurden abgetragen und im flacheren Hangfußbereich abgelagert, wobei die Mächtigkeit des erodierten Bereichs bis zu etwa 2 m beträgt. Weiters sind ein verlagertes Bachbett, ein kurzer Abschnitt eines unveränderten ufernahen Bereiches und vegetationsbedingte Höhenunterschiede festzustellen.



Abb. 6: Querprofil A-B (vertikale Überhöhung 1,5:1): Höhenwerte der uLFZ-DHM November 2014 (rot), Juni 2015 (blau) und Mai 2017 (grün).

Die Höhenwertdifferenzen der jeweiligen DHM zeigen in kartographischer Betrachtung (Abb. 7), dass der Großteil durch nicht markante Veränderungen oder aber durch als Transportareal dienende Bereiche gekennzeichnet ist. Dies trifft zumindest für den Zeitraum November 2014 bis Juni 2015 zu (Abb. 7a), während unter Einbeziehung der jüngsten Aufnahme von Mai 2017 durchaus kleinere Abtragungsstrukturen zum Vorschein kommen (Abb. 7c,e). Weitere Beispiele für kleinräumigere Ausnahmen sind jene in Abb. 7b,d,f, die größere Höhenunterschiede zeigen und den in Abb. 6 gezeigten Ablagerungsbereich und auch das veränderte Bachbett flächenhaft nachvollziehbar machen. In Abb. 8 wurde ähnlich mit den DHM und der graphischen Darstellung der Höhendifferenzen verfahren, mit dem wesentlichen Unterschied, dass die zur Unterscheidung von modellbedingten bzw. tatsächlichen (d. h. oberflächenbedingten) Höhendifferenzen herangezogene SD hier nicht mit 0,2 m sondern 0,1 m definiert wurde (vgl. Kapitel 4.1). Wie hier zu sehen ist, wird die Aussage für den bisher diskutierten Detailausschnitt (Abb. 8b,d,f) dadurch kaum verändert, während in den restlichen Bereichen nun Abtragungs- und Ablagerungsbereiche zutage treten, die vorher nicht als solche zu interpretieren waren. In beiden Fällen sind weite Teile des Untersuchungsgebietes abgesehen von kleinräumigeren Ausnahmen durch Höhendifferenzen gekennzeichnet, die innerhalb der einfachen SD liegen und daher entweder modellbedingt sein können oder aber Bereiche des Sedimenttransportes sind.



Abb. 7: Höhenwertdifferenzen der (a) uLFZ-DHM November 2014 und Juni 2015, (c) uLFZ-DHM Juni 2015 und Mai 2017 und (e) uLFZ-DHM November 2014 und Mai 2017, farbskaliert nach Vielfachen der einfachen SD (0,2 m, gelb) mit Ausnahme der jeweils letzten Kategorie (Werte größer der dreifachen SD), die sehr wahrscheinlich markante Oberflächenänderungen zeigen, bspw. Abtragungs- und Ablagerungsbereich in den Detaildarstellungen (b,d,f).



Abb. 8: Höhenwertdifferenzen der (a) uLFZ-DHM November 2014 und Juni 2015, (c) uLFZ-DHM Juni 2015 und Mai 2017 und (e) uLFZ-DHM November 2014 und Mai 2017, farbskaliert nach Vielfachen der einfachen SD (0,1 m, gelb) sowie Detaildarstellungen (b,d,f).

5 Fazit

Dieser Beitrag zeigt, dass (a) es bei der Auswahl der zur Qualitätsbeurteilung der DHM herangezogenen Kontrollpunkte wesentlich ist, ob Punkte an den Rändern der Modelle mitberücksichtigt werden oder ob nur auf jene im zentralen und für die Fragestellung relevanten Bereich Bezug genommen wird, dass (b) ein möglicher Lagefehler der Orthophotos an Bereichen in unmittelbarer Ufernähe, die keine Änderungen erwarten ließen, mittels NCC quantifiziert werden konnten. Aussagekräftige Ergebnisse sind somit ab horizontalen Differenzen von etwa > 0,1 m zu erwarten. (c) Durch eine veränderte Beurteilung der Größe des Schwellenwertes, der markante von nicht markanten vertikalen Änderungen unterscheidet (bei Verwendung aller Kontrollpunkte liegt dieser Wert bei etwa 0,5-0,6 m, bei reduzierter Kontrollpunktanzahl hingegen bei etwa 0,1-0,2 m), treten Strukturen hervor, die bisher als Transportareale bzw. modellbedingte Änderungen angesehen wurden und durch diese andersgeartete Beurteilung sind diese Strukturen als mögliche Erosions- und Sedimentationsareale zu bezeichnen. Eine zweifelsfreie Identifikation von Abtragungs- und Ablagerungsgebieten konnte durch Berücksichtigung der jüngsten Aufnahme in nennenswertem Ausmaß festgestellt werden.

6 Literaturverzeichnis

- BROWN, M. & LOWE, D.G., 2005: Unsupervised 3D Object Recognition and Reconstruction in Unordered Datasets. 5th International Conference on 3D Digital Imaging and Modeling, IEEE Comput. Soc. (Hrsg.), Ottawa, Ontario, Canada, 13.-16. Juni 2005, 1-8.
- CIAS, 2013: Image Correlation Software. http://www.mn.uio.no/geo/english/research/projects/icemass/cias, letzter Zugriff 18.12.2016.
- DIE PRESSE, 2017: Jugendlicher stirbt bei Unwetter in der Obersteiermark. http://diepresse.com/home/panorama/oesterreich/5264268/Jugendlicher-stirbt-bei-Unwetter-in-der-Obersteiermark, letzter Zugriff am 28.08.2017.
- EPOSA, 2017: EPOSA Echtzeitpositionierung Austria. http://www.eposa.at, letzter Zugriff 28.08.2017.
- KÄÄB, A. & VOLLMER, M., 2000: Surface Geometry, Thickness Changes and Flow Fields on Creeping Mountain Permafrost: Automatic Extraction by Digital Image Analysis. Permafrost and Periglacial Processes, 11, 315-326.
- KAUFMANN, V. & LADSTÄDTER, R., 2002: Spatio-temporal analysis of the dynamic behaviour of the Hochebenkar rock glaciers (Oetztal Alps, Austria) by means of digital photogrammetric methods. Grazer Schriften der Geographie und Raumforschung, 37, 119-140.
- ORF Österreichischer Rundfunk, 2017: Oberwölz: Aufräumen nach dem Unwetter. http://steiermark.orf.at/news/stories/2858904/, letzter Zugriff 28.08.2017.
- SASS, O., HARB, G., TRUHETZ, H., STANGL, J. & SCHNEIDER, J., 2015: Abschlussbericht Projekt Climcatch. Klima und Energiefonds Projekt B175084. https://static.unigraz.at/fileadmin/urbi-institute/Geographie/pictures/misc/climcatch/Sass-etal_ACRP4-ClimCatch_Endbericht.pdf, letzter Zugriff 26.09.2016.

- SCHNEIDER, J., STANGL, J., TRUHETZ, H., FIEGER, S., HARB, G. & SASS, O., 2016: Monitoring of Erosion and of Sediment Transport in the River Schöttlbach (Styria) with respect to Climate Change. Journal of Torrent, Avalanche, Landslide and Rock Fall Engineering, 177, 60-72.
- SEIER, G., STANGL, J., SCHÖTTL, S., SULZER, W. & SASS, O., 2017a: UAV and TLS for monitoring a creek in an alpine environment, Styria, Austria. International Journal of Remote Sensing, 38, 2903-2920.
- SEIER, G., SULZER, W., WECHT, M., SCHÖTTL S. & SASS, O., 2017b: Veränderungsdetektion eines Wildbaches mittels unbemanntem Luftfahrzeug. AGIT Journal für Angewandte Geoinformatik, 3, 332-342.
- SNAVELY, N., SEITZ, S.M., & SZELISKI, R., 2008: Modeling the World from Internet Photo Collections. International Journal of Computer Vision, **80**, 189-210.
- STANGL, J., RASCHER E. & SASS O., 2016: Comparative analysis of sediment routing in two different alpine catchments. Source-to-Sink Fluxes in Undisturbed Cold Environments, BEYLICH, A.A., DIXON, J.C. & ZWOLIŃSKI, Z. (Hrsg.), Cambridge University Press, Cambridge, 364-377.
- ULLMAN, S., 1979: The interpretation of Structure from Motion. Proc. R. Soc. London, **203** (Ser. B), 405-426.
- 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.
- WOCHE STEIERMARK, 2018: Unwetter: "Es ist zermürbend". https://www.meinbezirk.at/murtal/c-lokales/unwetter-es-ist-zermuerbend_a2884067, letzter Zugriff 28.12.2018.

Kostengünstige 3D-Dokumentation von Skulpturen des Barockbildhauers Philipp Jakob Straub

THOMAS MIKL¹ & VIKTOR KAUFMANN¹

Zusammenfassung: Dieser Beitrag beschreibt zwei kostengünstige Konzepte zur Erstellung von maßstabsgetreuen photo-texturierten 3D-Modellen von Skulpturen des Barockbildhauers Philipp Jakob Straub (1706-1774) unter Verwendung von Digitalkameras, einer Totalstation und kommerzieller 3D-Auswertesoftware. Die beiden Konzepte wurden im Zuge eines Auftrages des Instituts für Kunstgeschichte der Karl-Franzens-Universität Graz von der Arbeitsgruppe für Fernerkundung und Photogrammetrie des Instituts für Geodäsie an der Technischen Universität Graz entwickelt und für mehrere Skulpturen in Graz und in der benachbarten Südoststeiermark umgesetzt. Der Hintergrund für diesen Auftrag ist das EU-Projekt "Tracing the Art of the Straub Family", welches das Werk des erwähnten Barockbildhauers dokumentieren und für klassische und moderne Medien zugänglich machen soll.

1 Einleitung

Im Rahmen des EU-Projektes "Tracing the Art of the Straub Family" (TRARS 2018) sollten im Auftrag des Instituts für Kunstgeschichte der Karl-Franzens-Universität Graz von der Arbeitsgruppe für Fernerkundung und Photogrammetrie des Instituts für Geodäsie an der Technischen Universität Graz exemplarisch mehrere Skulpturen des Barockbildhauers Philipp Jakob Straub (1706-1774) in Graz und in der benachbarten Südoststeiermark dreidimensional vermessen und entsprechende Visualisierungen in Form von photo-texturierten 3D-Modellen bereitgestellt werden. Das übergeordnete Ziel des EU-Projektes ist es, das Werk des Bildhauers zu dokumentieren und dieses für klassische und moderne Medien zugänglich zu machen.

Die Herausforderung für die Umsetzung der Auftragsarbeit war es, ein kostengünstiges Konzept zu entwickeln, mit dem einerseits wichtige geometrische Kenngrößen der Skulpturen, wie ihre Höhe, Breite und Tiefe, bestimmt und andererseits auch photo-texturierte 3D-Modelle erstellt werden können. Bei den ausgewählten Skulpturen handelt es sich um lebensgroße Heiligendarstellungen aus Sandstein oder Holz, die zum Teil aufgrund ihrer exponierten Lage nur schwer zugänglich sind oder nur bei schlechten Lichtverhältnissen aufgenommen werden können. In diesem Beitrag werden die Arbeiten für vier ausgewählte Skulpturen exemplarisch erläutert (Tab. 1).

¹ Technische Universität Graz, Institut für Geodäsie, Arbeitsgruppe für Fernerkundung und Photogrammetrie, Steyrergasse 30/I, A-8020 Graz,

E-Mail: mikl@student.tugraz.at, viktor.kaufmann@tugraz.at

Name	Lage	Eigenschaften
Heiliger Leonhard	Mureckerstraße 18,	freistehende und frei zugängliche
	8083 St. Stefan im Rosental	Sandsteinfigur
Maria Immaculata	Mariahilferstraße 9,	Eckstatue aus Sandstein im ersten
	8020 Graz	Obergeschoss
Mater Dolorosa	Minoritenkirche Mariahilf,	Wandstatue aus Holz im Innenraum
	Marihilferplatz 3, 8020 Graz	einer Kirche
Engelsturz	Minoritenkirche Mariahilf,	zwei Statuen auf dem Giebel einer
	Marihilferplatz 3, 8020 Graz	Kirche in über 18 m Höhe

Tab. 1: Name, Lage und Eigenschaften der Skulpturen

Die folgende Abbildung zeigt photographische Aufnahmen der behandelten Skulpturen.



Abb. 1: Photographische Aufnahme der Statue des Heiligen Leonhard (a), der Maria Immaculata (b), der Mater Dolorosa (c) und der Darstellung vom Engelsturz (d)

2 Methoden

LUHMANN (2018) beschreibt unterschiedliche Methoden, um Skulpturen dreidimensional zu vermessen, wie z.B. Tachymetrie, terrestrisches Laserscanning, Streifen- und Musterprojektionssysteme und terrestrische Photogrammetrie.

Eine alleinige Aufnahme von wenigen diskreten Punkten mithilfe eines Tachymeters, um die gewünschten Kenngrößen daraus abzuleiten, erschien von vorhinein aufgrund der mangelnden Definitionsgenauigkeit als ungeeignet. Ein weiterer Nachteil der Tachymetrie ist, dass gewisse unbedingt notwendige Punkte wegen der exponierten Lage der Skulpturen nicht sichtbar oder durch reflektierende Oberflächen nicht messbar sein können. Da es sich bei den Skulpturen um Heiligendarstellungen handelt, welche auch vergoldete Elemente, wie z.B. einen Heiligenschein, vorweisen, wurde eine reine Erfassung der geometrischen Kenngrößen mithilfe eines Tachymeters ausgeschlossen.

Die dreidimensionale Vermessung von Skulpturen mithilfe eines terrestrischen Laserscanners hätte den Vorteil, dass sowohl das maßstabsgetreue photo-texturierte 3D-Modell als auch die gewünschten geometrischen Kenngrößen aus der Punktwolke abgeleitet werden können. Diese Methode kann jedoch ebenfalls an reflektierenden Oberflächen nicht verwendet werden. Des Weiteren eignen sich stationäre Laserscanner im gegebenen Anwendungsfall nicht, da die Instrumentenhöhe nach oben hin beschränkt ist und die Skulpturen meist über die Kopfhöhe ragen oder überhaupt auf einem höheren Niveau positioniert sind. Außerdem wäre die vollständige Erfassung der komplexen Oberfläche nur mit einer großen Anzahl an Aufstellungen möglich. Schlussendlich wurde aufgrund mangelnder Verfügbarkeit eines geeigneten Gerätes diese Methode ausgeschlossen.

Die Verwendung von kostengünstigen Streifen- und Musterprojektionssystemen, wie einem handgehaltenen 3D-Scanner, wäre für kleinere, frei zugängliche Skulpturen denkbar, jedoch sind diese für exponierte Skulpturen, wie sie im Rahmen dieses Auftrages vorliegen, nicht ohne zusätzliche Kosten bzw. Aufwand anwendbar.

Mit einer rein terrestrisch-photogrammetischen Aufnahme der Skulpturen unter Verwendung einer Stehleiter oder einer Hebebühne wären diese Modelle vollständiger. Die Bestimmung des Maßstabes des photogrammetrischen Modells kann durch Hinzunahme von Passpunkten oder bekannten Strecken erfolgen. Die benötigten Passpunkte können wiederum mithilfe eines Tachymeters eingemessen werden. Je nach Genauigkeitsanforderung und Zugänglichkeit der Skulpturen können markante natürliche Punkte oder signalisierte Punkte verwendet werden. Die Kombination aus terrestrisch-photogrammetrischer Aufnahme und Tachymetrie erwies sich unter den vorgegebenen Rahmenbedingungen als optimal, um die Aufgabenstellung der 3D-Modellierung und Kenngrößenermittlung zu lösen.

3 Umsetzung

Bei der Umsetzung des Auftrages wurden zwei Konzepte verwendet, die sich in der Art der verwendeten Kamera und der verwendeten Passpunkte unterscheiden.

Das erste Konzept ermöglicht eine sehr genaue und detaillierte Erfassung von frei zugänglichen Skulpturen, da hier einerseits eine Digitalkamera Nikon D800 mit einer Auflösung von 36 Megapixel und andererseits neben gut definierten, natürlichen Punkten auch signalisierte Passpunkte verwendet wurden. Dieses Konzept wurde nur an der Sandsteinfigur vom Heiligen Leonhard angewendet.

Das zweite Konzept unterscheidet sich im Gegensatz zum ersten, dass anstelle einer hochauflösenden Digitalkamera Nikon D800 eine Digitalkamera Nikon D300 mit einer Auflösung von 12 Megapixel und nur natürliche Passpunkte verwendet werden. Dieses Konzept eignet sich für Objekte, an denen aufgrund ihrer exponierten Lage eine Signalisierung nicht möglich oder einfach nicht erwünscht ist, und wurde auf die drei verbleibenden Skulpturen angewendet.

In den nachfolgenden Abschnitten wird die Umsetzung der beiden Konzepte erläutert.

3.1 Datenakquisition

Zur Erfassung der geometrischen Kenngrößen und zur Erstellung der 3D-Modelle der Skulpturen wurde jeweils eine terrestrisch-photogrammetrische und eine geodätische Aufnahme vorgenommen.

3.1.1 Terrestrisch-photogrammetrische Aufnahme

Bei der terrestrisch-photogrammetrischen Aufnahme wurden die beiden Digitalkameras Nikon D800 und Nikon D300 verwendet. Von jeder Skulptur wurden mehrere Aufnahmen von unterschiedlichen Bodenstandpunkten vorgenommen, um ihre Oberfläche möglichst vollständig abzudecken. Bei der Sandsteinfigur vom Heiligen Leonhard wurde eine einfache Stehleiter verwendet, um auch Blickwinkel von oben zu ermöglichen. Die photographischen Aufnahmen der exponierten Darstellung des Engelsturzes vom Bodenniveau konnten durch die Nutzung einer Hebebühne um Aufnahmen von höheren Standpunkten ergänzt werden. Eine vollständige Abdeckung der Oberfläche dieser Skulptur konnte dennoch nicht erreicht werden. Aufgrund der schlechten Lichtverhältnisse im Innenraum der Minoritenkirche Mariahilf wurde bei der Holzfigur der Mater Dolorosa der interne Blitz der verwendeten Kamera genutzt.

3.1.2 Geodätische Aufnahme

Zur Ableitung des Maßstabes der photogrammetrischen 3D-Modelle wurden bei allen Skulpturen einzelne, gut definierte, natürliche Punkte und bei der Sandsteinfigur des Heiligen Leonhard zusätzlich signalisierte Punkte eingemessen. Die Signalisierung erfolgte mit kleinen selbstklebenden Passmarken mit acht Sektoren. Bei der geodätischen Aufnahme wurden auch nutzbare Zusatzpunkte zur Bestimmung der geometrischen Kenngrößen der Skulpturen, wie ihre Höhe, Breite und Tiefe, aufgenommen. Die geodätischen Messungen der einzelnen Skulpturen wurden mithilfe der Totalstation TPS1201R von Leica Geosystems GmbH von zumindest einem Standpunkt durchgeführt. Bei der Sandsteinfigur vom Heiligen Leonhard und bei der Eckstatue der Maria Immaculata erfolgte die geodätische Aufnahme mit zwei zwangszentrierten Aufstellungen, um gut verteilte Passpunkte einmessen zu können. Für die vier Skulpturen wurden folgende Konzepte und Aufnahmekonfigurationen verwendet.

Name	Konzept	terrestrphotogr. Aufnahme		geodätische Aufnahme		
		Kamera	# Aufn.	Hilfsmittel	# Standpkt.	Art der Passpkt.
Heiliger	1	Nikon	226	Stehleiter	2	signalisiert und
Leonhard		D800				natürlich
Maria	2	Nikon	26	-	2	natürlich
Immaculata		D300				
Mater Dolorosa	2	Nikon	47	interner Blitz	1	natürlich
		D300				
Engelsturz	2	Nikon	(54)	Hebebühne	(3)	natürlich
		D300				

Tab. 2: Verwendete Konzepte und Aufnahmekonfigurationen der Skulpturen

Die terrestrisch-photogrammetrische und die geodätische Aufnahme des Engelsturzes erfolgte im Rahmen einer umfassenden Fassadendokumentation.

3.2 Auswertung

Die Auswertung der geodätischen und der terrestrisch-photogrammetrischen Aufnahme zur Bestimmung der geometrischen Kenngrößen und zur Erstellung der 3D-Modelle erfolgte mit kommerzieller Software. Dabei wurde bei beiden Konzepten in gleicher Art und Weise vorgegangen, indem in einem ersten Schritt die geodätische und in einem zweiten Schritt die terrestrischphotogrammetrische Aufnahme ausgewertet wurde.

3.2.1 Auswertung der geodätischen Aufnahme

In diesem Schritt wurden die diskreten Punktmessungen der geodätischen Aufnahme für die nachfolgende Auswertung der terrestrisch-photogrammetrischen Aufnahme aufbereitet. Zu Beginn wurden diese Punktmessungen mithilfe der Software Geosi VERM (GEOSI VERM 2018) als Polarpunkte bestimmt. Die resultierenden Punkte wurden im Anschluss in ein objektbezogenes Koordinatensystem transformiert.

3.2.2 Auswertung der terrestrisch-photogrammetrischen Aufnahme

Die Auswertung der terrestrisch-photogrammetrischen Aufnahme zur Bestimmung des phototexturierten 3D-Modells der Skulptur erfolgte mithilfe des Verfahrens Structure-from-Motion (SfM) der Software Agisoft PhotoScan (AGISOFT 2018). Das Grundprinzip dieses Verfahrens wird unter anderem in YI et al. (2014) beschrieben. Als Datengrundlage dienten die photographischen Aufnahmen der beiden Digitalkameras Nikon D300 und Nikon D800 im komprimierten JPG-Format, um die Datenmenge möglichst gering zu halten. Zu Beginn wurden die Aufnahmen durch Image-Matching mit einer anschließenden Bündeltriangulation vorläufig relativ orientiert und eine dünne Punktwolke erzeugt. Zur Ableitung des Maßstabes und zur absoluten Orientierung des photogrammetrischen Modells wurden die Passpunkte in den einzelnen photographischen Aufnahmen manuell eingemessen. Die Parameter der inneren Orientierung (Kamerakonstante, Lage des Bildhauptpunktes, Affinparameter und radial-symmetrische Verzeichnung mit K1 und K2) der Kamera wurden in Form einer Simultankalibrierung neu bestimmt. Bei Kleinformen der Skulpturen, wie dem Heiligenschein des Heiligen Leonhard, wurde der Hintergrund in den photographischen Aufnahmen maskiert, um den Rauschanteil in diesen Bereichen zu minimieren. Schließlich erfolgte die Erzeugung einer dichten Punktwolke. Aus dieser wurde mittels Dreiecksvermaschung ein digitales Oberflächenmodell und in weiterer Folge durch Texturierung ein maßstabsgetreues photo-texturiertes 3D-Modell generiert.

4 Ergebnisse

Aufgrund der Anwendung der zwei verschiedenen Konzepte sind die erstellten 3D-Modelle unterschiedlich genau und unterschiedlich detailliert.

Für die Sandsteinfigur vom Heiligen Leonhard wurde das erste Konzept angewendet, wodurch eine hohe Genauigkeit erreicht und ein detailliertes 3D-Modell erstellt werden konnte. Mit dieser Vorgehensweise wurde als relative Genauigkeitsangabe ein Reprojection Error (RE) von 0.98
Pixel erzielt, wobei 1 Pixel im Mittel 0.22 mm am Objekt entspricht. Die absolute Genauigkeit des maßstabsgetreuen 3D-Modells liegt bei einem Root Mean Square Error (RMSE) von 3 mm. Folglich können geometrische Kenngrößen der Skulptur als Abstand zwischen zwei diskreten Punkten der dichten Punkwolke mit einer Genauigkeit von 4.2 mm gemessen werden. Diese Genauigkeitsangabe für geometrische Kenngrößen ergibt sich durch folgenden Zusammenhang.

$\sigma = \sqrt{2} \cdot RMSE$

(1)

Das 3D-Modell ist bis auf kleinere Strukturen, wie einzelne Segmente des dünnen Heiligenscheins (Abb. 2b), welche bei der Dreiecksvermaschung der dichten Punktwolke mit der verwendeten Software Agisoft PhotoScan herausgefiltert wurden, vollständig und ist in Abbildung 2a dargestellt. Eine Visualisierung dieses 3D-Modells wurde mit der Software CloudCompare (CLOUDCOMPARE 2018) in der Form eines Videos erstellt (KAUFMANN 2018).



Abb. 2: Maßstabsgetreues photo-texturiertes 3D-Modell der Sandsteinfigur vom Heiligen Leonhard in einer Gesamtansicht (a) und in einer Detailansicht (b)

Bei den übrigen Skulpturen wurde nach dem zweiten Konzept vorgegangen, wodurch im Vergleich zur Anwendung des ersten Konzeptes geringere Genauigkeiten erreicht und weniger detaillierte 3D-Modelle bestimmt werden konnten. Aufgrund der exponierten Lage der Darstellung des Engelsturzes, welche sich in über 18 m Höhe befindet, konnten trotz Verwendung einer Hebebühne nur Aufnahmen von Standpunkten unterhalb der Skulptur gemacht werden, wodurch das erstellte 3D-Modell nicht vollständig ist (Abb. 3c). Für die Statuen der Maria Immaculata und der Mater Dolorosa konnten im Vergleich zur Darstellung des Engelsturzes nahezu vollständig 3D-Modelle erstellt werden (Abb. 3a & 3b).



Abb. 3: Maßstabsgetreues photo-texturiertes 3D-Modell der Statue der Maria Immaculata (a), der Mater Dolorosa (b) und der Darstellung vom Engelsturz (c)

In der folgenden Tabelle sind jeweils die relative und absolute Genauigkeit der 3D-Modelle, welche durch die Anwendung der beiden Konzepte bei den verschiedenen Skulpturen erreicht wurden, angegeben. Die angeführten Kenngrößen wurden aus den zugrundeliegenden dichten Punktwolken der erstellten 3D-Modelle bestimmt, womit sich die Genauigkeit dieser Messungen aus dem jeweiligen RMSE ergibt (Gl. 1).

Name	Konzept	RE [Pi-	RMSE	Kenngrößen der Gesamtfigur [cm		
		xel]	[mm]	Höhe	Breite	Tiefe
Heiliger Leonhard	1	0.98	3.0	294	118	92
Maria Immaculata	2	0.34	7.3	254	79	56
Mater Dolorosa	2	0.52	5.7	254	74	46
Engelsturz	2	0.47	14.7	448	232	236

Tab. 3: Genauigkeit und Kenngrößen der maßstabsgetreuen photo-texturierten 3D-Modelle der Skulpturen

Auffällig ist der RMSE des 3D-Modells der Darstellung des Engelsturzes, welcher im Vergleich zu den übrigen behandelten Skulpturen bei über 1 cm liegt. Dieses Ergebnis ist auf die erschwerten Aufnahmebedingungen bezüglich der exponierten Lage des Objektes in über 18 m Höhe und unter anderem auf die schlechtere Schnittgeometrie zurückzuführen.

5 Fazit & Ausblick

Die beiden Konzepte zur Bestimmung von geometrischen Kenngrößen und zur Erstellung von maßstabsgetreuen photo-texturierten 3D-Modellen von Skulpturen haben sich im Anwendungsfall für lebensgroße Statuen aus Sandstein oder Holz bewährt. Mit dem ersten Konzept können durch die Verwendung von signalisierten Passpunkten und einer hochauflösenden Digitalkamera Nikon D800 detaillierte maßstabsgetreue photo-texturierte 3D-Modelle von lebensgroßen, frei zugänglichen Sandsteinfiguren erstellt werden. Mit einer Stehleiter kann nahezu die gesamte Oberfläche der behandelten Statue abgedeckt und dadurch ein nahezu vollständiges 3D-Modell erstellt werden. Kleinformen, wie z.B. ein dünner Heiligenschein, werden bei der Dreiecksvermaschung der dichten Punkwolke mit der verwendeten Software Agisoft PhotoScan nicht vollständig erhalten.

Das zweite Konzept stellt eine gute Alternative zur aufwändigeren ersten Methode dar, da durch die Verwendung der Digitalkamera Nikon D300 die Datenmenge geringgehalten wird und ebenfalls gute Ergebnisse erzielt werden können. Die Beschränkung der Passpunkte auf gut definierte natürliche Punkte bringt im Vergleich zum ersten Konzept nur eine geringfügig verminderte Genauigkeit mit sich. Bei exponierten Statuen in mehreren Metern Höhe kann die Schnittgeometrie der photographischen Aufnahmen durch die Verwendung einer Hebebühne verbessert werden, jedoch sind hier die Möglichkeiten bezüglich der Vollständigkeit und Genauigkeit trotzdem begrenzt.

Das zweite Konzept könnte für zukünftige Anwendungen bezüglich des zeitlichen und finanziellen Aufwandes noch weiter minimiert werden, indem die Maßstabsermittlung mit zumindest einer bekannten Strecke auf oder im Nahbereich der untersuchten Skulptur im Objektraum beschränkt wird. Hierfür können gut definierte Strecken, welche mittels Maßband oder handgehaltenem Distometer im Objektraum eingemessen werden können, oder kalibrierte Maßstäbe (KERSTEN & LINDSTAEDT 2017) verwendet werden, wodurch keine Totalstation benötigt wird und somit sowohl der finanzielle als auch der zeitliche Aufwand bei der Datenakquisition verringert wird. Diese Vorgehensweise eignet sich jedoch nur für frei zugängliche Skulpturen und hat den Nachteil, dass bei Verwendung von wenig Passinformation über den Objektraum das erstellte 3D-Modell zum einen potentiell Verformungen aufweisen kann und zum anderen keine umfassende Kontrolle möglich ist.

6 Literaturverzeichnis

- AGISOFT, 2018: Agisoft PhotoScan User Manual Professional Edition, Version 1.4. http://www.agisoft.com/pdf/PhotoScan-pro_1_4_en.pdf, letzter Zugriff 23.11.2018.
- CLOUDCOMPARE, 2018: User documentation CloudCompare, Version 2.6.1. https://www.cloudcompare.org/doc/qCC/CloudCompare%20v2.6.1%20-%20User%20manual.pdf, letzter Zugriff 24.11.2018.
- GEOSI VERM, 2018: Geosi VERM Datenblatt. http://www.idc-edv.at/wpcontent/plugins/pdfjs-viewer-shortcode/pdfjs/web/viewer.php?file=http://www.idcedv.at/datenblaetter/GeosiVERM_V18.pdf&download=true&print=true&openfile=false, letzter Zugriff 24.11.2018.
- KAUFMANN, V., 2018: Homepage of Viktor Kaufmann (pseudonym Victor Mercator), Graz University of Technology, Austria. https://www.staff.tugraz.at/viktor.kaufmann/, letzter Zugriff 23.11.2018.
- KERSTEN, T. & LINDSTAEDT, M., 2017: Mit Kamera, Maßstab und Laptop Best Practice zur Erstellung von 3D-Modellen für eine Museumsausstellung am Beispiel vietnamesischer Kulturobjekte. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 26, T. Kersten (Hrsg.), 37. Wissenschaftlich-Technische Jahrestagung der DGPF, Webversion, 35-44.
- LUHMANN, T., 2018: Nahbereichsphotogrammetrie: Grundlagen, Methoden und Anwendungen. 4. überarbeitete und erweiterte Auflage, Wichmann VDE Verlag GmbH, Berlin/Offenbach.
- TRARS, 2018: Tracing the Art of the Straub Family. http://trars.eu, letzter Zugriff 23.11.2018.
- YI, G., JIANXIN, L., HANGPING, Q. & BO, W., 2014: Survey of structure from motion. Proceedings of 2014 International Conference on Cloud Computing and Internet of Things, 72-76.

Räumlicher Zusammenhang zwischen Sentinel-1 Polarisationsdaten und Vegetationshöhe in einem landwirtschaftlichen Gebiet in Niederösterreich

ISABELLA PFEIL^{1,2}, MARIETTE VREUGDENHIL², PETER DORNINGER³ & WOLFGANG WAGNER^{1,2}

Zusammenfassung: Satellitenbeobachtungen, digitale Höhenmodelle und Orthophotos aus Drohnen- und Paragleitflügen sowie in situ Messungen von Vegetationsparametern werden in dieser Studie verglichen, um einen möglichen räumlichen Zusammenhang zwischen dem aus Sentinel-1-Beobachtungen abgeleiteten Cross Ratio Index und Pflanzenhöhen zu finden oder auszuschließen. Die Studie wurde im Hydrological Open Air Laboratory (HOAL) durchgeführt, einem seit 2009 bestehenden landwirtschaftlichen Untersuchungsgebiet in Petzenkirchen (Niederösterreich). Die Ergebnisse zeigen, dass das Cross Ratio zwar einem ähnlichen saisonalen Verlauf wie die Pflanzenhöhen folgt, kleinräumige Unterschiede innerhalb von Feldern aber nicht von den Pflanzenhöhen beeinflusst werden.

1 Einleitung

Weltweit werden die Auswirkungen des Klimawandels immer spürbarer, sei es in Form von Überschwemmungen, Meeresspiegelanstieg, Stürmen, Wasserknappheit oder Dürre. Auch in Österreich folgen Temperatur-Rekordjahre aufeinander, die durch Hitze und Trockenheit zu Ertragsverlusten in der Landwirtschaft führen (BUNDESMINISTERIUM FÜR NACHHALTIGKEIT UND TOURISMUS 2018). Es ist daher erstrebenswert, Entwicklungen im Wachstum des angebauten Getreides großflächig und mit hoher räumlicher Auflösung zu beobachten. Die 2014 und 2016 gestarteten Copernicus-Satelliten Sentinel-1A und -1B messen den Rückstreukoeffizienten (Backscatter) der Erdoberfläche mit einer noch nie dagewesenen räumlichen und zeitlichen Auflösung. Da die Backscattermessungen unter anderem von der Vegetation beeinflusst werden, wird es möglich, Variationen der Vegetation auf Feldebene zu analysieren. Noch ist aber nicht zur Gänze geklärt, zu welchem Teil das Backscatter von der Vegetationsstruktur und dem Vegetationswassergehalt beeinflusst wird. In dieser Studie wird untersucht, ob auch die Vegetationshöhe Einfluss auf die räumliche Verteilung des Backscatters nimmt. Dies soll dabei helfen, das von den Sentinel-1-Satelliten beobachtete Signal besser zu verstehen und interpretieren zu können, um schließlich genauere Rückschlüsse auf Änderungen im Pflanzenwachstum zu ermöglichen.

¹ TU Wien, Centre for Water Resource Systems, Karlsplatz 13, A-1040 Wien, E-Mail: isabella.pfeil@geo.tuwien.ac.at

² TU Wien, Forschungsgruppe Fernerkundung, Department für Geodäsie und Geoinformation, Gußhausstraße 27-29, A-1040 Wien,

E-Mail: mariette.vreugdenhil@geo.tuwien.ac.at, wolfgang.wagner@geo.tuwien.ac.at

³ 4D-IT GmbH, Emil Kögler-Gasse 13, A-2511 Pfaffstätten, E-Mail: p.dorninger@4d-it.com

2 Das Hydrological Open Air Laboratory (HOAL)

Seit 2009 besteht in einem landwirtschaftlichen Einzugsgebiet in Petzenkirchen, Niederösterreich, das Hydrological Open Air Laboratory (HOAL) (BLÖSCHL et al. 2016). Auf 66 ha werden hydrologische Experimente durchgeführt, mit dem Ziel, neue Erkenntnisse über räumliche und zeitliche Änderung von Parametern wie Evapotranspiration, Erosion, Sedimenttransport, Bodenfeuchte und Vegetationsdynamik und deren treibende Kräfte zu gewinnen. Die im HOAL ansässigen Landwirte bewirtschaften dabei ihre Felder gänzlich unabhängig von den durchgeführten Experimenten und dokumentieren akkurat die Art der Bepflanzung, Zeitpunkte der Aussaat und Ernte sowie Düngerausbringung und Bodenbearbeitung. Da 2017 und 2018 Winterweizen und Mais das HOAL dominierten, liegt der Fokus dieser Studie auf diesen zwei Getreidearten. Detaillierte Informationen über das HOAL sind in Tabelle 1 aufgelistet. Abb. 1 zeigt die Lage des HOAL in Niederösterreich und die Topographie im Einzugsgebiet.

	C C C C C C C C C C C C C C C C C C C
Koordinaten	48°9' N 15°9' O
Fläche	66 ha
Elevation	268-323 m
Mittlere Hangneigung	8%
Ackerland	87%
Häufigste Getreidearten	Wintergetreide (Weizen, Gerste), Mais, Raps
Andere Vegetation	Grasflächen, Laubwald, Nadelwald

Tab. 1: E	Eigenschaften	des HOAL
-----------	---------------	----------



Abb. 1: Lage des HOAL in Niederösterreich (links) und Topographie im Einzugsgebiet (rechts)

3 Datensätze

3.1 Sentinel-1

Die 2014 und 2016 gestarteten Copernicus-Satelliten Sentinel-1A und -1B bieten die Möglichkeit, Entwicklungen im Wachstum des angebauten Getreides großflächig zu beobachten (VREUGDEN-HIL et al. 2018). Horizontal und vertikal polarisierte Backscatter-Daten sind über Europa mit einer zeitlichen Auflösung von 1.5-4 Tagen und einer räumlichen Auflösung von 20 m verfügbar (BAUER-MARSCHALLINGER et al. 2018). Tabelle 2 zeigt ausgewählte Zeitpunkte, für die Sentinel-1 Backscatterdaten über dem HOAL verfügbar sind.

Tab. 2: Verfügbarkeit von Sentinel-1 Backscatter (nicht vollständig), DOM (Drohnenbefliegungen – 4D-IT) sowie DOM und Orthophotos (Paragleitflüge – AELuftbilder). Gelb hinterlegte Daten sind in weiter unten folgenden Abbildungen dargestellt.

Datum	Sentinel-1 Back- scatter	DOM (4D-IT)	DOM & Ortho- photo (AELuftb.)
2017-03-12			Х
2017-03-13	Х		
2017-04-06	Х		
2017-04-09			Х
2017-04-30	Х		Х
2017-05-22		Х	
2017-05-23			Х
2017-05-24	Х		
2017-06-04			Х
2017-06-05	Х		
2017-06-15			Х
2017-06-17	Х		
2017-06-20		Х	
2017-07-11	Х		
2017-07-12			X
2017-07-31		Х	
2017-08-04	Х		Х
2017-08-28	Х		
2017-08-29		Х	
2017-09-09	Х		X
2017-09-28		Х	
2018-10-03	Х		
2017-10-14			X
2017-10-15	Х		
2018-05-07	Х	Х	
2018-07-18	Х		
2018-07-20		Х	
2018-07-30	Х		
2018-08-03		Х	
2018-09-13		Х	
2018-09-16	Х		

3.2 4D-IT – DOM

Während der Wachstumssaisonen 2017 und 2018 wurden von der 4D-IT GmbH mittels RGB-Luftbildauswertung von Drohnenbefliegungen mit dem 4DU-CamCopter fünf (2017) bzw. vier (2018) digitale Oberflächenmodelle (DOM) des HOAL erstellt. Die Flugtage sind in Tabelle 2 aufgelistet.

3.3 AELuftbilder – DOM und Orthophotos

Zusätzlich zu den Drohnenüberflügen wurden 2017 im HOAL Paragleitflüge durchgeführt, mit dem Ziel, hochaufgelöste Oberflächenmodelle und Orthophotos des Untersuchungsgebiets zu erhalten. Beides liegt mit einer räumlichen Auflösung von 10 cm vor. Tabelle 2 zeigt die Zeitpunkte der Überflüge.

3.4 In situ Daten

Am Ende jeden Jahres werden von den im HOAL tätigen Landwirten Daten bezüglich Pflanzenart, Aussaat, Ernte, Düngerausbringung und sonstiger Bodenbearbeitung für jedes Feld erhoben. Außerdem werden seit 2016, jeweils an Tagen mit Sentinel-1-Überflug, Pflanzenhöhen und -wassergehalt an bis zu neun verschiedenen Punkten im Untersuchungsgebiet gemessen. Diese wurden in der vorliegenden Studie als Referenz für die aus den Satelliten- und Luftbildern gewonnenen Daten herangezogen.

3.5 Sentinel-1 Cross Ratio (CR)

Der Cross Ratio (CR) Index berechnet sich aus dem Verhältnis von horizontal und vertikal polarisiertem Backscatter, und kann nachweislich für die Beobachtung des Getreidewachstums und dessen räumliche Variation innerhalb eines Feldes verwendet werden (KIM et al. 2012; PALOSCIA et al. 1999; VREUGDENHIL et al. 2018). Hohe CR-Werte erhält man bei einem hohen Anteil an horizontal polarisiertem Backscatter, was charakteristisch für hohe Volumsstreuung, verursacht durch die Vegetation, ist. Allerdings beeinflussen auch Bodenrauigkeit sowie Pflanzenhöhe und Pflanzenwassergehalt den Index. Um Speckle-Effekte (Rauschen) zu vermeiden, wurden die Backscatterdaten für diese Studie mittels eines Lee-Filters geglättet.

3.6 Absolute Höhen aus DOM

Absolute Vegetationshöhen wurden aus den DOM mittels Differenzenbildung zwischen verschiedenen Zeitpunkten abgeleitet. Zusätzlich zur Verfügbarkeit der DOM ist dabei der Wachstumsverlauf der verschiedenen Getreidearten entscheidend. Allgemein werden Wintergetreide (Winterweizen und -gerste, Raps) im Herbst des vorhergehenden Jahres angebaut und im Juli geerntet. Mais wird generell im April ausgesät und je nach Wetter und Sorte von August bis Oktober geerntet. Um absolute Pflanzenhöhen zu erhalten, wurde deshalb zuerst für jedes Feld der Zeitpunkt ermittelt, an dem das Feld brach lag (entspricht Vegetationshöhe 0), und von diesem DOM anschließend die Differenz zu den während der Wachstumsphase erstellten DOM berechnet.

3.7 Vergleich der Datensätze

Der Vergleich der Datensätze erfolgte auf zwei Arten: 1) eine visuelle Analyse der räumlichen Sentinel-1 CR Daten, absoluten DOM-Höhen und Orthophotos (qualitativer Vergleich); 2) eine Analyse von Mittelwert und Standardabweichung der Sentinel-1 Daten und absoluten DOM-Höhen pro Feld (quantitativer Vergleich).

4 Ergebnisse

4.1 Qualitativer Vergleich

Anhand von Sentinel-1 CR, absoluten Höhen und Orthophotos wurde ein visueller Vergleich von räumlichen Mustern in verschiedenen Feldern des HOAL durchgeführt. Abb. 2 zeigt diese drei Datensätze für Mai 2017. In allen drei Datensätzen sind klar die einzelnen Getreidearten unterscheidbar: die im Mai noch sehr niedrigen Maispflanzen führen zu einem niedrigen CR, von dem sich die höheren CR-Werte über den Weizenfeldern deutlich abheben. Auch der Raps ist eindeutig erkennbar, der 2017 im Westen des HOAL angepflanzt wurde. Im Mai 2017, aber auch zu allen anderen Zeitpunkten sind in den Orthophotos und Höhenmodellen jene Stellen erkennbar, die von häufigem Oberflächenabfluss und damit einhergehendem spärlichen Bewuchs gekennzeichnet sind (markiert durch orange Kästchen in Abb. 2c). Auch im CR sind an diesen Stellen leichte Änderungen zu erkennen. Höhenunterschiede innerhalb eines Feldes, wie zum Beispiel im Falle des großen Feldes im Süden, auf dem 2017 zwei verschiedene Sorten Winterweizen angebaut wurden, scheinen jedoch die räumliche CR-Verteilung nicht zu beeinflussen. Abb. 3 zeigt CR, absolute Höhen und ein Orthophoto von Ende Juli bzw. Anfang August 2017. Zu diesem Zeitpunkt wurde der Weizen bereits geerntet, der Mais ist nun an die zwei Meter hoch. Das zeichnet sich durch hohe CR-Werte in den Maisfeldern aus. Obwohl der Weizen bereits geerntet wurde, sind räumliche Muster auf den Weizenfeldern erkennbar. Diese können durch verbleibende Pflanzenreste auf den Feldern erklärt werden, oder durch die Bodenfeuchte der Felder, die nun möglicherweise den CR-Index dominiert.

4.2 Quantitativer Vergleich

Um eine quantitative Analyse der räumlichen Variabilität von CR und Pflanzenhöhen durchführen zu können, wurden pro Feld zu verschiedenen Zeitpunkten Mittelwert und Standardabweichung von CR und absoluten Pflanzenhöhen berechnet. Abb. 4 zeigt den zeitlichen Verlauf dieser Metriken für ein Weizenfeld und Abb. 5 für ein Maisfeld. Zusätzlich zu den Mittelwert- und Standardabweichungszeitserien sind auch die Zeitpunkte der Düngerausbringung und Ernte (nur Weizen) angegeben. In beiden Abbildungen lässt sich ein Zusammenhang zwischen den Mittelwerten erkennen. Phasen mit hohem CR stimmen zeitlich überein mit hohen Pflanzenhöhen und umgekehrt. Im Maisfeld steigt allerdings zuerst der Mittelwert des CR an, bevor mit einer zeitlichen Verzögerung von 1-2 Wochen auch die Pflanzenhöhe steigt.

Im Fall der Standardabweichung ist ein solcher Zusammenhang nicht erkennbar. Im Weizenfeld zeigt die Standardabweichung der Pflanzenhöhen einen ähnlichen Verlauf wie der Höhen-Mittelwert, während im Maisfeld die Standardabweichung über mehrere Monate stark variiert. Ähnliche Verläufe von Mittelwert und Standardabweichung wurden auch für andere Mais- und Weizenfelder beobachtet. Diese Ergebnisse stützen die Annahme, dass die räumliche Verteilung des CR stärker von der Vegetationsstruktur beeinflusst wird als von der Pflanzenhöhe.



a) Sentinel-1 CR (2017-05-24). Einheit: [dB * 100]





b) Absolute Höhen (2017-05-22). Einheit: [m]

Abb. 2: Sentinel-1 CR (a), absolute Höhen (Differenz zwischen DOM-20170522 und DOM-20170731) (b) und Orthophoto (c). Im Orthophoto sind jene Stellen orange markiert, die von häufigem Oberflächenabfluss betroffen sind und deswegen weniger Bewuchs aufweisen

c) Orthophoto (2017-05-23). Orange Kästchen zeigen Stellen mit häufigem Oberflächenabfluss

5 Diskussion

In den obigen Abschnitten wurde gezeigt, dass ein Zusammenhang zwischen mittlerem CR und mittlerer Pflanzenhöhe pro Feld besteht. Ein Zusammenhang zwischen den jeweiligen Standardabweichungen, welche als Maß für die räumlichen Variationen der jeweiligen Parameter verwendet wurden, konnte nicht gefunden werden. Es lässt sich also sagen, dass räumliche Änderungen im CR eher von der Vegetationsstruktur beziehungsweise dem Vegetationswassergehalt als von der Vegetationshöhe ausgelöst werden. Ebenso bedeutet ein Zusammenhang zwischen mittlerer Pflanzenhöhe und mittlerem CR nicht unbedingt, dass es wirklich die Höhe ist, die das CR beeinflusst. Vielmehr ist zu den Zeiten mit großen Höhen normalerweise auch der Wassergehalt höher beziehungsweise tritt mehr Volumsstreuung auf, was zu einem höheren CR Index führt. Abb. 6 und 7 zeigen den in ausgewählten Feldern gemessenen in situ-Vegetationswassergehalt (VWC). Sowohl im Weizen- als auch im Maisfeld verhält sich der VWC proportional zum von den Satelliten beobachteten CR.

Für diese Analyse waren zwei verschiedene Höhenmodell-Datensätze verfügbar: der eine wurde mittels Drohnenbefliegungen, der andere mittels Paragleitflügen erstellt. Während sich der Drohnen-Datensatz durch hohe Genauigkeit im gesamten Untersuchungsgebiet auszeichnet, gibt es im Paragleit-Datensatz besonders an den Rändern des HOAL Ungenauigkeiten von bis zu mehreren Metern. Betrachtet man jedoch nur einzelne Felder, sind beide Datensätze gleich gut nutzbar. Ein Vorteil des in dieser Studie verwendeten Paragleit-Datensatz liegt in der besseren zeitlichen Verfügbarkeit gegenüber den Drohnenbefliegungen. Beide Methoden können nur bei gutem Wetter und möglichst wenig Wind durchgeführt werden; besonders bei den Paragleitflügen kann im Falle von starken Windböen und ungünstiger Thermik nicht nur die Technik, sondern auch der ausführende Pilot zu Schaden kommen.



a) Sentinel-1 CR (2017-08-04). Einheit: [dB * 100]



b) Absolute Höhen (2017-07-31). Einheit: [m]



Abb. 3: Sentinel-1 CR (a), absolute Höhen (Differenz zwischen DOM-20170731 und DOM-20170522) (b) und Orthophoto (c)

c) Orthophoto (2017-08-04)



Abb. 4: Mittelwert und Standardabweichung von CR und absoluter Pflanzenhöhe (Weizenfeld; Höhe abgeleitet von 4D-IT DOM). Die vertikalen Linien zeigen die Zeitpunkte der Düngerausbringung (grün) und Ernte (orange) auf diesem Feld



Abb. 5: Mittelwert und Standardabweichung von CR und absoluter Pflanzenhöhe (Maisfeld; Höhe abgeleitet von AELuftbilder DOM). Die vertikalen Linien zeigen die Zeitpunkte der Aussaat (braun), Düngerausbringung (grün) und Ernte (orange) auf diesem Feld



Abb. 6: Verlauf von Sentinel-1 CR und in situ-VWC in einem Weizenfeld (Wachstumsperiode 2017)



Abb. 7: Verlauf von Sentinel-1 CR und in situ-VWC in einem Maisfeld (Wachstumsperiode 2017)

6 Fazit

Über zwei Jahre (2017 und 2018) wurden in einem landwirtschaftlichen Versuchsgebiet in Petzenkirchen (Niederösterreich) Drohnenbefliegungen und in situ-Messungen der Vegetation durchgeführt. Anhand dieser Luftbild- und in situ-Datensätze wurde in dieser Studie untersucht, inwieweit das von den Sentinel-1-Satelliten gemessene Backscatter, insbesondere der daraus abgeleitete Cross Ratio Index, von der Pflanzenhöhe beeinflusst wird. Dazu wurde ein qualitativer Vergleich von Orthophotos mit den räumlichen Mustern des CR und der aus digitalen Oberflächenmodellen abgeleiteten absoluten Pflanzenhöhen angestellt. Zusätzlich wurden Zeitserien der Mittelwerte und Standardabweichungen aller pro Feld gemessenen CR- und Höhenwerte erstellt und verglichen. Während zwischen den Zeitserien der Mittelwerte durchaus ein Zusammenhang besteht, scheinen die räumlichen Muster von CR und absoluten Pflanzenhöhen voneinander unabhängig. Das bekräftigt die Vermutung, dass der CR Index empfindlicher auf die Pflanzenstruktur und den Pflanzenwassergehalt als auf Pflanzenhöhen ist.

7 Literaturverzeichnis

- BAUER-MARSCHALLINGER, B., FREEMAN, V., CAO, S., PAULIK, C., SCHAUFLER, S., STACHL, T., MODANESI, S., MASSARI, C., CIABATTA, L, BROCCA, L. & WAGNER, W., 2018: Toward Global Soil Moisture Monitoring with Sentinel-1: Harnessing Assets and Overcoming Obstacles. IEEE Transactions on Geoscience and Remote Sensing, (99), 1-20.
- BLÖSCHL, G., BLASCHKE, A., BROER, M., BUCHER, C., CARR, G., CHEN, X., EDER, A., EXNER-KIT-TRIDGE, M., FARNLEITNER, A., FLORES-OROZCO, A., HAAS, P., HOGAN, P., KAZEMI AMIRI, A., OISMÜLLER, M., PARAJKA, J., SILASARI, R., STADLER, P., VREUGDENHIL, M., WAGNER, W. & ZESSNER, M., 2016: The Hydrological Open Air Laboratory (HOAL) in Petzenkirchen: a hypothesis-driven observatory. Hydrology and Earth System Sciences, 20(1), 227-255.
- BUNDESMINISTERIUM FÜR NACHHALTIGKEIT UND TOURISMUS, 2018: Grüner Bericht 2018 Bericht über die Situation der österreichischen Land- und Forstwirtschaft.
- KIM, Y., JACKSON, T., BINDLISH, R., LEE, H. & HONG, S., 2012: Radar vegetation index for estimating the vegetation water content of rice and soybean. IEEE Geoscience and Remote Sensing Letters, 9(4), 564-568.
- PALOSCIA, S., MACELLONI, G., PAMPALONI, P., & SIGISMONDI, S., 1999: The potential of C- and L-band SAR in estimating vegetation biomass: the ERS-1 and JERS-1 experiments. IEEE Transactions on Geoscience and Remote Sensing, 37(4), 2107-2110.
- VREUGDENHIL, M., WAGNER, W., BAUER-MARSCHALLINGER, B., PFEIL, I., TEUBNER, I., RÜDIGER, C. & STRAUSS, P., 2018: Sensitivity of Sentinel-1 Backscatter to Vegetation Dynamics: An Austrian Case Study. Remote Sensing, 10(9), 1396.

Multi-Temporal Aerial Imagery for Automatic Mapping and Trend Analysis

ROLF BECKER¹

Abstract: The enormous number of existing aerial photographs - partly also satellite data - document the development of our environment for about 100 years, verifiable, unbiased and free from any ideological or political influence. Development trends are identifiable, not only for specific, isolated areas, but also regarding the reciprocal interference of various trends whose potential effects are only revealed by them. A comprehensive analysis of time-phase aerial images can significantly enhance our knowledge of the development and their causes over the last 100 years, evidence that should not be neglected, but on the contrary help us to predict up-coming events and their consequences. Considering the increasingly complicated environmental issues, this paper to calls attention to the potential contribution of aerial photography for trend identification and analysis globally and with specific attention to periods and areas where Satellite images are not available.

1. Introduction

Newton's 1st Law of Motion: "Every object in a state of uniform motion tends to remain in this state unless an external force is applied to it." This law applies to everything. Left by itself, action, procedures, development, etc. follow a trend that may turn into a routine, a habit. The law could be rewritten as 'Everything follows a trend as long as it is not influenced by another trend".

Trends are the baseline of our actions and ambitions, they define the current norm. Trends can be followed, set, modified or whatsoever, but above all, they must not be ignored. Trends can be dangerous and deceiving, in particular if they have not been identified. Having not come to our attention, they lull us into complacency and make us blind to lateral thinking. But equally, trends if identified point to where we will end up, right or wrong and by understand the consequence of past actions help up make the correct decisions going forward.

Trends must be identified, their origin, their purpose and their consequence. This calls for Trend Awareness! We must search for them, analyse them and find out how they influence each other, and to what extent.

Trend Analysis either conscientiously or sub-conscientiously provides us with hindsight, an indispensable element for prediction, which is part of everything we engage in. Trends can be difficult to detect or identify as they do not happen but develop. They start almost imperceptibly and have a tendency to grow on us. By the time they become apparent, their origin might be no longer traceable. They get turned into a routine and are simply neglected as 'a matter of fact!'. By then, it might be too late to reverse, or even stop a trend.

Trends are identified by analysing some quantifiable measurements over time. In this way we gain evidence of slow-moving action or changes that under normal conditions appear to be static. Likewise, multi-temporal (time-series) aerial photography can provide these data points may so

¹ Geodyn Technology, Suite 623, Building 6EA, DAFZA, Dubai, UAE, E-Mail: rolf.becker@geodyn.com

far have not been identified, nor examined. The analysis of this photography provides a unique possibility to identify a trend curve, pointing to its origin, as well as to the time to come! Prediction is the ultimate decision-making element. It accompanies us from the perception of an activity to its final (eventual) implementation (realization). Prediction need is to be preceded by Hindsight. As are cannot foresee the future, we must predict the most likely scenario. This applies equally when we go on holiday, choose the profession we want to train for, or design the water supply for a new township. Prediction requires to review the past in order to identify development trends and assess by what they might be influenced by. Trend Analysis is an important prerequisite for predicting the circumstances prevailing at the time a projected measure is to take effect. It provides an important basis for planning and development projects. Prediction is an essential, all-encompassing planning element, needed right up to the time a decision must be taken, or action to start. Prediction must not simply rely on reports, but foremost on reliable, confirmable and unprejudiced evidence. In order to looks ahead, we need to assess trends and actions in the light of current awareness.

2. Aerial Survey, Photogrammetry

An enormous number of existing aerial photography – and lately also satellite imagery – trace the development of our environment for over 100 years, impartially and free of ideological or political influences. In addition, they are a testimony (the witness) to our achievements, the ups and the downs. From these imagery trends can be identified, not only for specific, isolated events, but they reveal comprehensively the reciprocal effects and influences of actions that have been taken. Aerial Photography documents the state of environment at the time the photography was taken in an irrefutable way. This can be done with higher accuracy and reliability, than by any comparable mean. Moreover, Aerial photography depict object in their natural surroundings, which greatly enhances their identification and purpose. Aerial Photography shows details in their relative position. Photogrammetry can map this detail with great accuracy.

We can map the environment as it existed at the time the photograph has been taken. This does not only apply to physical objects, like houses, streets, fences, but can also identify objects or conditions derived from their relative position and neighbourhood, e.g. railway station, parks, industrial/ residential/ commercial zones, flooding, avalanche protection measure, water shortage, noise control measures, danger area of all sorts, accessibility and much more. All these elements that can be extracted form aerial imagery and managed in a GIS.

Much of what we would like to know today had not been measured or predicted previously. Although much detail was extracted into maps by photointerpretation from the original imagery, only a fraction of the information in the imagery was extracted. Moreover, the maps created were application specific with limited details and much of this data has been lost, mostly because it was not digitally recorded or maintained.

The concept of using multi-temporal aerial photography is not new. However, up to now, the required data have not been available nor accessible, at least not to the extent as is required to reach conclusive results. In view of the enormous data quantities required, only computerized analysis achieve the trend identification.

Worldwide, a conservative estimate of number of aerial photography frames is hundreds of millions. For instance, the US Department of Agriculture holds in its archives over 40 million frames, and the US National Security Agency (NSA) holds an archive of over 25 million. The U.S. National Archives (NARA) holds multiple agency-specific archives, including a globally-distributed archive of 28.5 million trimetrogon photo triplets, an estimated half of which were captured by the U.S. Defence Intelligence Agency between 1940 and 1945. Most photos were captured at 1:40,000 scale and cover 10 km wide swaths (Fig. 1, LEFSKY 2018). In most counties in the world the state and national mapping agencies also have archives collected at least decadal frequency covering their lands.



Fig. 1: Section of NARA Photo Index (Approx. 250 frames out of 28 million)

Converted into digital form this data alone would let us to look at how our world has change over the last 80 years. In combination with satellite imagery, this would provide us indisputable evidence for answering question, for instance in such much disputed topic of climate changes. Aerial Photography also contributes to document a Nation's Heritage. For its safeguard, it is essential that this data become accessible, something that can only be assured thought converting them into a digital format. However, the bulk of this aerial photography is presently being stored in archives all over the World, still awaiting to be converted into a digital format. To date the lossless conversion cost of about \$10 per frame has made creating a digital version an impossible undertaking. The required equipment for such conversion has become scarce and no longer available for bulk processing.

Regardless of any immediate need, it is unconceivable that this data remains inaccessible. Already many of these archived images have being lost due to organizations not realizing their value or not being able to adequately maintain the film in the required large cooled spaces. In the foreseeable future the necessary personal with the skills to handle the film might no longer be in a position to do so. Apart from, that with age some aerial film deteriorates to the stage where they can no longer be safely handled, e.g. the Cellulose Nitrate Film which over time potentially become explosive. About 30 years ago, with the introduction of digital photogrammetry, there was a considerable drive for digitizing of aerial photography to enable digital stereo compilation of the films as digital cameras were not performant enough. Since about 15 years, digital cameras and satellite imagery dominate spatial data acquisition, pushing the need for analogue/digital conversion for operational purposes into the background. This freed the established digitizing capacity to digitize the film from the archives, but it soon became apparent, that the established scanning capacity was insufficient to cope with bulk digitizing of archived aerial films. This was aggravated by (a) the slow conversion rates of 3 to 6 frames per hour, (b) photogrammetric scanners no longer being produced, (c) the raising maintenance costs of the few remaining photogrammetric quality digitizer. Today it is estimated to be less than 30 photogrammetric scanned still in production worldwide. This has made analogue to digital conversion of aerial film even less affordable. In addition, procedures for automatic georeferencing and data extraction from aerial photography were not in sight making the digital images hard to use. This has led to most organizations discontinued to systematically digitize their Aerial Photography Archives. This situation prevails in countries almost all over the World. As a consequence, estimate over 90% of all aerial films still await to be digitized.

3. By today, this scenario has drastically changed

3.1 Analogue/digital Conversion

In spring 2019, after a development of over three years, GeoDyn is releasing Prompt Capture GenTwo. This new equipment and procedures shorten the time required for Analogue/Digital conversion 50-fold and reduces the cost by an order of magnitude. This will provide the economic ability to digitize millions of frames per year with a resolution of 15micron, high geometric accuracy and 14bit dynamic range where the full information content of aerial film can be permanently captured and made accessible. This alone can be considered a paradigm shift.

3.2 Automated Georeferencing

Technological development in massively parallel image matching and bundle block adjustment, the existence of global, accurate high-resolution image basemaps and digital terrain models along with refined procures now enable efficient determination of the interior and exterior orientation of all imagery, even when only limited photo indexes exit. By combining utilizing the correlation between multi-temporal imagery that relative and absolute accuracy of the georeferencing is improved and can be iteratively refined. Inexpensive cloud storage and infrastructure has made storage and dissemination of imagery affectable with the cost of storage continuing to drop

3.3 Time Series Image Maps



Fig. 2. Multi-temporal imagery and trend analysis of Lafto. Courtesy: Ethiopian Mapping Agency

The georeferenced imagery can be served through the web as dynamic image service that provides access to the full information content and metadata of the imagery. These time-series of digital image maps, provide detail-matching accuracy that can be used for interdisciplinary trend analysis and to evaluate the extent of mutual interference (Fig. 2). These will feed a wide range of Apps for instant review and analysis of the imagery access to the full information content and metadata of the imagery. These time-series of digital image maps, provide detail-matching accuracy that can be used for interdisciplinary trend analysis and to evaluate the extent of mutual image maps, provide detail-matching accuracy that can be used for interdisciplinary trend analysis and to evaluate the extent of mutual interference. These will feed a wide range of Apps for instant review and analysis of the imagery.

3.4 Automatic (vector) Image Classification and Object Detection from Aerial Imagery

It is to be foreseen that within a few years procedure will become available for the automatic feature identification using Artificial Intelligence. The rapid advance of deep learning for spatial application is resulting in machine learning that can automate the both the identification of objects from scanned aerial imagery as well as change detection and the inference of land use. Mapping organizations that have maintained maps captured from the imagery can utilize the exiting mapping as training data to enable the algorithm to quickly learn and classify feature with higher accuracy. The results are attributed temporal (i.e. Time-phased) map layer, providing interdisciplinary spatial data in GIS, thereby applicable in many conceivable domains. All features are referenced spatially, i.e. geographically as well as temporally. Thus, a complete inventory of all objects can be created automatically. This procedure does not only identify physical objects, compounded (derived) definitions, such as: danger zone, residential area, orchard, vine yard football field, railway station, areas subject to inundation, etc.

4. Spatial Insights

BigData processing along with improvements in spatial data analytics are enabling the rapid trend detection modelling and prediction. The techniques being developed enable the detection of trends and anomalies. By combining the millions of observations with many other geospatial variables such a distance to roads and towns, terrain form, distance to transportation corridors, weather and seasons, the algorithm employs machine learning to determine correlations between possibly unforeseen events to determine the trends, the change in the trends and the influencers of these trends. Based on the huge data volume, a very high statistical accuracy can be obtained. These models can then be extrapolated to predict future state as well the influence newly planned actions may have.

5. Conclusion

Huge amounts of aerial photography, recently complemented by satellite imagery, can provide irrefutable evidence of developments of the past and thereby improve on the prediction for the near future. This becomes possible not only because of advanced in photogrammetric procedures and image analysis involving AI and Mechanical learning, but notably, because new analogue/digital image conversion technology now overcomes the bottleneck in the conversions of the huge number of aerial photographs held in archives all over the World. With emphasis on the climate change problematic it is recommended that the Photogrammetric Community involve itself in this thematic.

At this occasion, I am not addressing this audience as Austrians, Germans or Swiss, not even as Europeans, but as a Community that has the mandate to maintain the authoritative data sources that our nations have spent billions on. As a community we can contribute through Aerial survey in a joint effort to work towards the UN Sustainable Development Goals (SDG) as per Agenda 2030 adopted by all UN Members in 2015 and reiterated during the United Nations Geospatial Information Congress last December in Deqing/China. Our generation owes it to the next to ensure

all that has happened since the turn of the twentieth century is appropriately recorded so we can learn from the past helping us to make better decisions moving forward.

Digital, multi-temporal (time-series) images from Europe, the US, the Middle East, and many African countries are featured in the article, and may be provided interested parties to prospective students for study purposes (images samples will become accessible over the net).

6. Bibliography

- LEFSKI, M., 2018: Global Index RG_373, Colorado State, Personal interview and data exchange concerning the U.S. National Archive Record Group 373, Archives of the Defence Intelligence Agency, December 1st.
- NECSOIU, M., DINWIDDIE, C.L., WALTER, G.R., LARSEN, A., STOTHOFF, S.A., 2013: Multi-temporal Image Analysis of Historical Aerial Photographs and Recent Satellite Imagery Reveals Evolution of Water Body Surface Area and Polygonal Terrain Morphology in Kobuk Valley National Park, Alaska. Environmental Research Letter, **8**, 07-25.
- NOCERINO, E., MENNA, F. & REMONDINO, F., 2012: Multi-Temporal Analysis of Landscapes and Urban Areas. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **39**(B4), 85-90.
- ATIAS, P., KAIMARIS, D. & STYLIANIDIS, E., 2011: Change Detection in Historical City Centers using Multi-source Data: The Case of Historical Center of Nicosia – Cyprus. 23rd Int. CIPA Conference, Prague, Czech Republic, 6 pages.
- RATCLIFFE, I.C. & HENEBRY, G.M., 2004: Using Declassified Intelligence Satellite Photographs with Quickbird Imagery to Study Urban Land Cover Dynamics: a case study from Kazakhstan. Annual Proceedings ASPRS, **198**, 1-10.
- SEVARA, C. & VERHOEVEN, G., 2018: Surfaces from the Visual Past: Recovering High-Resolution Terrain Data from Historic Aerial Imagery for Multitemporal Landscape Analysis. J Archaeol Method Theory, 25, 611-642.
- TARANTINO, C., CASELLA, F., ADAMO, A., LUCAS, R., BEIERKUHNLEIN, C. & BLONDA, P., 2019: Ailanthus Altissima Mapping from Multi-temporal very High Resolution. ISPRS Journal of Photogrammetry and Remote Sensing, 147, 90-103.

UAV-Based Multispectral Data for Tree Species Classification and Tree Vitality Analysis

MAX KAMPEN¹, STEFAN LEDERBAUER², JAN-PETER MUND³ & MARKUS IMMITZER²

Abstract: The presented research work tested the suitability of remote sensing via unmanned aerial vehicles (UAV) and multispectral data as a low-cost solution for forestry applications. High-resolution 6-band multispectral aerial image sequences obtained via UAV acquisition flights in three research areas in Austria were photogrammetrically processed to coherent datasets by the use of Structure-from-Motion technology. Resulting orthomosaics, digital surface and digital terrain models in resolutions from 2.5 - 5 cm were utilised in the subsequent object-based image analysis combined with a Random Forest classification modelling approach. The methodology facilitated tree species classification and vitality analysis with additional predictive mapping.

Our study delivered promising results for the application of UAV-based multispectral datasets in operational forestry. The developed methodology performed reliably in species classification and despite inherent research issue-related uncertainties showed great potential in the detection of differences in tree vitality.

1 Introduction

Sustainable management of natural environments requires spatial information in high temporal and spatial resolution. Detailed information on tree species distribution and forest health represent essential parameters. Time-consuming terrestrial surveys are not able to provide near-term information on large scale and are often related to high costs. Forest management therefore widely relies on the supporting implementation of conventional remote sensing technology. Data obtained by earth observation satellites and aerial flight surveys are frequently used in forestry applications and are adequate for numerous research objectives (FASSNACHT et al. 2016; IMMITZER & ATZBERGER 2014; IMMITZER et al. 2012, 2018; IMMITZER et al. 2016a; IMMITZER et al. 2016b; LAUSCH et al. 2013). However, the mostly coarse resolution of freely available satellite imagery of earth-orbiting satellites as well as complex planning and high costs of airplane-based surveys restrict their applicability for several investigations and research approaches. Furthermore, strong weather-dependency, for example in recurrent overcast conditions, diminishes the reliability of data supply.

In times of climate change, flexibility is of crucial importance when monitoring forest disturbances like bark beetle (IMMITZER & ATZBERGER 2014) or fungus infestations (KIRISITS & FREINSCHLAG

¹ University of Sustainable Development (HNE), Schicklerstraße 5, D-16225 Eberswalde, E-Mail: max.kampen1988@gmail.de

² University of Natural Resources and Life Sciences (BOKU), Institute of Surveying, Remote Sensing and Land Information – IVFL, Peter-Jordan-Straße 82, A-1190 Vienna,

E-Mail: stefan.lederbauer@boku.ac.at & markus.immitzer@boku.ac.at

³ University of Sustainable Development (HNE), Faculty of Forest and Environment, Schicklerstraße 5, D-16225 Eberswalde, E-Mail: jan-peter.mund@hnee.de

2012) in order to develop mitigation strategies and initiate timely countermeasures. Changing climatic conditions and resulting severe weather events increase the vulnerability of trees (FACCOLI 2009) and thus infestation risk (FAHSE & HEURICH 2011; IMMITZER & ATZBERGER 2014; LINDNER et al. 2010). The development of suitable monitoring methods for the localisation of stressed and hence potential host trees or outbreak epicentres has gained largely in importance.

Gathering information about tree species distribution in forest management systems that approximate close-to-nature silvicultural practices is of utmost interest. The development towards ecologically and politically preferred heterogeneous mixed forests over pure stands entails traditional field-based assessment methodologies for tree species distribution, like stand estimation, as inapplicable. Generating operational forest inventories of these mixed forests requires cost-effective and accurate acquisition of detailed information on species distribution (GOODBODY et al. 2017).

Numerous scientific publications indicate the introduction of a flexible low-cost solution in unmanned aerial vehicle-based (UAV) multispectral data acquisition and processing for multitemporal monitoring of forest disturbances and species distribution mapping (FRANKLIN & AHMED 2017; FRANKLIN et al. 2017; GOODBODY et al. 2017; LEHMANN et al. 2015; LISEIN et al. 2015; MICHEZ et al. 2016; NÄSI et al. 2015; NEVALAINEN et al. 2017; TORRESAN et al. 2016). Application of UAV requires comparatively low planning efforts, as they are portable and easily deployed for data acquisition under suitable weather conditions and various terrains. Multispectral camera systems covering photosynthetically active wavelength spectra of visible light, as well as near infrared (NIR) parts of the electromagnetic spectrum in sub-decimetre resolution are highly suitable for the characterisation of vegetation and therefore described as an increasingly viable forestry application (GOODBODY et al. 2017; TORRESAN et al. 2016). Furthermore, they enable calculation of classification-enhancing raster layers like spectral band ratios, vegetation indices, variance-based metrics (e.g. PCA – principal component analysis) or structural indices.

The study aims to develop an object-based Random Forest (RF) modelling approach for tree species classification and vitality analysis based on UAV-based multispectral imagery from three investigation sites in Austria. Image datasets were obtained from an European ash (*Fraxinus excelsior*, L.) seed plantation where the fungal pathogen *Hymenoscyphus pseudoalbidus* causes ash dieback disease, a stand of Norway spruce (*Picea abies*, (L) Karst.) with a monitored infestation by the European spruce bark beetle, and a heterogeneous mixed forest stand. We explore suitability and limitations of obtained image datasets and evaluate the applicability of multispectral image data and applied methodology in regard to respective research questions:

- 1) Are differences in vitality of individual European ash trees detectable using multispectral UAV-imagery and applied methodology, and how well do the results explain reference data from a ground-based damage assessment?
- 2) Is it possible to reliably differentiate infested from not infested Norway spruce trees with the multispectral UAV-imagery and applied methods?
- 3) Is a sufficiently accurate tree species classification and distribution mapping feasible on the basis of present image data and developed methods?

2 Methodology

The workflow of the presented research work is divided into different thematic sections, such as data acquisition via UAV flights, photogrammetric processing of the obtained image sequences, and object-based image analysis with RF classification which is finally used for predictive mapping (Fig. 1).



Fig. 1: General workflow showing main steps of the conducted research project

2.1 Investigation Areas

The investigated seed plantation of European ash was established close to Feldkirchen an der Donau, a municipality in the Austrian state of Upper Austria. The area is situated at 48°20'13.4"N, 14°02'50.2"E and has an average altitude of 270 m above the Adriatic (m AA). Averaged annual rainfall quantity stretches from 53.5 mm in February to a maximum of 124.9 mm in July. The flat area is characterised by hardwood floodplain forests along the river Danube in the south and the mountainous region of Mühlviertel in the north.

Decreasing tree population size due to ash dieback disease caused by the fungal infection with *Hymenoscyphus pseudoalbidus* is a known and widespread problem in Austria since it was first reported in 2005 (HALMSCHLAGER & KIRISITS 2008). The severity of the disease is ubiquitous and threatens the future of the European ash as an economically and also ecologically valuable tree species (KIRISITS & FREINSCHLAG 2012). Concerned with the possible extinction of the European ash, researchers established several seed plantations to conduct research on clones of individual ash trees with desirable traits for timber production, which could express signs of natural genetic resistance towards the pathogenic fungus.

The Norway spruce stand (48°53'12.1"N, 15°09'48.6"E), infested by the European spruce bark beetle, and the mixed forest stand proposed for tree species classification (48°52'32.6"N, 15°09'13.5"E) are situated near Heidenreichstein in the northwestern part of the Waldviertel region within the Austrian state Lower Austria. Located at an average altitude of 620 m AA, bedrocks of

the plateau region are mainly base-poor silicates and prevalent climate conditions are considered continental (PROVINCIAL GOVERNMENT OF LOWER AUSTRIA 2015). According to ZAMG (2018), annual mean precipitation in the Waldviertel region only amounts to 400 mm with a maximum of 100 mm in July. Due to low precipitation and temperature increase in spring and summer, spruce stands in this area are endangered by insect pests like European spruce bark beetle or Spruce wood engraver.

Before countermeasures in form of tree extractions were carried out, reference data on infested and healthy spruces were collected for the purpose of testing detectability of differences in tree vitality with the acquired multispectral images for the whole stand (N = 652).

In the mixed forest stand, Norway spruce and Scots pine (*Pinus sylvestris*, *L*.) can be regarded as the predominant species. European silver fir (*Abies alba*, *Mill*.) is also occurring in large numbers, whereas European beech (*Fagus sylvatica*, *L*.) and European larch (*Larix decidua*, *Mill*.) were present with 13 and 9 individuals respectively. Multispectral images of the area were recorded for the development of a tree species classification and segmentation approach for small-scale species distribution mapping.

2.2 Data Acquisition

We utilised a modified DJI S900 lightweight multirotor UAV equipped with the multispectral camera AIRPHEN (Focal length: 8 mm; Resolution: 1280 x 960 pixel), comprising six synchronized global shutter sensors arranged in a 2 x 3 array for multispectral image acquisitions. Recording wavelengths were predefined at 450 nm (BLUE), 570 nm (GREEN), 675 nm (RED), 710 nm (RED EDGE 1), 750 nm (RED EDGE 2) and 850 nm (NIR). A linked GPS antenna added positional information to the image metadata and facilitated and enhanced subsequent steps in the photogrammetric image processing. Flight planning ensured a front-overlap of 80 % with a sideoverlap of 40 % of the image size. Images were captured in an interval of 2 seconds at a flight speed of 3 m/s (KRAUSE et al. 2016). Spectral reference measurements were taken with a calibrated PSR-2500 full range portable spectroradiometer using differently coloured felt targets in white, black, light grey, dark grey, light green and dark green. Calibration data was used for correcting the obtained imagery to reflectance values during the photogrammetric image processing.

2.3 Photogrammetric Processing

Photogrammetric image processing was realised in Agisoft PhotoScan Professional 1.3.2, developed by Agisoft LLC, St. Petersburg, Russia. This software provides an automated image processing pipeline for aerial imagery using Structure from Motion technology which is widely applied in recent forestry-related research (DANDOIS & ELLIS 2013; LEHMANN et al. 2015; LISEIN et al. 2015; MOHAN et al. 2017; NÄSI et al. 2015; TOMAŠTÍK et al. 2017; TORRESAN et al. 2016). The image sequences were processed into coherent datasets, providing structural as well as spectral information in the form of a digital surface model, a digital terrain model and reflectance orthomosaics of each investigation area.

2.4 Object-Based Image Analysis (OBIA)

Processing and modelling procedures for the OBIA workflow have been developed within the open-source environment of RStudio (Version 1.0.153) (RSTUDIO TEAM 2016) using the open-

source statistical programming language R (Version 3.4.3) (R CORE TEAM 2017). Map products of this study were created using ESRI ArcGIS® ArcMap 10.3 software, which was also used for digitising procedures.

Image analysis and modelling approaches were developed (Fig. 2) using the following R packages: *randomForest v4.6-12* (LIAW & WIENER 2002); *raster* (HIJMANS 2017); *caret* (KUHN et al. 2017); *matrixStats* (BENGTSSON 2017). Additionally, the open-source software Orfeo ToolBox (OTB Version 6.2.0) (MICHEL et al. 2015) was utilised via system commands in RStudio for tree crown segmentation trials with Large-Scale-Mean-Shift (LSMS) algorithm and the calculation of Haralick textural features.



Fig. 2: OBIA workflow showing detailed steps of image data processing with reference to applied software packages

2.5 Training Polygons & Segmentation

Individually selected trees were manually delineated on colour-infrared orthomosaics and digitised in ESRI ArcMap 10.3. Resulting crown polygon shapes, including illuminated as well as shadowed crown parts, served as masks for the ensuing OBIA feature extraction process. Branches in lower tree levels or transition areas to adjacent trees were excluded.

Sunny conditions during data acquisition at the species classification site and resulting directional shadowing effects generated large inhomogeneity in radiometric values at crown level. To account for expected misclassifications, a mixed shade class was introduced, which considered independent of the surrounding tree species. Additionally, trees were delineated by using Large-Scale Mean-Shift (LSMS) segmentation in an attempt to automatize ensuing predictive mapping. The LSMS segmentation from the Orfeo ToolBox (MICHEL et al. 2015) was used via system commands in Rstudio. Numerous other vegetation-based remote sensing applications using the LSMS method (EINZMANN et al. 2017; IMMITZER et al. 2016b; NG et al. 2016, 2017) already presented excellent results.

2.6 Feature Layer Generation & Data Extract

Based on the generated multilayer orthomosaics, several additional feature layers were calculated. Camera band settings allowed for spectral enhancement through calculation of 5 band ratios, 13 vegetation indices, as well as principal components (PCA) and Haralick textures that delivered additional datasets for enhanced classification procedures.

Tree crown polygons created for training and prediction (segmentation and manual digitizing) were utilised in the feature extraction using the *extract* function embedded in the *raster* package for R. Orthomosaics, feature layers from vegetation indices, as well as from PCA and Haralick texture feature calculation served as input for a data processing function that extracts data at locations with intersecting data and tree polygons. Based on extracted values, the following statistical measures were calculated for each individual tree polygon: *Mean, Standard Deviation, Median, 75th, 90th* and *95th Percentile.* Compiled in a cumulative table, the derived statistics were used as a feature input dataset for the RF classifier.

2.7 Random Forest Classification

For the object-based classification of the extracted data, a RF (BREIMAN 2001) classifier was utilised. A recursive feature selection approach, which uses a stepwise elimination of the least important input features based on the particular *Mean Decrease Accuracy* ranking (importance measure), was applied (IMMITZER et al. 2016a; SCHULTZ et al. 2015; TOSCANI et al. 2013). In conclusion, the model with the lowest number of input features, reaching a minimum of 97.5 % of the maximum overall OOB (out-of-bag) accuracy, was declared as the best model (EINZMANN et al. 2017).

3 Results & Discussion

3.1 European Ash Seed Plantation

Five different ash dieback intensities (0 = no dieback, 1 = 0 - 5 %, 2 = 5 - 20 %, 3 = 20 - 50 %, 4 = 50 - 80 %) were classified with overall accuracies (OA) of 61.7 % (Tab. 1) in the European ash seed plantation. This result underperforms the overall accuracy of 71.0 % presented in another publication analysing ash damage class classification based on WorldView-2 data (WASER et al. 2014). Spectral in-class variability hampered the achievement of reliable classification results in the modelling procedure, which entails low correlation with the available reference data. It remains unclear if reference data from the ground-based assessment was inadequate for analysing differences that result from spectral datasets, or if the remotely sensed differences are related to other factors like intraspecific variability, thus such results do not reflect the damage intensity professionally assessed in the field. Therefore a final evaluation of our findings is difficult and requires further research.

Reference Data								
	CLASS	0	1	2	3	4	Total	User's accuracy
	0	19	9	0	1	0	29	65.5%
	1	21	79	15	4	0	119	66.4%
Classification	2	0	5	4	4	1	14	28.6%
	3	0	2	4	2	1	9	22.2%
	4	0	0	0	0	4	4	100.0%
	Total	40	95	23	11	6	175	
Producer's accuracy		47.5%	83.2%	17.4%	18.2%	66.7%		
Overall accuracy		61.7%						
	Карра	0.338						

Tab. 1: Confusion matrix of the best-performing RF model for ash damage classification

The binary logistic regression model for the differentiation of trees with damage rates below or above 5 % delivered satisfying results with an OA of 89.1 % (Tab. 2). Feature selection indicated good separability of the damage classes when utilising reflectance and Haralick texture measures based on RED and both RED EDGE spectral bands, as well as the Red Edge NDVI.

(**************************************								
		Reference	Data					
	CLASS	0	1	Total	User's accuracy			
C1 .c	0	130	14	144	90.3%			
Classification	1	5	26	31	83.9%			
	Total	135	40	175				
	Producer's accuracy		65.0%					
Overall accuracy		89.1%						
	Карра	0.666						

Tab. 2: Confusion matrix of RF model used in binary logistic regression (0 = damage levels below 5 %; 1 = damage levels above 5 %)

In reference to the first research question, differences in the vitality of individual European ash trees are very well detectable in the multispectral data. The detected differences do not correlate well with the in-field reference data from ground-based damage assessments of ash dieback.

3.2 European Spruce Bark Beetle Infestation

Results from the Norway spruce stand with a previously detected infestation with the European spruce bark beetle showed good separation of two reference-based training classes for *infested* and *not infested* individuals. The best Random Forest model reached an OA of 86.6 % with constant user's and producer's accuracies (Tab. 3). Thus it was subsequently used for predictive mapping (Fig. 3). Most features selected in the RF model process were based on red, red edge and NIR wavelengths, which underlines the importance of these spectral wavelengths for the detection of differences in vitality between individual trees (JONES & VAUGHAN 2010; LILLESAND et al. 2014; MASAITIS et al. 2013; OLLINGER 2011).

Reference Data							
	CLASS	Not Infested	Infested	Total	User's accuracy		
Classification	Not Infested	35	5	40	87.5%		
Classification	Infested	6	6 36		85.7%		
	Total	41	41	82			
1	Producer's accuracy	85.4%	87.8%				
Overall accuracy		86.6%					
	Kappa	0.732					

The majority of marked *infested* individual spruce trees are located along the main roads and a skidding trail at the northeastern border of the stand. This is likely related to a drainage canal alongside the road and the influence of road clearing measures in winter. The central area is predominantly classified as *not infested* with occasional *infested* tree clusters.

A verification of the predictive map (Fig. 3) was performed through professional in-field evaluation of the Norway spruce stand. In the entire stand, there were no further signs of bark beetle activity detected, which suggests a successful prevention measure of bark beetle propagation applied in the stand earlier. The formerly infested southwestern corner of the stand was completely cleared as a preventive countermeasure to avoid further dispersal of *Ips typographus* throughout the stand.



Fig. 3: Predictive distribution map for infested and not infested Norway spruce

During an in-field inspection, several characteristics that could explain signs of impaired fitness were detected. We discovered two marked trees with signs of mechanical bark damage and external resin flow, whereas surrounding trees that were classified as healthy did not show any signs of disturbance. Furthermore, three trees classified as *infested* have probably been naturally coppiced (top rupture or break of the central shoot) at a young life stage which led to growing of two main trunks as a result. Marked trees in other areas did not show any signs of stress that would be observable in a ground-based assessment. The extraction of two spruces with sparse crowns and one-directional crown growth for a professional examination of the bark and debarked wood did not show any signs of insect pest activity.

The detected tree stress-inducing conditions might affect reflectance characteristics at tree crown level, but assigning impaired fitness levels of tree crowns to certain disturbance factors remains difficult. Applying the designed and tested tree selection method, no infested tree remained undetected.

3.3 Tree Species Classification

Beyond the tested detection of infestations and spectral signals of diseases, our research delivered very satisfactory results for the individual species classification and spatial distribution mapping. Species differentiation methods developed in this study, based on UAV imagery, demonstrated plausible and promising results with a high OA of 95.8 % (Tab. 4). The lower producer's accuracy for European larch is likely due to the small sample size of the class.

	-							
Reference Data								
CLASS	SPRUCE	PINE	LARCH	FIR	BEECH	SHADE	Total	User's accuracy
SPRUCE	29	1	2	1	0	0	33	87.9%
PINE	0	29	0	0	0	0	29	100.0%
LARCH	0	0	6	0	0	0	6	100.0%
FIR	0	0	0	29	0	0	29	100.0%
BEECH	1	0	0	0	13	0	14	92.9%
SHADE	0	0	1	0	0	30	31	96.8%
Total	30	30	9	30	13	30	142	
Producer's accuracy	96.7%	96.7%	66.7%	96.7%	100.0%	100.0%		
Overall accuracy	95.8%							
Карра	0.948							

Tab. 4: Confusion matrix of segment-based RF model used in predictive species distribution mapping

As described in other publications, the spectral overlap of several species in the red edge and nearinfrared band (Fig. 5) requires adding features to the model based on visible spectra. In blue wavelengths, the emissivity of European silver fir shows lower values than all other species, whereas Scots pine showed substantially lower reflectance in the GREEN band. The implementation of a Blue Ratio (WASER et al. 2014) and VDVI (WANG et al. 2015; XUE & SU 2017) in the RF model underlines the assumption of a good species separation potential in visible wavelengths. These findings are congruent with results of other studies on tree species classification (FENG et al. 2015; GITELSON et al. 2003; IMMITZER et al. 2012; IMMITZER, VUOLO, et al. 2016; NG et al. 2016). Contributing features based on red, red edge and near-infrared wavelengths were the multispectral indices RENDVI and NDRE.



Fig. 4: Reflectance signatures of the five occurring tree species and shade class

In order to guarantee correct segment-based predictive mapping, the RF classifier was retrained with LSMS segments containing the centroids of the manually generated tree crown polygons. The evaluation of the resulting predictive species cover map (Fig. 4) showed remarkable congruency with ground reference from field observations.

Noteworthy overclassification in the predictive map was only observed in European larch, which shall be dispersed across the entire stand according to the predictive map (Fig. 5). In the course of the ground assessment individuals of European larch were only found in the northeastern quarter of the investigation area. Despite some other minor misclassifications like the confusion of a vigorous Norway spruce at the northern edge of the stand with the beech class, or intensely illuminated segments of coniferous tree crowns classified as European beech, the distinction between species showed reliable performance.



Fig. 5: Predictive species distribution map of tree species study site

4 Conclusion & Outlook

Our study results revealed high potential in the detection of infested or physiologically stressed trees, whereas an identification or assignment of the stress causes remains complicated and requires additional research. Concerning the species classification approach, our research delivered satisfactory results and achieved sufficiently accurate and reliable tree species classification and ensuing distribution mapping.

Feature selection showed that predictors based on red, red edge and near infrared reflectance were the most important ones for the detection of differences in tree vitality. This was similar for species classification which also relied on additional predictors based on reflectance in the visible region of the electromagnetic spectrum. From 18 tested band ratios and vegetation indices, only the PSRI, RENDVI, SAVI, NDRE, VDVI and Blue Ratio were used in the Random Forest classifier image processing method presented in this study. ADAMCZYK & OSBERGER (2015) mentioned differences in suitability of vegetation indices for threshold-based OBIA classification and the majority of selected indices in this study is also reported in other research publications (DASH et al. 2017; EINZMANN et al. 2017; WASER et al. 2014; XUE & SU 2017). Haralick texture indices were part of every best-performing Random Forest model which shows the importance and utility of textural information on crown level for classification of tree species and vitality differences.

Limitations in the spectral range covered by multispectral sensors might account for any of the occurring uncertainties. Earlier stages of physiological stress or smaller spectral differences could be detectable with better reliability using hyperspectral sensors that can observe physiological changes in wavelengths not covered by multispectral cameras. Hyperspectral cameras are however related to higher purchasing and operating costs. Additionally, the implementation of thermal imagery could enhance separability of different tree species and health conditions. Reflectance information from short-wave infrared (SWIR) and emissivity information from the thermal infrared (TIR) regions could emphasise characteristic differences for vitality analysis and species classification that are not perceivable with multispectral sensor equipment.

Altogether a significant reduction in expenses and the flexibility in application make UAV paired with multispectral sensor equipment a viable tool for the detailed remote sensing based monitoring of forested areas on small to medium scale. Hybridization of vertical take-off and landing and fixed-wing UAV are likely to expand the application range to much larger areas in the future (DEMPEWOLF et al. 2017). Furthermore, the technology could find use as a flexible support for established satellite remote sensing systems (e.g. Sentinel-2) and enhance or validate such research results with additional ground truth verification products in near-time.

5 References

- ADAMCZYK, J. & OSBERGER, A., 2015: Red-edge vegetation indices for detecting and assessing disturbances in Norway spruce dominated mountain forests. International Journal of Applied Earth Observation and Geoinformation, **37**, 90-99.
- BENGTSSON, H., 2017: matrixStats: Functions that Apply to Rows and Columns of Matrices (and to Vectors). R Package Version 0.52.2.

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

- DANDOIS, J. P. & ELLIS, E. C., 2013: High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. Remote Sensing of Environment, 136, 259-276.
- DASH, J. P., WATT, M. S., PEARSE, G. D., HEAPHY, M. & DUNGEY, H. S., 2017: Assessing very high resolution UAV imagery for monitoring forest health during a simulated disease outbreak. ISPRS Journal of Photogrammetry and Remote Sensing, **131**, 1-14.
- DEMPEWOLF, J., NAGOL, J., HEIN, S., THIEL, C. & ZIMMERMANN, R., 2017: Measurement of withinseason tree height growth in a mixed forest stand using UAV imagery. Forests, 8(7), 231.
- EINZMANN, K., IMMITZER, M., BÖCK, S., BAUER, O., SCHMITT, A. & ATZBERGER, C., 2017: Windthrow detection in european forests with very high-resolution optical data. Forests, **8**(1), 21.
- FACCOLI, M., 2009: Effect of Weather on Ips typographus (Coleoptera Curculionidae) Phenology, Voltinism, and Associated Spruce Mortality in the Southeastern Alps. Environmental Entomology, **38**(2), 307-316.
- FAHSE, L. & HEURICH, M., 2011: Simulation and analysis of outbreaks of bark beetle infestations and their management at the stand level. Ecological Modelling, **222**(11), 1833-1846.
- FASSNACHT, F. E., LATIFI, H., STEREŃCZAK, K., MODZELEWSKA, A., LEFSKY, M., WASER, L. T., STRAUB, C. & GHOSH, A., 2016: Review of studies on tree species classification from remotely sensed data. Remote Sensing of Environment, 186, 64-87.
- FENG, Q., LIU, J. & GONG, J., 2015: UAV Remote sensing for urban vegetation mapping using random forest and texture analysis. Remote Sensing, 7(1), 1074-1094.
- FRANKLIN, S. E. & AHMED, O. S., 2017: Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data. International Journal of Remote Sensing, 39(15-16), 5236-5245.
- FRANKLIN, S. E., AHMED, O. S. & WILLIAMS, G., 2017: Northern Conifer Forest Species Classification Using Multispectral Data Acquired from an Unmanned Aerial Vehicle. Photogrammetric Engineering & Remote Sensing, 83(7), 501-507.
- GITELSON, A. A., GRITZ, Y. & MERZLYAK, M. N., 2003: Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. Journal of Plant Physiology, **160**(3), 271-282.
- GOODBODY, T. R. H., COOPS, N. C., MARSHALL, P. L., TOMPALSKI, P. & CRAWFORD, P., 2017: Unmanned aerial systems for precision forest inventory purposes: A review and case study. The Forestry Chronicle, **93**(01), 71-81.
- HALMSCHLAGER, E. & KIRISITS, T., 2008: First report of the ash dieback pathogen Chalara fraxinea on Fraxinus excelsior in Austria. Plant Pathology, **57**(6), 1177.
- HIJMANS, R. J., 2017: raster: Geographic data analysis and modeling. R Package Version 2.6-7.
- IMMITZER, M. & ATZBERGER, C., 2014: Early Detection of Bark Beetle Infestation in Norway Spruce (Picea abies, L.) using WorldView-2 Data. Photogrammetrie Fernerkundung Geoinformation, (5), 351-367.
- 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., BÖCK, S., EINZMANN, K., VUOLO, F., PINNEL, N., WALLNER, A. & ATZBERGER, C., 2018: Fractional cover mapping of spruce and pine at 1 ha resolution combining very high and medium spatial resolution satellite imagery. Remote Sensing of Environment, 204, 690-703.
- IMMITZER, M., STEPPER, C., BÖCK, S., STRAUB, C. & ATZBERGER, C., 2016a: Use of WorldView-2 stereo imagery and National Forest Inventory data for wall-to-wall mapping of growing stock. Forest Ecology and Management, 359, 232-246.
- IMMITZER, M., VUOLO, F. & ATZBERGER, C., 2016b: First experience with Sentinel-2 data for crop and tree species classifications in central Europe. Remote Sensing, **8**(3), 166.
- JONES, H. G. & VAUGHAN, R. A., 2010: Remote sensing of vegetation: principles, techniques, and applications. Oxford University Press.
- KIRISITS, T. & FREINSCHLAG, C., 2012: Ash dieback caused by Hymenoscyphus pseudoalbidus in a seed plantation of Fraxinus excelsior in Austria. Journal of Agricultural Extension and Rural Development, **4**(9), 184-191.
- KRAUSE, S., HARTMANN, F., MUND, J.-P., 2016: UAV Workflow Optimization for the Acquisition of High Quality Photogrammetric Point Clouds in Forestry. GI_Forum 2016, Vol.1 Page: 72-84
- KUHN, M., WING, J., WESTON, S., WILLIAMS, A., KEEFER, C., ENGELHARDT, A., COOPER, T., MAYER, Z., KENKEL, B., TEAM, R. C., BENESTY, M., LESCARBEAU, R., ZIEM, A., SCRUCCA, L., TANG, Y., CANDAN, C. & HUNT, T., 2017: caret: Classification and Regression Training. R Package Version 6.0-78.
- LAUSCH, A., HEURICH, M., GORDALLA, D., DOBNER, H. J., GWILLYM-MARGIANTO, S. & SALBACH, C., 2013: Forecasting potential bark beetle outbreaks based on spruce forest vitality using hyperspectral remote-sensing techniques at different scales. Forest Ecology and Management, **308**, 76-89.
- LEHMANN, J. R. K., NIEBERDING, F., PRINZ, T. & KNOTH, C., 2015: Analysis of unmanned aerial system-based CIR images in forestry-a new perspective to monitor pest infestation levels. Forests, **6**(3), 594-612.
- LIAW, A. & WIENER, M., 2002: Classification and Regression by randomForest. R News, **2**(3), 18-22.
- LILLESAND, T., KIEFER, R. W. & CHIPMAN, J., 2014: Remote sensing and image interpretation. John Wiley & Sons.
- LINDNER, M., MAROSCHEK, M., NETHERER, S., KREMER, A., BARBATI, A., GARCIA-GONZALO, J., SEIDL, R., DELZON, S., CORONA, P., KOLSTRÖM, M., LEXER, M. J. & MARCHETTI, M., 2010: Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. Forest Ecology and Management, 259(4), 698-709.
- LISEIN, J., MICHEZ, A., CLAESSENS, H. & LEJEUNE, P., 2015: Discrimination of deciduous tree species from time series of unmanned aerial system imagery. PLoS ONE, **10**(11), e0141006.
- MASAITIS, G., MOZGERIS, G. & AUGUSTAITIS, A., 2013: Spectral reflectance properties of healthy and stressed coniferous trees. IForest, **6**, 30.
- 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.
- MICHEZ, A., PIÉGAY, H., LISEIN, J., CLAESSENS, H. & LEJEUNE, P., 2016: Classification of riparian forest species and health condition using multi-temporal and hyperspatial imagery from unmanned aerial system. Environmental Monitoring and Assessment, **188**(3), 1-19.
- MOHAN, M., SILVA, C. A., KLAUBERG, C., JAT, P., CATTS, G., CARDIL, A., HUDAK, A. T. & DIA, M., 2017: Individual tree detection from unmanned aerial vehicle (UAV) derived canopy height model in an open canopy mixed conifer forest. Forests, 8(9), 340.
- NÄSI, R., HONKAVAARA, E., LYYTIKÄINEN-SAARENMAA, P., BLOMQVIST, M., LITKEY, P., HAKALA, T., VILJANEN, N., KANTOLA, T., TANHUANPÄÄ, T. & HOLOPAINEN, M., 2015: Using UAVbased photogrammetry and hyperspectral imaging for mapping bark beetle damage at treelevel. Remote Sensing, 7(11), 15467-15493.
- NEVALAINEN, O., HONKAVAARA, E., TUOMINEN, S., VILJANEN, N., HAKALA, T., YU, X., HYYPPÄ, J., SAARI, H., PÖLÖNEN, I., IMAI, N. N. & TOMMASELLI, A. M. G., 2017: Individual tree detection and classification with UAV-Based photogrammetric point clouds and hyperspectral imaging. Remote Sensing, 9(3), 185.
- NG, W. T., MERONI, M., IMMITZER, M., BÖCK, S., LEONARDI, U., REMBOLD, F., GADAIN, H. & ATZBERGER, C., 2016: Mapping Prosopis spp. with Landsat 8 data in arid environments: Evaluating effectiveness of different methods and temporal imagery selection for Hargeisa, Somaliland. International Journal of Applied Earth Observation and Geoinformation, **53**, 76-89.
- NG, W. T., RIMA, P., EINZMANN, K., IMMITZER, M., ATZBERGER, C. & ECKERT, S., 2017: Assessing the potential of sentinel-2 and pléiades data for the detection of prosopis and vachellia spp. In Kenya. Remote Sensing, 9(1), 74.
- OLLINGER, S. V., 2011: Sources of variability in canopy reflectance and the convergent properties of plants. New Phytologist. Wiley/Blackwell (10.1111).
- PROVINCIAL GOVERNMENT OF LOWER AUSTRIA, 2015: Naturschutzkonzept Niederösterreich. St. Pölten.
- R CORE TEAM, 2017: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.
- RSTUDIO TEAM, 2016: RStudio: Integrated Development for R. Boston, MA: RStudio Inc.
- SCHULTZ, B., IMMITZER, M., FORMAGGIO, A. R., SANCHES, I. D. A., LUIZ, A. J. B. & ATZBERGER, C., 2015: Self-guided segmentation and classification of multi-temporal Landsat 8 images for crop type mapping in Southeastern Brazil. Remote Sensing, 7(11), 14482-14508.
- TOMAŠTÍK, J., MOKROŠ, M., SALOŠ, S., CHUDY, F. & TUNÁK, D., 2017: Accuracy of photogrammetric UAV-based point clouds under conditions of partially-open forest canopy. Forests, 8(5), 151.
- TORRESAN, C., BERTON, A., CAROTENUTO, F., DI, S. F., GIOLI, B., MATESE, A., MIGLIETTA, F., ZALDEI, A., WALLACE, L., TORRESAN, C., BERTON, A., CAROTENUTO, F., DI, S. F., GIOLI, B., MATESE, A., MIGLIETTA, F., VAGNOLI, C. & CAROTENUTO, F., 2016: Forestry applications of UAVs in Europe : a review Forestry applications of UAVs in Europe : a review. International Journal of Remote Sensing, 1-21.
- TOSCANI, P., IMMITZER, M. & ATZBERGER, C., 2013: Wavelet-based texture measures for objectbased classification of aerial images. Photogrammetrie - Fernerkundung - Geoinformation, 2013(2), 105-121.

- WANG, X., WANG, M., WANG, S. & WU, Y., 2015: Extraction of vegetation information from visible unmanned aerial vehicle images. Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering, 31(5), 152-159.
- WASER, L. T., KÜCHLER, M., JÜTTE, K. & STAMPFER, T., 2014: Evaluating the potential of worldview-2 data to classify tree species and different levels of ash mortality. Remote Sensing, **6**(5), 4515-4545.
- XUE, J. & SU, B., 2017: Significant remote sensing vegetation indices: A review of developments and applications. Journal of Sensors. Hindawi.

ZAMG, 2018: Klimanormalwerte Österreich 1981–2010. Wien.

Automatisierte Detektion von Waldbrandflächen mit Sentinel-2 Satellitendaten

MIRA WEIRATHER¹

Zusammenfassung: Die meisten Katastrophen des letzten Jahrzehnts waren an klimatisch bedingte Naturgefahren wie Überschwemmungen, Stürme, Hitzewellen oder Trockenperioden geknüpft. Prognosen zeigen, dass durch die globale Erwärmung die Gefahr einer erhöhten Anzahl von Waldbränden steigt. Wälder leisten einen großen ökologischen, sozialen und ökonomischen Beitrag für die Umwelt. Während für manche Wälder Feuer sogar zum Erhalt der natürlichen Artenzusammensetzung beitragen, sind sie für andere Waldökosysteme gefährlich und können auch zur Änderung des Landschaftsbildes führen. Die neuerdings auftretenden Mega-Waldbrände, unkontrollierbare Feuerstürme, sind jedoch für beide Arten gleichermaßen gefährlich. Um die exakte Ausdehnung eines Waldbrandes zu ermitteln, hat sich die Fernerkundung als geeignetes Mittel etabliert. Besonders neue Technologien wie die frei verfügbaren Sentinel-2 Satelliten eignen sich gut für die Anwendung bei der Waldbranderfassung. Aus diesen Daten wird in dieser Arbeit der Normalized Burn Ration (NBR) berechnet, ein multispektraler Index, der sich besonders für die Unterscheidung zwischen verbrannter und nicht verbrannter Fläche eignet. Anhand dreier Studiengebiete, die in unterschiedlichen biogeographischen Regionen der EU liegen, wird die vollautomatisierte Prozesskette entwickelt und getestet. Die Genauigkeitsanalyse zeigt, dass mit der Prozesskette gute bis sehr gute Ergebnisse erzielt werden konnten: Die Kappa Koeffizienten reichen von 0,72 bis 0,89.

1 Einleitung

Statistiken globaler Rückversicherungen wie die Münchner Rück belegen, dass die meisten Katastrophen des letzten Jahrzehnts an klimatisch bedingte Naturgefahren wie Überschwemmungen, Stürme, Hitzewellen und Trockenperioden geknüpft waren (MUNICHRE 2018). Aufgrund des Klimawandels wird die Wahrscheinlichkeit solcher Extremereignisse künftig noch weiter ansteigen (IPCC 2013). Hitzewellen und Trockenperioden schaffen ideale Bedingungen für den Ausbruch und die schnelle Ausbreitung von Waldbränden. Wie die großen Waldbrände des Jahres 2017 in Kalifornien und im Mittelmeerraum aber auch 2018 in Deutschland und in Schweden zeigen, müssen wir damit rechnen, dass Waldbrände eine zunehmend größere Gefahr für Umwelt und Gesellschaft darstellen. Wälder machen mehr als 40 % der gesamten Landoberfläche aus (BARREDO et al. 2015) und bieten viele Ökosystemdienstleistungen in allen Bereichen, von der Bereitstellung von Ressourcen wie Holz und Fasern über Klimaregulierung bis hin zu seiner erholsamen Wirkung. Während manche Ökosysteme für ihre natürliche Artenzusammensetzung und -erhaltung auf Waldbrände angewiesen sind (HIRSCHBERGER 2012), können Waldbrände für andere Ökosysteme schwerwiegende Folgen wie einen Wandel des Landschaftsbildes haben (JUSTICE & KORO-NTZI 2001). Von den sich in den letzten Jahren verstärkt ausbreitenden Mega-Waldbränden, die wegen ihrer starken Intensität kaum oder gar nicht unter Kontrolle gebracht werden können, sind

¹ Technische Universität München, Fachgebiet für Waldinventur und nachhaltige Nutzung, Hans-Carlvon-Carlowitz-Platz 2, D-85354 Freising, E-Mail: weiratherm@yahoo.de

jedoch feuerempfindliche und feuerabhängige Ökosysteme gleichermaßen negativ betroffen (HIRSCHBERGER 2012). Um umfangreiche Analysen über die Konsequenzen der Waldbrände machen zu können, wie beispielsweise über die Menge des dabei freigesetzten Kohlenstoffs, ist in erster Linie die genaue Ermittlung der verbrannten Fläche erforderlich. Dafür hat sich die Fernerkundung als geeignetes Mittel herausgestellt. Mehrere Initiativen widmen sich der Erfassung der globalen verbrannten Fläche, viele auf Grundlage von MODIS Daten mit einer Auflösung von 500 m, wie beispielsweise das Standardprodukt von MODIS für die globale verbrannte Fläche MCD45A1 (MOUILLOT et al. 2014). BIRD, ein vom DLR entwickelter deutscher Kleinsatellit, lieferte von 2001 bis 2006 Daten für ein Frühwarnsystem von Waldbränden (ALPERS 2018). Mit dem Start der Sentinel-2 Satelliten im Jahr 2015 stehen nun Satellitenbilddaten in einer sehr viel besseren räumlichen Auflösung (10, 20 und 60 m) frei zur Verfügung. Damit können auch kleinere Brände erfasst werden, durch die hohe zeitliche Auflösung von 2,5 bis 5 Tagen und die damit erhöhte Wahrscheinlichkeit eines ausreichend nahe am Feuer gelegenen, wolkenfreien Satellitenbildes, eignen sich die Sentinel-2 Satelliten besonders zur Erfassung von Waldbränden (ESA 2018b). Da das Multispectral Instrument (MSI), der Sensor der Sentinel-2 Satelliten, auch Bilder im Nahen (NIR) und Mittleren Infrarot (SWIR) aufnimmt, lässt sich der für Landsat TM entwickelte Normalized Burn Ratio (NBR) gut auf Sentinel-2 übertragen.

$$NBR = \frac{NIR - SWIR}{NIR + SWIR}$$

Der NBR ist ein spektraler Index, der sich besonders gut zur Ermittlung verbrannter Flächen eignet, da er neben dem Nahen Infrarot auch aus dem Mittleren Infrarot berechnet wird. Auf diese Art berechnete Indizes zeigen eine höhere Genauigkeit bei der Ermittlung der Brandflächen als Indizes, die das Mittlere Infrarot nicht verwenden (ROGAN & FRANKLIN 2001).

Ziel dieser Arbeit ist es, auf der Basis von den frei verfügbaren Sentinel-2 Daten eine Prozesskette zur Ermittlung der exakten Brandfläche nach einem Feuerereignis zu entwickeln. Im Fokus steht dabei die Vollautomatisierung der Prozesskette und die ausschließliche Verwendung von Open Source Software und von Open Data. Die Arbeit ist Teil von I-REACT, einem europäischen Forschungsprojekt, das durch das EU-Programm Horizont 2020 finanziert wird. Beim Verfassen dieser Arbeit ist I-REACT noch in der Entwicklung.

2 Methodik

Für die Ermittlung der verbrannten Fläche wurde in dieser Arbeit der NBR verwendet. Als Eingangsdaten wurden die Bänder 8A (NIR) und 12 (SWIR) mit einer räumlichen Auflösung von 20 m der Sentinel-2 Satelliten hergenommen. Die Auswahl der Bänder erfolgte in Anlehnung an die entsprechenden Landsat TM/ETM+ Bänder, für die der NBR ursprünglich entwickelt wurde. Für die Genauigkeitsanalyse wurden die Karten des Copernicus Emergency Management Service (EMS) verwendet, aufgrund derer auch die Auswahl der Studiengebiete erfolgte. Zusätzlich wurden die Studiengebiete so ausgewählt, dass sie in drei verschiedenen biogeografischen Regionen liegen. Biogeografische Regionen sind Gebiete mit ähnlichen Eigenschaften von Lebensräumen und Arten (EEA 2002). Außerdem sollten alle drei Brände im Jahr 2017 ausgebrochen sein. Dadurch fiel die Wahl auf einen Waldbrand in der pannonischen Region (Ungarn), einen in der mediterranen Region (Frankreich) und einen in der alpinen Region (Spanien). Um zu analysieren, welche Landnutzungstypen durch die Feuerereignisse verbrannt sind, wurde die Dominant Tree Type Karte 2015 des JRC für das Studiengebiet in der pannonischen Region verwendet, für die anderen beiden Regionen das Landbedeckungsprodukt CORINE Land Cover (CLC) 2012 von Copernicus.

Für die Festlegung des Schwellenwertes des NBR, ab dem ein Pixel als verbrannt gilt, wurde ein empirischer Ansatz verfolgt: Für jedes der drei Studiengebiete wurden verschiedene Schwellenwerte getestet und derjenige verwendet, der auf einer visuellen Beurteilung basierend die besten Ergebnisse zeigte.



Abb. 1: Arbeitsschritte der entwickelten Prozesskette

Um die Vollautomatisierung der Prozesskette zu garantieren, wurde sie in der Amazon Elastic Compute Cloud (EC2) der Amazon Web Services (AWS) implementiert. Abb. 1 zeigt den Workflow der vollautomatisierten Prozesskette. Brandflächenkarten werden nur dann erzeugt, wenn sie von einem autorisierten Benutzer angefordert werden. Diese Forderung löst die Aktivierung der Prozesskette aus. Der Auslöser enthält eine Punktkoordinate des Brandes und das Enddatum des Feuers. Die Punktkoordinate wird zu einem Interessensgebiet mit einem Radius von 8 km erweitert, da der genaue Umfang der Brandfläche noch nicht bekannt ist. Das Interessensgebiet wird mit der Ausdehnung der verfügbaren Satellitenbilder überlagert, um zu überprüfen, ob eine oder mehrere Kacheln benötigt werden. Anschließend werden die Ergebnisse nach ihrer Wolkenbedeckung geordnet. Gibt es ein Bild ohne Wolken, wird es für die weitere Verarbeitung verwendet. Ansonsten wird das Interessensgebiet nacheinander mit der von Sentinel mitgelieferten Wolkenmaske der Bilder abgeglichen, um ein Satellitenbild zu bekommen, das keine Wolken in der AOI aufweist. Im nächsten Schritt werden die Top-Of-Atmosphere (TOA)-Reflektanzen der benötigten Bänder mit einer Dark Object Subtraction (DOS) zu Bottom-Of-Atmosphere (BOA)-Reflektanzen kalibriert. Ist aufgrund des Interessensgebietes mehr als eine Szene für die vollständige Extraktion der Brandfläche nötig, werden die Bilder zusammengefügt. Die Extraktion der Brandfläche erfolgt anschließend in zwei Stufen. In der ersten Stufe wird die Prozesskette mit Bändern durchgeführt, die auf 250 m Auflösung herunterskaliert wurden, um einen groben Umriss der betroffenen Fläche zu erhalten. Nach Berechnung des NBR und Einsatz des Schwellenwertes wird die verbrannte Fläche des gesamten Bildes extrahiert. Aus dem resultierenden Binärbild werden zusammenhängende, eigenständige Pixelgruppen zu einzelnen Kategorien gruppiert. Nach Auswahl der gesuchten Fläche durch Verschneiden mit dem anfänglichen Interessensgebiet wird das Rasterbild vektorisiert. Um auch kleinere Brandflächen außerhalb der eigentlichen Brandfläche abzudecken, wird ein Polygonpuffer eingesetzt. In der zweiten Stufe werden die originalen Bänder mit 20 m Auflösung entlang der Außenlinie des Vektors aus Stufe eins ausgeschnitten. Daraus wird wiederum der NBR berechnet, mithilfe des Schwellenwertes die verbrannte Fläche extrahiert und anschließend kleine Pixelgruppen eliminiert. Das Ergebnis wird vektorisiert, für eine ansprechendere Außenlinie geglättet und als Vektoratei an das System des Projektes zurückgegeben.

Die Prozesskette wurde mit der Programmiersprache Python implementiert. Außerdem werden Open Source Werkzeuge von QGIS, GDAL, SAGA GIS und GRASS GIS benutzt.

3 Ergebnisse

Für alle drei Studiengebiete verlief die Prozesskette erfolgreich. Abb. 2 zeigt die Falschfarbenbilder mit den ermittelten verbrannten Flächen für alle drei Studiengebiete. Bereits im optischen Vergleich wird deutlich, dass die Prozesskette akkurate Ergebnisse geliefert hat. In der mediterranen Region wurden nicht verbrannte Nordhänge, die im Falschfarbenbild als Schatten dargestellt wer-



Abb. 1: Falschfarbenbilder (Bänder 8, 4, 3) nach dem Feuerereignis mit den ermittelten verbrannten Flächen (gelbe Außenlinie) der mediterranen, pannonischen und alpinen Region

den, auch tatsächlich nicht als verbrannte Fläche extrahiert. In der pannonischen Region konnten einige Inselpolygone innerhalb bzw. außerhalb der Hauptbrandfläche als nicht verbrannt bzw. verbrannt erkannt werden. In der alpinen Region wurden einige Flächen im nördlichen Teil des Brandes jedoch fälschlicherweise nicht als verbrannt erkannt. Zwei Straßen, die durch die verbrannte Fläche laufen, machen den Anschein, dass sie den Brand zum Teil durch ihre Wirkung als Barrieren aufgehalten haben. Zu sehen ist das entlang der einen Straße im nordöstlichen Teil des Brandes und in der Mitte des Gebietes, westlich der anderen Straße.

Tab. 1 zeigt die Ergebnisse der Genauigkeitsanalyse für alle drei Studiengebiete. Es wurden jeweils die Konfusionmatrizen anhand 1000 zufällig ausgewählter Verifikationsflächen innerhalb der Interessensgebiete berechnet. Daraus wurden die Gesamtgenauigkeit (GG), die Herstellergenauigkeit (HG), die Nutzergenauigkeit (NG) und Cohen's Kappa berechnet.

		·····	jj	
		Mediterrane Region	Pannonische Re- gion	Alpine Region
HG	Total	0,85	0,94	0,89
	Verbrannt	0,70	0,89	0,83
	Nicht verbrannt	1,00	0,98	0,95
NG	Total	0,94	0,95	0,85
	Verbrannt	0,99	0,96	0,97
	Nicht verbrannt	0,88	0,94	0,73
GG		0,91	0,95	0,87
Kappa		0,76	0,89	0,72

Tab. 1: Genauigkeitsparameter der Brandflächenkarten. HG = Herstellergenauigkeit; NG = Nutzergenauigkeit, GG = Gesamtgenauigkeit

Für die mediterrane Region nimmt die Gesamtgenauigkeit mit 0,91 einen hohen Wert ein, der Kappa Koeffizient beträgt allerdings nur 0,76. Das lässt sich darauf zurückzuführen, dass die Schatten der Nordhänge im westlichen Teil der verbrannten Fläche bei den als richtig angenommenen Referenzdaten fälschlicherweise als verbrannt dargestellt werden. Dies wird durch die geringe Herstellergenauigkeit der verbrannten Verifikationsflächen bestätigt: Viele der als tatsächlich verbrannt angenommenen Flächen wurden im Ergebnis der Prozesskette nicht als verbrannt extrahiert. Da die Referenzdaten an dieser Stelle falsch sind, kann angenommen werden, dass die Genauigkeit des Ergebnisses höher ist, als die Genauigkeitsparameter es zeigen. Für die pannonische Region konnten die höchsten Genauigkeiten erreicht werden. Sowohl die Hersteller- als auch die Nutzergenauigkeit liegen über 0,94. Der Kappa Koeffizient lieg t bei 0,89 und bestätigt damit die hohe Übereinstimmung zwischen dem Ergebnis der Prozesskette und den Referenzdaten. Mit einem Kappa Koeffizient von 0,72 zeigt die Brandfläche der alpinen Region die geringste Genauigkeit.



Abb. 3: Landbedeckung der verbrannten Flächen der mediterranen, pannonischen (CORINE LandCover 2012) und alpinen Region (Dominant Tree Type 2015, COPERNICUS)

Abb. 3 zeigt die Landbedeckungstypen innerhalb der verbrannten Flächen für die drei Studiengebiete. In der mediterranen Region macht Hartlaubvegetation den größten Anteil aus, gefolgt von Wald-Strauch-Übergansstadien, Flächen mit spärlicher Vegetation und natürlichem Grünland. In der pannonischen Region dominieren Sümpfe, der einzige andere Landbedeckungstyp ist natürliches Grünland. In der alpinen Region machen Laubwälder den größten Anteil aus, natürliches Grünland und Nadelwälder haben einen gleich großen Anteil.

4 Diskussion und Schlussfolgerung

Das entwickelte Verfahren konnte erste verlässliche Ergebnisse liefern. Hervorzuheben ist, dass es ausschließlich auf die Nutzung von frei verfügbaren Sentinel-2 Satellitenbilddaten, frei verfügbaren Programmierbibliotheken und eines oft verwendeten und somit bekannten und gut einschätzbaren spektralen Index baut und dabei vollautomatisiert abläuft. In der Cloud implementiert und durch einen autorisierten Benutzer ausgelöst, wird so schnelles und von menschlicher Interaktion unabhängiges Prozessieren ermöglicht. Darüber hinaus lässt die Auswertungskette Skalierbarkeit zu, eine wichtige Eigenschaft, wenn der Service nicht mehr durch einen Benutzer, sondern automatisiert durch z.B. Katastrophenwarnsysteme ausgelöst wird. Der Vergleich mit den Copernicus EMS Referenzkarten bestätigte, dass das Verfahren nahezu identische Ergebnisse liefert, diese aber sehr viel schneller erzeugt. Die zweistufige Herangehensweise, beginnend mit einer Herunterskalierung auf 250 m und weiterführender Analyse bei 20 m Auflösung zeigte gute Ergebnisse für die Studiengebiete. Für kleinere Flächen muss die Auswertungskette noch getestet und das Herunterskalieren dafür angepasst werden, da solche Brände bei einer zu geringen Auflösung nicht detektiert werden können, wie eine Studie von CORTI 2017 mit MODIS-Daten zeigte (CORTI et al. 2017). Da das Verfahren jedoch im Hinblick auf Monitoring im kontinentalen bis globalen Maßstab entwickelt wurde, sind die kleineren Flächen eher als Sonderfall einzustufen.

Für I-REACT ist es geplant, den Output der Auswertungskette für die Echtzeit-Kalibrierung von Waldbrandsimulationen während eines Ereignisses zu nutzen. Diese Simulationen sind wichtig, um die Feuerwehr beim Löschen des Feuers bestmöglich zu unterstützen. Des Weiteren können die erstellten Karten zur Validierung der Ergebnisse der Simulationsmodelle hergenommen werden. Für Forstbetriebe und Behörde n kann das Verfahren zu Zwecken der Prävention und Ursachenforschung genutzt werden. Eine Verbesserung des Service könnte die Verwendung von Sentinel-3 Daten sein, die mit einer räumlichen Auflösung von 300 m, aber täglicher Repetitionsrate (ESA 2018a) die Verfügbarkeit von wolkenfreien Szenen erhöhen würden. Sentinel-2 Daten sind in so einem Verfahren dann zur räumlich expliziten Lokalisierung im Sinne der oben beschriebenen zweiten Hierarchiestufe zu sehen.

5 Literaturverzeichnis

- ALPERS, M., 2018: DLR Raumfahrtmanagement. BIRD Erster Satellit zur Entdeckung von Waldbränden. https://www.dlr.de/rd/desktopdefault.aspx/tabid2440/3586_read-5329/, letzter Zugriff 25.07.2018.
- BARREDO, J., BASTRUP-BIRK, A., TELLER, A., ONAINDIA, M., FÉRNANDEZ DE MANUEL, B., MA-DARIAGA, I. & MIMO, S., 2015: Mapping and asessment of forest ecosystems and their services - Applications and guidance for decision making in the framework of MAES. EUR 27751 EN.
- CORTI, N., MENGISTIE, K. & SCHNEIDER, T., 2017: Brandflächen in Ecuadors Tropenwald per Satellit erfassen. AFZ-Der Wald, 19, 25-29.
- EEA, 2002: Europe's biodiversity biogeographical regions and seas. https://www.eea.eu-ropa.eu/publications/report_2002_0524_154909, letzter Zugriff 11.03.2018.

- ESA, 2018a: Revisit and Coverage. https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/revisit-coverage, letzter Zugriff 28.01.2018.
- ESA, 2018b: Sentinel-2. https://earth.esa.int/web/sentinel/missions/sentinel-2, letzter Zugriff 23.05.2018.
- HIRSCHBERGER, P., 2012: Wälder in Flammen Ursachen und Folgen der weltweiten Waldbrände. WWF Deutschland (Hrsg.), WWF Deutschland, Berlin.
- IPCC, 2013: Climate Change: The Physical Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschun, & P. Midgley, (Hrsg.), Cambridge, United Kingdom an New York, NY, USA: Cambridge University Press.
- JUSTICE, C. & KORONTZI, S., 2001: A Review of the Status of Satellite Fire Monitoring and the Requirements for Global Environmental Change Research. In F. Ahern, Global and Regional Vegetation Fire Monitoring from Space: Planning a Coordinated International Effort, 1-19.
- MOUILLOT, F., SCHULTZ, M., YUE, C., CADULE, P., TANSEY, K., CIAIS, P. & CHUVIECO, E., 2014: Ten years of global burned area products from spaceborne remote sensing - A review: Analysis of user needs and recommendations for future developments. International Journal of Applied Earth Observation and Geoinformation, **26**, 64-79.
- MUNICHRE, 2018: MunichRe Press Release. https://www.munichre.com/en/media-relations/publications/press-releases/2018/2018-0104-press-release/index.html, letzter Zugriff 28.01.2018
- ROGAN, J. & FRANKLIN, J., 2001: Mapping wildfire burn severity in Southern California Forests and shurblands using Enhanced Thematic Mapper imagery. Geocarta International, 22, 89-99.

Wildtier-Monitoring mit UAVs – Künstliche Intelligenz zur automatisierten Detektion von Infrarot-Signaturen

ADRIAN F. MEYER¹, NATALIE LACK & DENIS JORDAN

Zusammenfassung: Die Detektion von Wildtieren ist ein zentrales Monitoring-Instrument in Ökologie, Jagd, Forst- und Landwirtschaft. Bisherige Verfahren sind aufwändig, basieren häufig lediglich auf indirekten Nachweisen und bieten so oft nur eine grobe Schätzung der Bestände. Die in dieser Arbeit durchgeführte fernerkundliche Auswertung von UAV-Befliegungen über dem Südschwarzwald und der Nordwestschweiz zeigte, dass sich insbesondere Wärmebilddaten für eine Automatisierung der Wildtier-Detektion eignen. Hierzu wurde ein modernes Verfahren der künstlichen Intelligenz (Faster R-CNN) aufgebaut, das in der Lage ist, durch Training Eigenschaftsmerkmale aus markierten Tiersignaturen zu extrahieren. Für einige Tierarten (Hirsche, Ziegenartige, Wisente, Weidevieh) konnten in der anschließenden Anwendung (Inferencing) extrem robuste Detektionsergebnisse erreicht werden. Die effiziente Implementierung des Prototyps erlaubt eine Echtzeitanalyse von Live-Video-Feeds unter Feldbedingungen. Mit einer Detektionsrate von 92,8% pro Tier, bzw. 88,6% in der Klassifikation nach Tierart, konnte so gezeigt werden, dass die neue Technologie ein enormes Innovationspotential für die Zukunft des Wildtier-Monitorings aufweist.

1 Einleitung

Für Anwendungsfelder wie Populationsmanagement, Rehkitzrettung und Wildschadensprävention in Ökologie, Jagd, Forst- und Landwirtschaft ist es von entscheidender Bedeutung, eine möglichst präzise Erfassung der Wildtierbestände vornehmen zu können. Beim konventionellen Monitoring kommen derzeit zumeist Methoden zum Einsatz, die jeweils entscheidende Nachteile mit sich bringen (SILVEIRA et al. 2003): Zählkampagnen mit visueller Bestätigung (Scheinwerfersuche auf Waldwegen) sind enorm personalintensiv; Kamerafallen-Analysen decken nur einen kleinen Ausschnitt der Landschaft ab; Jagdquoten und Wildunfallstatistiken sind mit einem starken Bias verbunden; Peilsender sind zwar sehr genau, aber auch invasiv und in der Umsetzung aufwändig. Das Institut Geomatik (FHNW) kooperiert seit Januar 2018 mit der Stiftung Wildtiere des Aargauischen Jagdverbandes, um ein Verfahren zur Wildtierdetektion mittels UAVs (Unmanned Aerial Vehicles) zu entwickeln. Es wird untersucht, in wie weit automatisierte Fernerkundungs-methoden Vorteile gegenüber dem konventionellen Monitoring bieten, indem zeitliche oder personelle Ressourcen eingespart und somit Erhebungen präziser und vollständiger durchgeführt werden können (GONZALEZ et al. 2016). Zentrale Fragen, welche dieses Studie beantworten soll, stellen sich dabei hinsichtlich der Wahl von Sensorik und Trägersystemen, der generellen Sichtbarkeit von Tiersignaturen auf Infrarotluftbildern (z.B. die Robustheit gegenüber Verschattungen im Mischwald), sowie der Struktur eines möglichst performanten Algorithmus zur automatisierten Detektion und Klassifikation der Wildtierindividuen. Ein Ergebnis dieser Analyse ist ein Prototyp, der eine automatisierte Tierdetektion auf Luftbilddaten ermöglichen soll.

¹ Fachhochschule Nordwestschweiz, Institut Geomatik, Hofackerstr. 30, CH – 4132 Muttenz, E-Mail: adrian.meyer@fhnw.ch

2 Methode

2.1 Datenerhebung

Im Frühjahr 2018 wurden 27 Befliegungen über sieben naturnahen Wildgehegen mit heimischen Tierarten in der Nordwestschweiz und dem Südschwarzwald durchgeführt. Für jedes Gehege wurden mit dem Multikopter *senseFly Albris* oder dem Fixed-Wing-UAV *senseFly eBee* ca. 500 RGB-Bilder, 500 NIR-Multispektralbilder und 5000 TIR-Wärmebilder (radiometrische Thermogramme) generiert, um einen Technologievergleich zu ermöglichen (siehe Abb. 1). Die Aufnahmezeit (Februar/März) wurde so gewählt, dass der Wärmekontrast zwischen Tierkörper und zumeist bewaldeter Umgebung möglichst hoch ausfallen würde. Gleichzeitig sollte durch die laubfreie Vegetation möglichst wenig Verschattung entstehen.



Abb. 1: Links: Eingesetzte senseFly-Fluggeräte "eBee" (oben) und "Albris" (unten). Rechts: Typische Flugbahn mit der eBee (blau) über einem Tierpark (grün) mit den Auslösepositionen für Luftbilder (weiss). (Visualisierungen: GILLINS et al. 2018; GOOGLE 2018; SENSEFLY 2018)

Mit dem Fixed-Wing können große Gebiete unkompliziert mit austauschbaren Sensoren (RGB, NIR, TIR) erfasst werden, u.a. mit einer hochauflösenden Thermalkamera (*ThermoMap*, 640×512Px; max. 22ha bei 15cm/Px GSD und 100m AGL). Mit dem Multikopter kann zwar aufgrund dessen Schwebefähigkeit wesentlich flexibler und tiefer geflogen werden, allerdings weist die fest verbaute Thermalkamera eine deutlich niedrigere Auflösung auf (80×60Px). Das laute Rotorengeräusch bei einer niedrigen Flugbahn stellt zudem im Vergleich zum Fixed-Wing einen wesentlich stärkeren Eingriff in das Tierverhalten dar.

2.2 Vorprozessierung

Die sehr hoch aufgelösten RGB- und NIR-Aufnahmen (~3cm/Px GSD) eigenen sich gut für das Erstellen von Übersichtskarten in Form von Orthophotomosaiken, weisen jedoch häufig keinen ausreichenden Kontrast für die visuelle Erkennung von Tiersignaturen unter laubfreier Vegetation auf. Dies konnte im weiteren Verlauf der Studie auch durch terrestrische hyperspektrale Referenzmessungen (λ : 350-1000nm) an Waldboden, Vegetation und Tierkörpern verifiziert werden. Die Thermogramme hingegen zeigen kontrastreiche Signaturen einzelner Wildtiere (Abb. 2). Die Bilder eignen sich aber gleichzeitig kaum für den photogrammetrischen Bündelblockausgleich, da sich die Tiere in der Regel zwischen zwei Aufnahmen zu stark bewegen. In den relevanten Bildbereichen entsteht dadurch keine ausreichende Überlagerungstreue, sodass prozessierte TIR-Orthophotomosaike zusammenhängender Habitate häufig keine sichtbaren Signaturen enthalten. Für 39. Wissenschaftlich-Technische Jahrestagung der DGPF und PFGK19 Tagung in München – Publikationen der DGPF, Band 28, 2019

die automatisierte Analyse wurden die Thermogramme deshalb entweder direkt als nicht-orientierte Rohdaten weiterverarbeitet, oder einzeln per DSM-Projektion orthorektifiziert.

3 Analyse

3.1 Gestalt der thermalen Tiersignaturen

Sichtbare Änderungen im Erscheinungsbild der Signaturen wurden zunächst systematisch durch die Variation von Referenzparametern untersucht. So unterstützt eine flachere Aufnahmeperspektive die Tier-Identifikation durch einen menschlichen Betrachter (Abb. 2 links): Merkmale wie Kopf-Torso-Ratio oder Extremitäten treten deutlicher hervor. Die Abgrenzung der Individuen voneinander wird hingegen eher durch eine steilere Perspektive unterstützt.

In möglichst laubfreiem Mischwald reduziert dichtes Astwerk zwar teilweise den Signaturkontrast durch Konvektionswärmeverteilung und Abschirmung; Form, Umfang und die grundsätzliche Sichtbarkeit der Signaturen bleiben jedoch in weiten Teilen erhalten (Abb. 2 rechts).



Abb. 2: Thermogramme mit den Signaturen eines Damhirsch-Rudels (sechs Tiere, Blau 4°C, Rot 10°C). Links: Signaturen aus sechs verschiedenen Aufnahme-Winkeln im Vergleich. Mitte/Rechts: Signaturen neben und unterhalb einer laubfreien Esche im Vergleich.

3.2 Strategien zur automatisierten Signaturdetektion

Mehrere Strategien zur automatisierten Detektion von Signaturen wurden iterativ implementiert und hinsichtlich ihrer Klassifikationsgenauigkeit und Anwendbarkeit überprüft. Der klassische fernerkundliche Ansatz, mittels objektbasierter Bildanalyse Thermogramme beispielsweise in *Erdas Imagine Objective* zu klassifizieren, wurde dabei verworfen. Mit dieser Methode konnte aufgrund der Vielgestaltigkeit der Signaturen kein merkmalsbeschreibender Satz von Variablen gefunden werden, der eine Detektionspräzision von über 41% erreichen würde. Gefaltete neuronale Netzwerke (Convolutional Neural Networks, CNN) hingegen zeigten in den letzten Jahren eine außerordentliche Robustheit bei der Bildklassifikation mittels automatischer Merkmalsextraktion (SZEGEDY et al. 2016). In den Abschnitten 3.3 und 3.4 werden zwei CNN-Ansätze beschrieben, die auf unterschiedliche Art und Weise eine präzise Tierdetektion erreichen.

3.3 Rastersegment-Klassifikation mit dichotomen CNN

Ein dichotomes ("Zwei-Wege-Entscheidung") CNN mit einer Tiefe von 7 Neuronen-Layern (Abb. 3 Mitte) wurde mit *Keras* und *Tensorflow* unter *Python 3.6* aufgebaut. Es klassifiziert Rastersegmente orthorektifizierter Thermogramme per Inferencing in die Klassen "Tier" und "Nicht-Tier". Als Input-Layer wird eine 64×64Px-Matrix verwendet, welche bei maximal möglicher GSD

geoprozessierten 5×5m-Segmenten entspricht (Abb. 3 links). Nach ca. 3h Training auf Desktophardware kann für eine spezifische Befliegung eine hohe Klassifikationspräzision von ca. 90% erreicht werden (Abb. 3 rechts). Das Vorprozessieren der Thermaldaten (3D-Projektion auf DSM, Orthophoto-Generierung, Geoprozessierung) ist jedoch sehr zeit- und rechenintensiv und somit unter Feldbedingungen als nicht praktikabel automatisierbar einzustufen. Gerade bei zeitkritischen Anwendungen wie der Rehkitzrettung müssen Klassifikationsergebnisse im besten Fall bereits während des Flugs vorliegen. Das Inferencing auf Live-Rohdaten wäre diesen Limitationen nicht unterworfen. Durch die Rohdaten-Auflösung von 640×512Px bietet dieser Ansatz dem UAV-Operator bei der praxisnahen Anwendung aufgrund der 64×64Px-Inputauflösung jedoch nur ein grobes 10×8-Detektionsraster.



Abb. 3: Links: Ca. 10'000 5×5m Footprints als Inputkacheln, generiert aus 45 orthorektifizierten Thermo-grammen. Mitte: Schema des dichotomen neuronalen Netzwerks, Neuronenlayer in der violetten Markierung. Rechts: Klassifikation - 71 Kacheln Grün: "Tier"; Rest Rot: "Nicht-Tier".

3.4 Objekterkennung mittels R-CNN

Für die Live-Rohdaten-Interpretation sind *Faster Region-based Convolutional Neural Networks* (*Faster R-CNN*) besser geeignet. Modelle dieser Klasse können Objekte auf höher aufgelösten Gesamtbildern klassifizieren, indem *Regions of Interest (RoI)* durch iteratives *Region Proposal* lokalisiert werden. Auch können gleichzeitig verschiedene Klassen trainiert und erkannt werden. Eingesetzt wird ein *Inception-v2*-Netzwerk (Vgl. Abb. 4), das mit einer Tiefe von 42 Neuronen-Layern den Aufbau der Pyramidenzellen im visuellen Cortex von Wirbeltieren nachahmt. Durch das Vortraining mit 100'000 Alltagsbildern (sog. *COCO-Dataset*) können die Kantengewichte zwischen den Neuronen-Layern im spezifischen Training schneller und effizienter an neue Ziele für das Festlegen der *Bounding Boxes* angepasst werden. Das Modell gilt auch bei teilweise limitierten Hardware-Anforderungen noch als schnell und präzise (SZEGEDY et al. 2016).

Die Implementierung erfolgte mittels der *Tensorflow Object Detection Library* mit Unterstützung durch das *Nvidia CUDA/cuDNN* Deep-Learning-Framework zur Parallelisierung der GPU-Shaderkerne. Für das Training wurde ein Testdatensatz von ca. 600 Thermalbildern mit ca. 8'000 Tiersignaturen durch Zeichnen von ca. 1'800 Bounding Boxes manuell markiert. Nach ca. 12 Stunden Training (ca. 100'000 *Steps*) wurde der ca. 50 Mbyte große *Frozen Inference Graph* exportiert. Ein performanter, auf Python-Scripts basierender Prototyp wendet dieses Wissensschema per Inferencing auf neue Thermaldaten an.

39. Wissenschaftlich-Technische Jahrestagung der DGPF und PFGK19 Tagung in München – Publikationen der DGPF, Band 28, 2019



Abb. 4: Schematische Struktur des aufgebauten R-CNN (Subschema "Inception v2" aus ALEMI, 2016)

4 Ergebnis

Im Vergleich erwies sich die Objekterkennung mittels R-CNN aufgrund der Fähigkeiten, Rohdaten verwenden und mehrere Klassen gleichzeitig trainieren zu können, als der überlegenere Ansatz. In der prototypischen Umsetzung wurde deshalb diese Architektur verwendet.

Wird das Netzwerk nur auf die allgemeine Erkennung von Tieren trainiert (Abb. 5-A), so wird im Inferencing eine enorm hohe Detektionsquote von 92,8% pro Tier erreicht. Die Analysen zeigen sich bei der allgemeinen Detektion zudem relativ robust gegenüber Qualitätseinbußen bei den Eingabedaten (z.B. Bewegungsunschärfe und Verschattungen), da die Tiersignaturen zumeist in mindestens einem Frame korrekt erkannt werden und so für die Erfassung genutzt werden können.

	Allgemeine Detektion	Klasse Damwild*	Klasse Rotwild	Klasse Ziegen- artige	A Ziegenartige: 99% B
Detektierbarkeit pro Tier (mindestens in einem Frame)	92,8 %	96,9 %	84,2 %	93,8 %	Ter 99%
Detektierbarkeit pro Einzelframe	82,7 %	76,2 %	70,8 %	65,9 %	Auto: 99%
Anteil Falsch- Positive an allen Detektionen in der gleichen Klasse	0,5 %	21,7 %	6,4 %	0,9 %	C Provide cost
Anzahl Signaturen in der Auswertung	2 576	1 347	1 032	1 457	
Anzahl Bounding Boxes im Training	1 628	647	353	323	TDamwidd: 54%

Abb. 5: Auswertung der Detektionsergebnisse auf Video-Feeds, welche einen Live-Überflug simulieren. Links: Tabelle der Zählstatistiken (*Klasse Damwild enthält Damwild, Sikawild und Axiswild). Der Anteil Falsch-Negativer Detektionen ist nicht aufgeführt, da er dem Kehrwert der Detektierbarkeit entspricht. Rechts: Beispiele per Inferenz errechneter Bounding Boxes.

Bei der Klassifikation nach Tierart fällt die Präzision leicht ab (Abb. 5-BCD), übertrifft jedoch für die Klassen *Damwild*, *Rotwild* und *Ziegenartige* die Erfolgsraten konventioneller Erfassungs-methoden noch immer bei weitem. *Wisente* und *schottische Hochlandrinder* erreichen ebenfalls eine sehr hohe Präzision (Detektierbarkeit >90% pro Tier, >80% pro Frame bei n≈150 trainierten Bounding Boxes). Diese Werte sind jedoch nicht mit denjenigen von Wildtieren in naturnahen Mischwald-Gehegen vergleichbar, da sowohl für das Training, als auch für das Inferencing nur Daten in offener Weidelandschaft zur Verfügung standen (Vgl. Weide in Abb. 5-A; Mischwald in Abb. 5-BCD). Die Klassen *Schwarzwild*, *Menschen* und *kleinere Säugetiere* erreichten aufgrund einer zu niedrigen Anzahl an Trainingsdaten (n<60) keine ausreichende Klassifikationspräzision.

5 Diskussion und Ausblick

UAV-basierte Infrarot-Thermographie mit State-of-the-Art-Verfahren des Deep Learnings zu verbinden, weist auf das Potential hin, Effizienz- und Qualitätssteigerungen bei der Populationsschätzung zu erreichen. Der aktuelle Standard – ein aufwändiger, personalintensiver Prozess der Scheinwerfer-Suche, bei welchem viele Kilometer Waldwege abgefahren werden, um nur einen kleinen, unbekannten Anteil der Tiere zu kartieren – ließe sich so mit modernen Methoden der Mustererkennung ergänzen. Der implementierte Prototyp erreicht auf mobiler Hardware (2016 Consumer-Grade Laptop) eine Inferencing-Performanz von ca. 8 FPS. Damit ist das System so effizient, dass es während des Flugs auf einen thermalen Live-Videofeed angewandt werden kann. Diese vielversprechenden Ergebnisse zeigen somit, dass eine Ablösung der klassischen Erfassungsmethoden für bestimmte Gebiete in Zukunft denkbar ist.

Ein weiterer wichtiger Anwendungsfall ist die Rehkitzrettung. In Mähwiesen versteckte Rehkitze fallen aufgrund ihres Drückreflexes häufig den Mähdreschern zum Opfer. Werden heute Thermal-UAVs eingesetzt, läuft der Prozess jedoch noch weitgehend manuell ab. Zudem ist das Training der Piloten zur Signatur-Erkennung aufwändig. Die vorgestellte Softwareautomatisierung kann die UAV-basierte Rehkitzrettung in Zukunft daher erheblich verfügbarer machen.

Bei der Wildschadensprävention stehen zumeist Wildschweinrotten im Fokus, die sich von außen unsichtbar einen Rückzugsort in Ackerkulturen trampeln. Die Verortung der Tiere ist mit dieser Technologie noch vor der Entstehung eines größeren Schadens möglich. Sowohl für die Wildschadensprävention, als auch für die Rehkitzrettung bedarf es zum operativen Einsatz noch zusätzlicher Trainingsdaten. Sind diese erhoben und markiert, kann das bestehende Deep-Learning-Netzwerk per *Fine Tuning* weitertrainiert werden und so auf dem bereits erworbenen Wissensschatz aufgebaut werden.

Durch die schnellen Fortschritte in der UAV-Technologie ist es durchaus denkbar, dass kleinere Multikopter bald leiser fliegen können und somit weniger Störungen des Tierverhaltens auslösen. Mit einer niedrigeren Flughöhe und stärkeren Sensoren wären diese in der Lage, noch bessere Thermogramme zu generieren, was wiederum die Signaturklassifikation für das neuronale Netzwerk erleichtert. Denkbar wäre die Identifikation weiterer Individualmerkmale der bereits analysierten Arten wie Alter und Geschlecht, oder aber auch eine Ausweitung der Analyse auf kleinere Spezies wie Dachse, Feldhasen und Füchse, sowie seltene Spezies wie Luchse und Wölfe.

6 Literaturverzeichnis

ALEMI, A., 2016: Improving Inception and Image Classification in Tensorflow. Google AI Blog. GOOGLE, 2018: Google Earth Pro 7.3.1, Luftbild-Textur: GeoBasis-DF/BKG, 2017-08-07.

GONZALEZ, L., MONTES, G., PUIG, E., JOHNSON, S., MENGERSEN, K. & GASTON, K., 2016: Unmanned Aerial Vehicles (UAVs) and Artificial Intelligence Revolutionizing Wildlife Monitoring and Conservation. Sensors, 16 (1), Beitrag 97. 39. Wissenschaftlich-Technische Jahrestagung der DGPF und PFGK19 Tagung in München – Publikationen der DGPF, Band 28, 2019

- GILLINS, D., PARRISH, C., GILLINS, M., H. & SIMPSON, C., 2018: Eyes in the Sky: Bridge Inspections with Unmanned Areal Vehicles. Oregon Dept. of Transportation, SPR 787 Final Report.
- SENSEFLY, 2018: https://www.sensefly.com/drone/ebee-mapping-drone/ (6.5.2018).
- SILVEIRA, L., JACOMO, A. & DINIZ-FILHO, J., 2003: Camera trap, line transect census and track surveys: a comparative evaluation. Biological Conservation, **114** (3), 351-335.
- SZEGEDY, C., VANHOUCKE, V., IOFFE, S., SHLENS, J. & WOJNA, Z., 2016: Rethinking the inception architecture for computer vision. IEEE Conference on Computer Vision and Pattern Recognition, 2818-2826.

Trennung von parkenden und am Verkehr teilnehmenden Fahrzeugen basierend auf einer automatischen Verkehrserfassung aus Luftbildern

JULIA KNÖTTNER¹, DOMINIK ROSENBAUM² & ANSGAR BRUNN¹

Zusammenfassung: In diesem Beitrag wird eine Methode entwickelt, die auf Basis der automatischen Verkehrserfassung des Deutschen Zentrums für Luft- und Raumfahrt e.V. (DLR), stehende Fahrzeuge in "parkend" und "am Verkehr teilnehmend" unterteilt. Bei der Erfassung, die Teilbereich des Projektes VABENE++ ist, werden während eines Helikopterfluges Fahrzeuge im Bild detektiert und über eine Bildsequenz getrackt, sodass anschließend für jedes wiedergefundene Fahrzeug mehrere Positionen vorliegen. Im Zuge dessen werden zahlreiche Verkehrsparameter gewonnen. Anhand dieser Daten wird die finale Unterscheidung mithilfe von Fuzzy-Logik realisiert. Neben den gegenwärtigen Umgebungseinflüssen werden dabei auch die benachbarten Fahrzeuge mit einbezogen.

1 Motivation

Am Deutschen Zentrum für Luft- und Raumfahrt in Oberpfaffenhofen e.V. (DLR) werden für das Projekt VABENE++ Methoden zur offiziellen Umsetzung notwendiger Rettungslogistik bei Katastrophen und Großveranstaltungen entwickelt (ROSENBAUM et al. 2011, GSTAIGER et al. 2015, GSTAIGER et al. 2016). Dazu zählen das Verkehrsmonitoring und die Lageerfassung. Im Zuge dessen werden während eines Fluges Verkehrsdaten aus Bildsequenzen extrahiert und anschließend zu einer mobilen Bodenstation übertragen.

Die Aufnahme, die im Burst-Mode realisiert wird, erfolgt mit einem vom DLR entwickelten 4k-Kamerasystem (KURZ et al. 2014). Innerhalb eines Bursts, der sich alle sieben Sekunden aus einer Serie von drei Bildern zusammensetzt, werden alle abgebildeten Fahrzeuge detektiert und anschließend getrackt. Die Fahrzeugdetektion erfolgt bei jedem ersten Bild. Das Tracking bezieht sich im Gegensatz dazu auf ein Bildpaar innerhalb eines Bursts. Für jedes getrackte Fahrzeug existieren infolgedessen mehrere Positionen (LEITLOFF et. al 2014).



Abb. 1: Im Zuge der automatischen Verkehrserfassung des DLR wird eine Visualisierung der Fahrgeschwindigkeit und Fahrtrichtung ermöglicht.

¹ Hochschule für angewandte Wissenschaften, Labor für Photogrammetrie & Fernerkundung, Röntgenring 8, D-97070 Würzburg, E-Mail: [Julia.Knoettner; Ansgar.Brunn}@gmx.de

² Deutsches Zentrum für Luft- und Raumfahrt e.V., Institut für Methodik der Fernerkundung, Münchener Str. 20, 82234 Weßling, E-Mail: Dominik.Rosenbaum@DLR.de

Anhand dieser gewonnenen Daten ist eine sofortige Ermittlung zahlreicher Verkehrsparameter möglich. Dazu zählen neben der Geschwindigkeit auch die zurückgelegte Strecke und die Fahrtrichtung. Mit Hilfe dieser Daten ist im Falle einer Katastrophe oder Großveranstaltung eine sofortige Darstellung der aktuellen Verkehrssituation möglich.

Damit das VABENE-System allerdings zukünftig auch für eine Betrachtung der Parkraumbelegung innerhalb der Großstädte eingesetzt werden kann, wird in diesem Beitrag eine Methodik entwickelt, die basierend auf der vorangehenden Verkehrsprozessierung alle stehenden Fahrzeuge aufsucht und anschließend eine Unterteilung in "am Verkehr teilnehmend" und "parkend" vollzieht. Darüber hinaus ist durch diese Methodik auch eine Bereinigung der Verkehrsdaten möglich, in dem die innerorts am Straßenrand entlang der Fahrtrichtung parkenden Fahrzeuge, die bisher als stehender Verkehr miterfasst wurden, aus den eigentlichen Verkehrsdaten entfernt werden. Dabei bleiben verkehrsbedingt stehende Fahrzeuge, wie z.B. im Stau oder vor einer roten Ampel, weiterhin in den Verkehrsdaten enthalten. Diese Klassifizierung der Fahrzeuge in "parkend" und "am Verkehr teilnehmend" basiert neben der Fuzzy-Logik auch auf vermessungstechnischen Grundüberlegungen und wurde unter Verwendung der Programmierbibliothek OpenCV vollzogen. Daher war neben der Statistik auch ein Bezug zur Geoinformatik ein wichtiger Grundstein für die Entwicklung.

2 Methodik

Als Datengrundlage dienen die vorab ermittelten Verkehrsparameter. Die finale Unterscheidung wird hauptsächlich unter Betrachtung der Nachbarfahrzeuge auf demselben Straßenabschnitt realisiert.



Abb. 2: Die Methodik für die Unterscheidung in "am Verkehr teilnehmend" und "parkend" unterliegt dem dargestellten Prozessdiagramm.

Aus diesem Grund werden Fahrzeugschlangen generiert, die sich aus hintereinanderfahrenden Fahrzeugen desselben Straßenabschnittes zusammensetzen. Darüber hinaus ist jede generierte Fahrzeugschlange parallel zur Mittelachse der Straße und verfügt nur über Fahrzeuge derselben Fahrtrichtung. Im Anschluss daran wird die endgültige Unterteilung mithilfe von Fuzzy-Logik vollzogen. Hierfür wird separat für die Fahrzeugschlangen und Fahrzeuge ohne Nachbarschaft ein Fuzzy-System aufgebaut. Innerhalb eines solchen Systems fließen die Verkehrsparameter und die unmittelbare Umgebung der Fahrzeuge in die endgültige Bewertung mit ein.

2.1 Identifikation der Fahrzeugschlangen

Um Nachbarschaftsbeziehungen zwischen den Fahrzeugen untereinander herstellen zu können, wird zu Beginn eine Delaunay Triangulation durchgeführt. Alle Fahrzeugpositionen werden hierbei zu Dreiecken vermascht. Das Ergebnis ist ein Delaunay Graph, der aus Knoten und Kanten besteht (WESSLER & RÖPCKE 2015: 17). In diesem Fall repräsentieren die Knoten die Fahrzeugpositionen und die Kanten die Verbindungslinien zwischen den benachbarten Fahrzeugen. Unter der Annahme einer dichten Fahrzeugfolge richten sich die Kanten des Graphens vorzugsweise entlang der Fahrwege aus. Weitere Kanten kreuzen die Bereiche zwischen den Straßen.



Abb. 3: Mithilfe der Delaunay Triangulation werden alle Fahrzeuge miteinander vernetzt.

Unter Verwendung der Kanten werden für alle Fahrzeuge die unmittelbaren Nachbarfahrzeuge ermittelt. Um den Delaunay Graph anschließend auf Fahrzeugschlangen reduzieren zu können, wird für jedes Fahrzeug der zugehörige Straßenabschnitt lokalisiert. Im Anschluss daran wird für jede NAVTEQ Linie die zugehörige Fahrtrichtung ermittelt.



Abb. 4: NAVTEQ Straßenabschnitte stellen nahezu die Mittelachse der Straße dar.

Daraufhin wird der Graph auf Fahrzeugschlangen reduziert. Dabei wird das Ziel verfolgt, nur Verbindungslinien zwischen Fahrzeugen zu erhalten, die demselben Straßenabschnitt angehören. Das Nachbarfahrzeug sollte sich demzufolge direkt vor oder dahinter mit derselben Fahrtrichtung befinden. Um dies sicherzustellen, müssen Kanten aus dem Delaunay Graph eliminiert werden. Hierfür werden folgende Kriterien aufgestellt, die eine solche Kante erfüllen muss:

- Kantenrichtung ≈ NAVTEQ Richtung
- Fahrtrichtung \approx NAVTEQ Richtung
- Pro NAVTEQ ID: Pro Fahrzeug nur ein Nachbarfahrzeug
- Pro NAVTEQ ID: Fahrzeug fungiert nur einmal als Nachbar

Die Kantenrichtung ist der Richtungswinkel (GRUBER & JOECKEL 2014, S. 40) vom Fahrzeug zum Nachbarfahrzeug. Dieser Winkel muss nahezu mit der NAVTEQ Richtung übereinstimmen. Dasselbe betrifft die Fahrtrichtung. Für die Erzeugung von Fahrzeugschlangen ist es notwendig pro Fahrzeug nur ein zugehöriges Nachbarfahrzeug zuzulassen. Ebenso darf ein Fahrzeug nur einmal als Nachbarfahrzeug fungieren. Da eine Fahrzeugschlange eine nahezu gerade verlaufende Linie darstellt, ist diese zusätzliche Bedingung erforderlich. Falls eine Kante einer dieser Kriterien nicht erfüllt, wird das Nachbarfahrzeug aus der bestehenden Nachbarschaftsbeziehung ausgeschlossen. Demzufolge werden Fahrzeugschlangen generiert, die nahezu parallel zur Mittelachse der Straße verlaufen.



Abb. 5: Die Fahrzeuge werden zu Fahrzeugschlangen gruppiert.

2.2 Identifikation von Straßenkreuzungen und Ampeln

Für die Unterscheidung in "parkend" und "am Verkehr teilnehmend" ist des Weiteren ein Bezug zu den umliegenden Straßenkreuzungen und Ampeln erforderlich. Fahrzeuge, die sich im nahen Umkreis einer Straßenkreuzung oder Ampel befinden nehmen mit einer höheren Wahrscheinlichkeit am Verkehr teil. Im Gegenzug zu den Ampeln, deren Positionen bekannt sind, müssen die Straßenkreuzungen auf Basis der NAVTEQ Daten erstmalig mittels Template Matching lokalisiert werden. Auf Basis dieser Daten wird im Anschluss geprüft, ob sich Fahrzeuge innerhalb eines solchen Umkreises befinden. Hierfür wird ein separates Binärbild erzeugt, in diesem die Ampel- und Straßenkreuzungen als Kreise repräsentiert werden. Der jeweilige Radius wird durch die Fahrbahnanzahl der umliegenden Straßenabschnitte beeinflusst. Daraufhin werden die Pixelwerte der Fahrzeuge im Binärbild abgegriffen.

2.3 Fuzzy-Logik

Die finale Unterscheidung in parkend und am Verkehr teilnehmend wird abschließend mithilfe der Fuzzy-Logik vollzogen. Für die Fahrzeugschlangen und Fahrzeuge ohne Nachbarschaft wird dabei ein separates Fuzzy-System aufgebaut. Anhand definierter Regeln werden innerhalb eines Systems verschiedene Sachverhalte miteinander verknüpft und bewertet (BOTHE 1995, S.1). Dies sind die Fahrgeschwindigkeit, Fahrzeugdichte innerhalb einer Fahrzeugschlange, Straßenkreuzungen, Ampeln, sowie die Straßen- und Fahrbahnkategorie. Das System erhält und übergibt scharfe Werte. Im Zuge der Fuzzifikation und Defuzzifikation erfolgt eine Umwandlung der Werte in das jeweils andere System, sodass innerhalb des Systems mit unscharfen Werten gearbeitet wird (BOTHE 1995: 142).

Im Zuge der Fuzzifikation werden die Fahrzeugschlangen und Fahrzeuge innerhalb verschiedener Fuzzy-Systeme betrachtet. Für die Bewertung der örtlichen Gegebenheiten werden bei der Fuzzifikation linguistischen Variablen definiert. Die Variable Geschwindigkeit ist Teil des Systems. Unter Berücksichtigung der Genauigkeit in der Fahrzeugerfassung werden in dieser Arbeit alle Fahrzeuge mit einer Geschwindigkeit von weniger als 5km/h als stehend angesehen (HINZ et al. 2007). Da Straßenkreuzungen und Ampeln das Fahrverhalten der Verkehrsteilnehmer beeinflussen, ist dies ebenso eine linguistische Variable. Diesbezüglich erfolgte für alle Fahrzeug eine Prüfung, ob in unmittelbarer Umgebung eine Straßenkreuzung oder Ampel existiert. Der Sachverhalt wird anschließend unter Berücksichtigung der Anzahl der Fahrbahnen im System mitberücksichtigt. Eine weitere Einflussgröße ist die Straßenkategorie. Die Straßenabschnitte wurden bezüglich ihrer Funktion und der zulässigen Fahrgeschwindigkeit verschiedenartig klassifiziert (DIGITAL MAP TECHNOLOGY 2017). Die Straßenkategorie ist ausschlaggebend für die besondere Berücksichtigung von Fahrzeugen innerhalb von Wohngebieten. Diese werden innerhalb des Systems speziell gewichtet. Der letzte wichtige Bestandteil ist die Fahrzeugdichte innerhalb der Fahrzeugschlange. Diese Variable ist nur Teil des Systems der Fahrzeugschlangen und gibt Auskunft darüber, wie dicht die Fahrzeuge hintereinanderstehen oder -fahren. Stehen Verkehrsteilnehmer an einer Straßenampel oder Kreuzung sehr dicht hintereinander, so kann davon ausgegangen werden, dass diese verkehrsbedingt stehen. Ein hierfür geeigneter Dichtewert liegt bei mindestens 0,15 (VAN AERDE & RAKHA 2002).

Zur Verknüpfung der einzelnen Sachverhalte werden Fuzzy-Regeln in einer *Wenn-Dann*-Form formuliert (NOLL 2009: 49). Innerhalb jeder Regel werden die linguistischen Variablen mit einem "und" miteinander verknüpft.

	THEN			
Geschwindigkeit	Dichte	Einfluss	Straßenkategorie	Parkend
Langsam	Niedrig	Niedrig	-	Hoch
Langsam	Niedrig	Hoch	-	Niedrig
Langsam	Hoch	Hoch	-	Niedrig
Schnell	-	-	-	Niedrig
Langsam	Hoch	Niedrig	-	Hoch
Langsam	Niedrig	Mittel	Niedrig	Hoch
Langsam	Niedrig	Mittel	Hoch	Niedrig
Langsam	Hoch	Mittel	Hoch	Niedrig
Langsam	Hoch	Mittel	Niedrig	Hoch

Tab. 1: Regeln für Fahrzeugschlangen

	THEN		
Geschwindigkeit	Ein-	Straßenkategorie	Parkend
Langsam	Niedrig	-	Hoch
Langsam	Hoch	-	Niedrig
Langsam	Mittel	Hoch	Niedrig
Langsam	Mittel	Niedrig	Hoch
Schnell	-	-	Niedrig

Tab. 2: Regeln für Fahrzeuge ohne Nachbarschaft

Im Anschluss daran wird der Zugehörigkeitsgrad zur Regelbasis ermittelt. Die Vektordimension entspricht dabei der Anzahl der jeweiligen Regeln. Hierbei werden die Werte für die in der Regel stehenden linguistischen Terme ermittelt. Da es sich hierbei um eine *"und"-Ver*knüpfung handelt, wird für den Zugehörigkeitsgrad zur jeweiligen Regel das Minimum von allen Werten herangezogen (NOLL 2009: 64). Im Zuge der Akkumulation werden die Teilergebnisse mit einer *Oder-*Verknüpfung zu einem endgültigen Fuzzy-Set kombiniert. Da es sich bei "parkend" oder "am Verkehr teilnehmend" um eine *Oder-*Verknüpfung handelt, wird der Maximum-Operator verwendet (MAYER et al. 1993: 79). Die Ausgangsvariable hierfür ist die linguistische Variable "Parkend" mit den linguistischen Termen "niedrig" und "hoch", wobei "niedrig" für am Verkehr teilnehmend" und "hoch" für parkend steht.



Abb. 6: Mithilfe der linguistischen Variable "Parkend" wird die finale Unterscheidung in "am Verkehr teilnehmend" und "parkend" vollzogen.

Mithilfe der daraus resultierenden Fuzzy-Menge werden für die Fahrzeuge die scharfen Ausgangswerte ermittelt. Für diese Defuzzifikation wird die Maximum Mittelwertverfahren verwendet, sodass sich der scharfe Ausgangswert aus dem arithmetischen Mittel der Maxima ergibt (BOTHE 1995: 146).

3 Prozesskette & Experimente

Die Methode wurde in das operationelle System der Verkehrserfassung aus Luftbildsequenzen übernommen. Die modular aufgebaute Prozesskette besteht somit aus einem Kameramodul, einem Orthoprozessor, dem Verkehrsdetektor (MATTYUS et al. 2013, LEITLOFF et al. 2014) und dem hier vorgestellten Prozess zur Trennung von parkenden und stehenden Fahrzeugen. Diese Module laufen in Echtzeit auf den Computern des 4k Systems an Bord des Hubschraubers. Über den im 4k Kamerasystem installierten C-Band Datenlink werden dann die erzeugten Verkehrsdaten samt Klassifikation in "parkend" oder "stehend" nahezu verzögerungsfrei zu einer

Bodenstation übertragen. Von dort aus können diese den Nutzern des Systems, wie z.B. den Einsatzkräften während eines Großereignisses, zur Verfügung gestellt werden.

Das Kameramodul verbindet dabei die Luftbildaufnahmen mit den Globalen Navigationssatellitensystem (GNSS)- und Fluglagedaten, um die Bilder im nächsten Schritt im Orthoprozessor unter Berücksichtigung eines geeigneten Digitalen Höhenmodells zu georeferenzieren und zu orthorektifizieren. Der auf Maschinen-Lernalgorithmen basierende Verkehrsprozessor detektiert im nächsten Schritt in den orthorektifizierten Luftbildern die Fahrzeuge, welche in überlappenden Bereichen der Luftbildsequenzen dann per template-matching getrackt werden. Unterstützt werden die Detektion und das Tracking durch die Berücksichtigung einer Straßendatenbank, wie NAVTEQ oder OpenStreetMap. Somit liegen für die detektierten und getrackten Fahrzeuge die Parameter wie Position, Fahrtrichtung und Geschwindigkeit vor, welche dann im nächsten Schritt in "parkend" oder "verkehrsbedingt stehend" klassifiziert werden. Demnach ist es möglich, die Verkehrsdaten von den parkenden Fahrzeugen zu bereinigen oder Parkräume entlang der Straßen zu identifizieren.



Abb. 7: Mithilfe der Daten ist eine Identifikation der Parkräume möglich.

Durch einen Defekt am Fluglage- & GNSS System, der während des zur Evaluierung der Online-Prozesskette durchgeführten Testflugs am 4.10.2018 auftrat, konnte diese mit der hier vorgestellten Erweiterung bisher noch nicht im Flug getestet werden. Die erweiterte Prozesskette wurde jedoch ausführlich im Postprocessing mit Archivdaten von einem Flug mit dem 4k System am 30.3.2017 über München getestet. Dabei wurden 588 Bilder prozessiert, die über 4 verschiedenen Kreuzungen mit Lichtsignalanlagen zwischen 13:21 Uhr und 13:50 Uhr UTC aufgenommen wurden.

4 Fazit & Ausblick

Die Unterscheidung unterliegt dem Einfluss der Verkehrserfassung, da mögliche Fehldetektionen in das Endergebnis mit einfließen. Darüber hinaus kann das Ergebnis durch die Bildung von Nachbarschaftsbeziehungen ebenso negativ beeinträchtigt werden. Ist innerhalb einer Fahrzeugschlangen eine Bewegung, so werden alle Fahrzeuge als "am Verkehr teilnehmend" markiert. Die Genauigkeit der NAVTEQ Linien ist ebenso erforderlich, da anhand dieser Daten die Straßenkreuzungen ermittelt werden. Eine Situation vor einer Ampel ist in den nachfolgenden Abbildungen dargestellt.



Abb. 8: Fahrzeuge, die vor einer Ampel verkehrsbedingt stehen werden vom Algorithmus als "am Verkehr teilnehmend" klassifiziert.

Die vorderen Fahrzeuge stehen, wobei die hinteren Fahrzeuge eine geringe Fahrgeschwindigkeit aufweisen. Die Fahrzeuge auf der mittleren Spur wurden von der automatischen Verkehrserfassung nicht erkannt. Da die Fahrzeuge sich direkt vor einer Ampel befinden und darüber hinaus dicht hintereinanderstehen, werden sie mithilfe der Methodik als "Am Verkehr teilnehmend" markiert. Das Resultat zeigt die rechte Abbildung. Die am Verkehr teilnehmenden Fahrzeuge werden hier als grüne Punkte dargestellt.

In diesem Beitrag wurde ein gut geeigneter Algorithmus entwickelt, mit dem die Verkehrsdaten neben einer Verkehrssimulation zukünftig auch zur Betrachtung der Parkraumbelegung innerhalb der Großstädte genutzt werden können.

5 Literaturverzeichnis

- BOTHE, H., 1995: Fuzzy Logic Einführung in Theorie und Anwendung, 2. Auflage, Springer-Verlag, Berlin.
- DAM, E.: Digital Map Technology Has Arrived, http://www.cicomra.org.ar/cicomra2/expocomm/tutorial%208%20dam-%20navteq.pdf, letzter Zugriff 21.11.2017.
- GRUBER, F. & JOECKEL, R., 2014: Formelsammlung für das Vermessungswesen, 17. Auflage, Springer Vieweg, Wiesbaden.
- GSTAIGER, V., RÖMER, H., ROSENBAUM, D. & HENKEL, F., 2015: Airborne Camera System for Real-Time Applications - Support of a National Civil Protection Exercise. The Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., 40(7/W3), 1189-1194.
- GSTAIGER, V., RONALD NIPPOLD, R. & KIEFL, R., 2016: Forschungsprojekt VABENE++: Verkehrsmanagement bei Großereignissen und Katastrophen. Im Einsatz (6), 34-36, Verlag Stumpf & Kossendey GmbH
- HINZ, S., KURZ, F., WEIHING, D., SUCHANDT, S., MEYER, F. & BAMLER, R., 2007: Evaluation of Traffic Monitoring based on Spatio-Temporal Co-Registration of SAR Data and Optical Image Sequences. Photogrammetrie Fernerkundung Geoinformation, 2007(5), 309-325.
- KURZ, F., ROSENBAUM, D., MEYNBERG, O. & MATTYUS, G., 2014: Real-time mapping from a helicopter with a new optical sensor system. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 23, 1-8.
- LEITLOFF, J., ROSENBAUM, D., KURZ, F., MEYNBERG, O. & REINARTZ, P., 2014: An Operational System for Estimating Road Traffic Information from Aerial Images. Remote Sensing, 6(11), 11315-11341.
- MATTYUS, G., KURZ, F. & ROSENBAUM, D., 2013: A real-time optical airborne road traffic monitoring system. Hungarian Association for Image Processing and Pattern Recognition, 645-656.
- MAYER, A., MECHLER, B., SCHLINDWEIN, A. & WOLKE, R., 1993: Fuzzy Logic. 1. Auflage, Addison-Wesley, Bonn.

- NOLL, P., 2009: Statistisches Matching mit Fuzzy Logic. 1. Auflage, Vieweg+Teubner, Marburg.
- ROSENBAUM, D., BEHRISCH, M., LEITLOFF, J., KURZ, F., MEYNBERG, O., REIZE, T. & REINARTZ, P., 2011: An airborne camera system for rapid mapping in case of disaster and mass events. Proceedings of the Earth Observation for Global Change, 1-5.
- VAN AERDE, M. & RAKHA, H., 2002: Multivariate Calibration of Single Regime Speed-Flow-Density Relationships. IEEE Xplore, DOI: 10.1109/VNIS.1995.518858.
- WESSLER, M. & RÖPCKE, H., 2015: Graphen und Netzwerktheorie. 1. Auflage, Carl Hanser Verlag München, München.

Classification of 3D Point Clouds using Deep Neural Networks

LUKAS WINIWARTER^{1,2} & GOTTFRIED MANDLBURGER^{1,3}

Abstract: The use of deep neural networks, i.e. neural networks with multiple layers, has significantly improved the accuracy of classification and regression tasks in many disciplines, including computer vision. Here, especially convolutional neural networks (CNNs) that are able to extract features representing local neighborhoods in 2D images are employed. The direct transfer of this approach to 3D point cloud data requires a rasterization and manual selection of attributes that are represented in the raster.

We present a neural network based on PointNet by QI et al. (2017a) that instead directly uses 3D point clouds, along with any attributes attached to the points, as input, considering neighborhoods of points. On the ISPRS 3D Semantic Labeling Benchmark Vaihingen, we achieve an overall accuracy of 80.6 %. On the airborne laser scanning point cloud of the Federal State of Vorarlberg, Austria, we achieve accuracies in urban areas of up to 95.8 %, observing a strong correlation with land cover. Hence, we conclude that our approach learns a representation of neighborhood mostly equivalent to the current manual selection process, with the option of further improvement by use of additional data such as echo width or RGB information.

1 Introduction

Topographic 3D point clouds representing the surface of the earth and objects thereon can be acquired with different methods. The most established ones are Dense Image Matching (DIM) based on multi-view stereo images and topographic laser scanning or LiDAR (Light Detection And Ranging), respectively. Both methods lead to big amounts of data, often in the order of several terabytes. Especially with advances in sensor technology and automation concerning image matching, these volumes have grown even further. To make use of these data, classification is key. By classification, we refer to the assignment of semantic labels to points, in this case, on a per-point basis (OTEPKA et al. 2013). The nature of these labels depends on the task at hand. In our case, they represent types of objects commonly found on the earth's surface: vegetation, buildings, water, ground, and others (ASPRS 2011).

Section 2 shows that neural networks have been used in point cloud classification, however, current methods do not fully exploit their potential (YOUSEFHUSSIEN et al. 2018). Like with other classifiers, existing neural network approaches rely on the user to select and calculate a number of features characterizing the local neighborhoods of every point. A study on the relevance of these

¹ Technische Universität Wien, Department für Geodäsie und Geoinformation, Gusshausstraße 25-27, A-1040 Wien

² Universität Heidelberg, Geographisches Institut, Im Neuenheimer Feld 368, D-69120 Heidelberg, E-Mail: lukas.winiwarter@uni-heidelberg.de

³ Universität Stuttgart, Institut für Photogrammetrie, Geschwister-Scholl-Straße 24D, D-70174 Stuttgart, E-Mail: gottfried.mandlburger@ifp.uni-stuttgart.de

features was published by WEINMANN et al. (2013), showing that the structure tensor, i.e. the matrix of the second-order moments of the point distribution, is key. To achieve scale-, translationand rotation-invariance, the moments are centralized and the eigenvalues of the structure tensor are calculated. Ratios or other simple algebraic combinations of these eigenvalues represent measures like *linearity* or *planarity*, with values ranging between zero for non-linear/non-planar objects to one for elongated/planar distributions of points, respectively.

The user does not only have to decide on the specific features, but also on the neighborhood and attributes, these features are calculated from. For the neighborhood, two main approaches exist: a fixed distance (spherical) query, or a fixed number (k-Nearest-Neighbors, kNN) query (WEIN-MANN et al. 2014). Other attributes in the point cloud may come from the sensor itself (e.g. amplitude, echo width, echo number etc.) or from data fusion (e.g. RGB values from an orthophoto). In addition to the spatial distribution of the points, also the spatial distribution of these attributes can be considered for classification (OTEPKA et al. 2013).

2 State of the Art

Many different classification algorithms are commonly applied to point clouds (GRILLI et al. 2017). Here, we present two prominent machine learning classifiers, Random Forest and Support Vector Machine. Subsequently, we introduce neural networks, including their application on point clouds.

2.1 Supervised Classification

Supervised classification is based on training data. Training data is input data for the classification, where the expected result, which is the class of the data point, is known *a priori*. This reference classification can be provided by other classification algorithms, by manual classification, by a combination of the two, or by simulated data. A classifier trained on such data will attempt to approximate this reference as good as possible by minimizing the training error. This sometimes leads to questionable results, especially if the reference classification contains errors or inconsistencies. The latter is often referred to as classification under label noise. Furthermore, two classes that cannot be separated in feature space will result in outputs where the classes are mixed up or the class with less presence in the training data (i.e. prior probability) does not appear in the output (FRÉNAY & VERLEYSEN 2014).

2.2 Support Vector Machine

Support Vector Machines, or SVMs, are binary classifiers that separate a dataset into two classes. The basic idea is to find a best linear separation boundary (i.e. a hyperplane) between the classes in feature space. Since the restriction to linearly separable classes limits this classifier, the input data is transferred to a space of higher dimensionality. Here, a better linear separation might be possible. To save on computation cost, the "kernel trick" is used, where the data does not have to be transformed itself. In the high-dimensional space, the hyperplane is fitted so that the distance

from the hyperplane to the nearest feature point (the support vector) of each class is minimized (HEARST et al. 1998).

2.3 Random Forest

Random forests are based on decision trees. In a decision tree, an input dataset is fed through threshold classifiers represented by Boolean expressions, so-called nodes (e.g. "amplitude > 50"). Every node has two outputs (true/false), which are connected to new nodes. The end-nodes, called leaves, represent the class the dataset is assigned to, allowing more than two target classes.

Classic decision trees require threshold settings in every node. They also do not provide any measure of uncertainty for their classes. Random forests consist of multiple trees, each having been picked as the best one from a number of randomly generated trees. The input dataset is then passed through all trees simultaneously, which vote on the class. This leads to non-linear boundaries as well as a probability measure of the output class. For example, if 60 out of 100 trees vote for class A, the probability of the point belonging to class A is 60 %. Even though single decision trees are not very good classifiers, the aggregation over the forest leads to reasonable results (BREIMANN 2001).

2.4 Neural Networks

Neural networks represent another type of machine learning algorithms. They emulate the way the nervous system works by connecting multiple atomic units, the neurons. Every neuron has a number of input values I from the previous neurons, which are aggregated into an output O, that is passed to the next neurons. The simple neural model of the *Perceptron* uses a weighted sum of the inputs and a bias. The weights w and the bias b are not known *a priori*. They are initialized randomly and are then adapted in the machine learning process, i.e., during training (ROSENBLATT 1958). To overcome the linearity restraint of the weighted summation, the output value is fed through a non-linear activation function a, such as a sigmoid function, before it is passed to the next neuron. A single neuron can be described mathematically as in Eqn. 1 (GOODFELLOW et al. 2016, p. 171).

$$0 = a(\sum_{i=1}^{n} w_i \cdot I_i + b) \tag{1}$$

In a multi-layer perceptron, a very basic neural network, the neurons are organized in layers, where every neuron from the first layer is connected to every neuron from the second layer, and so on. The first layer is referred to as the input, the last layer as the output of the network. All layers in between are so-called hidden layers. The amount of hidden layers defines the depth of the network, coining the term "Deep Learning", while the amount of neurons per layer the width, which may vary from layer to layer (GOODFELLOW et al. 2016: 169). In this simple structure, a whole layer can be represented mathematically as a matrix operation with the weights summarized in a coefficient matrix *A*, as shown in Eqn. 2. The activation function *a* is applied to every entry of the vector resulting from the matrix multiplication. For multiple layers, this operation is chained together.

$$\vec{O} = a \left(A \cdot \vec{I} + \vec{b} \right) \tag{2}$$

For training, a reference sample is fed through the network, and the output is compared with the expected output. From this, an error measure ("loss") is created, e.g. using a cross-entropy function. The effect of the weights on this loss can be calculated as the partial derivative of the loss with respect to the weights and biases. For this, repeated application of the chain rule, as well as a differentiable activation function are required. Using gradient descent, the weights are adapted in a way to find a minimum value for the loss (GOODFELLOW et al. 2016: 82 f.). Advanced optimization algorithms such as the Adam optimizer incorporate a momentum approach, where not the weights directly are updated, but rather a velocity vector, which then, in turn, changes the weight values. The Adam optimizer also includes a rescaling of the gradients, making use of the variance of the weight velocity updates, i.e. the second order moments (GOODFELLOW et al. 2016: 308).

2.5 Convolutions and Convolutional Neural Networks

A convolution is a mathematic operation where a function is moved over another function. The result of the convolution is the integral of the multiplied functions with a given offset (i.e. movement position) for one of the functions. In discrete form, the integral is replaced by a sum, as shown in Eqn. 3. Convolutions are commutative, i.e. the order of the functions does not matter.

$$(f * g)(x) = \sum_{x} (f(x) * g(t - x))$$
 (3)

In practice, one function, the so-called kernel, is limited in domain, i.e. the summation is only over a small part of the whole function. Furthermore, the concept can easily be extended into 2D, where the kernel is represented by a small window, a matrix that is moved over the data, e.g. an image (GOODFELLOW et al. 2016, pp. 237f.). These kernels extract information on the local neighborhood of a pixel, e.g. the slope using the well-known Sobel kernel. The values in the kernel are commonly referred to as weights.

In Convolutional Neural Networks (CNNs), the kernel is not selected for a specific purpose *a priori*, instead the weights are adjusted in the training phase. Since the kernel is moved over the input data, the weights are shared between many input values. In a neural network, this leads to fewer values that have to be trained. LECUN et al. (1989) applied this concept on multiple levels (incorporating neighborhoods of different extents) to classify handwritten ZIP codes with previously unseen success. A local pooling function is applied to extract relevant data after each convolution. This function reduces the size of the image by selecting relevant information in a small region. For example, every 2x2-pixel submatrix might be replaced by the maximum value, as shown in the example in Figure 1 (GOODFELLOW et al. 2016: 339-345).

2.6 PointNet

Point clouds are inherently unordered, irregular sets of points in space. These properties hinder the direct input of points to a neural network, since the order of the neurons matters very much. QI et

al. (2017a) published a method called PointNet able to deal with these properties. They approximate a function f acting on a point set by two functions g, and h, where g is a symmetric function (e.g. the maximum or the mean) and h is a function applied on the (relative) coordinates and attributes of every point, as shown in Eqn. 4. Furthermore, h is represented by a multi-layer perceptron.

$$f(\{x_1, x_2, \dots, x_n\}) = g(h(x_1), h(x_2), \dots, h(x_n)) = g(h_1, h_2, \dots, h_n)$$
(4)

This ensures that every input point x_i is treated equally, and the output does not depend on the order of the input points. The result of *h* and *g* are vectors, representing local features aggregated from different points in the neighborhood. In the case of the max-function, the first element of *g* may have the value from point *p*, the second element from point *q*, and so on.

We can compare PointNet to a CNN. In both cases, a local neighborhood is used to calculate new features. As Figure 1 shows, weights are applied to the local neighborhood, before a pooling operation selects relevant or representative information from the convoluted neighborhood. These similarities suggest that PointNet may profit from the same benefits as a CNN in applications where local surroundings of a data point (i.e. a pixel or a 3D point) are important.



Fig. 1: Comparison of PointNet (left) and a Convolutional Neural Network (right). Blue values are learned in the training phase, while green values are selected to be representative by the pooling function.

PointNet allows the creation of a neighborhood feature vector for a local set of points. This vector can be created on a number of scales (i.e. different neighborhood input sets) for every point in the point cloud. In PointNet++, QI et al. (2017b) suggested, however, that the values representing large-scale neighborhoods are not subject to volatile changes. Therefore, they sub-sample the point cloud using farthest distance sampling to save processing cost. On this sub-sampled point cloud, the neighborhood is defined with respect to the original point set. This process is repeated on multiple levels, each using the sub-sampled points of the previous level for the neighborhood aggregation, as shown in Figure 2. In addition, the local neighborhood feature vectors are used in the computation of the more global ones.

To obtain a classification on the original point cloud, the features have to be propagated back to the original point set. This is done on each level, where the three nearest points of the higher-level point cloud are selected and the features are interpolated, weighted by the inverse distance from the point. These propagated values are fed through another multi-layer perceptron. The result is a set of neighborhood feature vectors for each point of the original point cloud.

These vectors are concatenated and fed through a fully connected layer. To avoid overfitting the training dataset, a dropout layer is applied during training. Here, every neuron has a fixed probability of returning nothing (i.e. zero) to the next layer. Finally, a softmax regression is carried out to obtain probability values for each class per point. The class with the highest probability is then assigned to the point along with the per-class probabilities.



Fig. 2: Subsampling of the point cloud using farthest point sampling and calculation of feature vectors based on neighborhoods of increasing size.

3 Methodology

We adapted the algorithm of PointNet++ to work with data from Airborne Laser Scanning (ALS) in a workflow titled alsNet. The code for this workflow is available at <u>https://github.com/lwini-war/alsNet</u>. While PointNet and PointNet++ mainly focused on the classification of CAD-objects represented by point clouds, the semantic segmentation was only a side-result of their study. They also applied their methods to scanned datasets of indoor environments. Their applications, how-ever, remain in a local scope (QI et al. 2017b).

To apply PointNet++ to ALS data, we exploited a property of this data: ALS data follows the earth's surface, which is close to a 2-manifold; therefore, regular 2D-batches were cut out of the

dataset. To minimize possible edge effects, these batches were circular in shape, created by sampling the nearest 200,000 points around a grid point. We spaced the grid points such that every ALS point was covered by at least one batch, but four batches on average. The limitation of 200,000 points is introduced by the processing: Since the neighborhood query, the interpolation and the subsampling are performed on the graphics card, the GPU Memory (11 GB on an NVidia GeForce GTX 1080 Ti) was a limiting factor.

After the per-batch classification, we averaged the class probabilities for each point over the batches, and assigned the class with the highest average probability to the point.

To avoid long training times, we also tested using an already trained model as initial values for training on a new dataset (transfer learning).

4 Data

We tested the algorithm on different datasets, including a large-scale airborne laser scan and the 3D Semantic Labeling Challenge Benchmark of the ISPRS (GERKE 2014).

4.1 ALS Vorarlberg

The ALS dataset of the federal state of Vorarlberg covers about 2700 km² with a point density between 10 pts/m² and 20 pts/m². This results in about 10^{10} points, of which about $8 \cdot 10^8$ were selected for training and $6 \cdot 10^8$ for validation. The covered areas include urban, suburban, agricultural, forested, and alpine terrain. Along with the point coordinates, the amplitude of the returned signal, the number of echoes, and the echo number were provided for each point.

The reference classification in this dataset was derived using the automatic workflow in *Ter-raScan*, *TerraSolid*, and *TerraModeler* (*Terrasolid Ltd*.). In addition, classification was checked manually and corrected if needed, especially with respect to cable car lines. The classification follows the ASPRS LAS Specification (TOPOSYS 2014; ASPRS 2011).

Because of the unbalanced class frequency in the dataset, only batches containing a large variety of classes were used for training. This was achieved by calculating the standard deviation of the class histogram for each batch. A threshold on this standard deviation was set at 20 %, effectively selecting 3,781 of the original 13,255 batches for training.

4.2 ISPRS Benchmark Vaihingen

The Vaihingen dataset, provided as a benchmark for 3D semantic labelling (i.e. classification), consists of about 750,000 data points for training, spanning about 360 by 360 meters. NIE-MEYER et al. (2014) manually created reference classes based on an existing 2D classification. As in the Vorarlberg dataset, amplitude information, the number of echoes and the echo number are provided with the dataset (GERKE 2014). We note here that the relatively small number of samples makes the application of machine learning algorithms, especially neural networks, difficult. Therefore, we did not apply any filtering on the training set batches.

5 Results

On the Vorarlberg dataset, we observed a strong spatial correlation of the classification accuracy. This lead to overall accuracies ranging between 95.8 % for urban areas (average over batches in a $2.5 \times 2.5 \text{ km}^2$ tile), between 82.6 % and 86.7 % in villages surrounded by forested mountains and down to 63.6 % in the case of high alpine terrain. Additionally, we observed the spatial correlation on a smaller scale when looking at the accuracies per batch, as shown in Figure 3.

Looking at the results in more detail, Figure 4 shows a batch containing a power line, a ditch with vegetation and a bus terminal with buses and street furniture. While the buses and vegetation are classified well, especially the street furniture (yellow) does not appear in the estimation at all. Similarly, Figures 5, 6, 7 and 8 show the classification results on different areas and land cover types of the Vorarlberg dataset. On the ISPRS dataset, we achieved an overall accuracy of 80.2 %, a mid-level result compared to the competitors in the benchmark. A full confusion matrix of the classification result on this dataset is shown in Figure 9.

An interesting observation was made on the variance of the class probabilities per batch for the points: Points with a high variance (receiving different class labels from different batches) seem to correlate with erroneous classification. We conclude that we can use this variance as a measure of the quality of the classification, possibly along with probability surplus (i.e. the difference in probability of the most likely and the second most likely class).



Fig. 3: Overall accuracies per batch (200,000 points each) in "Buchboden". A strong spatial correlation can be observed in the rocky terrain in the south and southeast. Background map: Land Vorarlberg – data.vorarlberg.gv.at. Coordinates: MGI/GK West (EPSG: 31254)



Fig. 4: A batch showing a bus terminal and a power line: reference (left), differences (center) and estimation (right). The buses (grey) are classified correctly, the station building (red) partially correct. There is a patch of water (blue) present in the estimation, and the power line pylon (light blue) is mostly classified as vegetation (shades of green). Overall, these pylons were misclassified, but the power lines (where present) were classified quite well.



Fig. 5: Large factory building that is mostly misclassified as ground and low vegetation: reference (left), differences (center) and estimation (right). The maximum search radius of 15 m does not cover the building. In a strip of around 7.5 m along the border of the building, the classification is correct. Furthermore, street furniture (yellow, in this case a fence) is misclassified as vegetation, and some vegetation points along the riverbank are misclassified as ground.



Fig. 6: All-vegetation batch showing the problematic separation between low vegetation and ground: reference (left), differences (center) and estimation (right). alsNet largely overestimates the ground points, whereas the reference shows much more vegetation.

L. Winiwarter & G. Mandlburger



Fig. 7: Reference (left), differences (center) and estimation (right). On the river banks (yellow rectangle), the estimation shows much more ground points than the reference classification. The other areas in this batch (single buildings, trees) are classified well.



Fig. 8: A batch covering coniferous forest: reference (left), differences (center) and estimation (right). The few misclassified points again appear where low vegetation points are close to the ground. This is also caused by shadowing effects in the forest, where single points do not have close neighbors, leading to ambiguity.

6 Discussion and Outlook

alsNet has shown the ability of a neural network acting directly upon the point cloud to learn a representation of neighborhood information comparable to manually selected and extracted features. The benefit of end-to-end Deep Learning is that the features will automatically be optimized for the task at hand, given by the training samples. In addition, alsNet allows the computation of distribution parameters not solely on the geometry, but also on the other attributes of the point cloud.

The highest misclassification rates appear at the intersection of low vegetation and ground points, especially in areas where occlusion leads to isolated points. We expect that the use of more attributes that might be available from the sensor, namely the *echo width*, would aid in the separation of these two classes. Furthermore, attributes acquired by different sensors such as RGB values from a camera may be fused with the point cloud to provide additional information.

		Power 499 (0%)	Low Veg. 95108 (23%)	Imp. Surf. 104198 (25%)	Car 1176 (0%)	Estimated Fence/Hedge 353 (0%)	Roof 104824 (25%)	Facade 4981 (1%)	Shrub 31764 (8%)	Tree 68819 (17%)
	Power 600 (0%)	64.2% 385	0.2% 1	0.0% 0	0.0% 0	0.0% 0	22.0% 132	0.5% 3	1.8% 11	11.3% 68
Ground truth Tree Shruh Facade Roof FenceMeidee Car Imp. Surf. Low Veo.	Low Veg. 98690 (24%)	0.0% 0	79.0% 77999	9.3% 9209	0.0% 5	0.0% 4	0.6% 579	0.1% 88	7.1% 6978	3.9% 3828
	Imp. Surf. 101986 (25%)	0.0% 0	8.0% 8184	91.1% 92937	0.0% 18	0.0% 0	0.2% 171	0.0% 25	0.6% 627	0.0% 24
	Car 3708 (1%)	0.0% 0	2.3% 86	28.4% 1054	30.1% 1116	1.2% 44	0.1% 2	0.2% 8	35.4% 1313	2.3% 85
	Fence/Hedge 7422 (2%)	0.0% 0	10.7% 796	6.2% 463	0.0% 3	4.0% 296	0.5% 40	0.1% 11	53.8% 3994	24.5% 1819
	Roof 109048 (26%)	0.1% 98	1.2% 1331	0.0% 36	0.0% 1	0.0% 0	91.3% 99515	0.6% 625	1.6% 1757	5.2% 5685
	Facade 11224 (3%)	0.0% 5	8.2% 916	1.3% 149	0.2% 20	0.0% 0	19.8% 2218	34.1% 3827	16.8% 1883	19.7% 2206
	Shrub 24818 (6%)	0.0% 6	19.4% 4821	1.3% 325	0.0% 11	0.0% 9	1.6% 405	0.5% 119	39.5% 9814	37.5% 9308
	Tree 54226 (13%)	0.0% 5	1.8% 974	0.0% 25	0.0% 2	0.0% 0	3.2% 1762	0.5% 275	9.9% 5387	84.5% 45796

Fig. 9: Confusion matrix for the ISPRS Benchmark dataset. Correctly represented classes are shown on the main diagonal in green, whereas off-diagonal elements show misclassification. For example, 22 % of the power line points are misclassified as roof points (first row, sixth column). The main classes of low vegetation (24 % of the dataset), impervious surface (25 %), roof (26 %) and tree (13 %) are each classified with an accuracy of over 79 %. The highest misclassification in terms of absolute points is the estimation of shrubs as trees, followed by low vegetation classified as impervious surface.

The transfer of a trained model showed that the network is very susceptible to changes in point density and point patterns. When re-training the network, however, only a relatively small dataset is needed to achieve adequate results. The need for big training datasets could therefore be overcome by the use of simulated data (for which the reference is known exactly) and only very little real training data.

alsNet can be seen as a proof-of-concept for applying Deep Learning directly on topographic 3D point clouds by using an adapted version of PointNet++, considering characteristics of ALS data. The additional information obtained (probability, variance of probability over batches) can be further investigated. The extension of the neighborhood concept to 4D (i.e. time series of 3D point clouds) could be applied to analyses of terrain deformations and quantification thereof.

7 Acknowledgments

The neural network was trained and evaluated in part on hardware donated by NVIDIA at the Universities of Stuttgart and Heidelberg and on the Vienna Scientific Cluster VSC-3.
8 References

- ASPRS, 2011: LAS Specification. Version 1.4 R13. URL: https://www.asprs.org/wp- content/up-loads/2010/12/LAS_1_4_r13.pdf, last access 17.06.2018.
- BREIMAN, L., 2001: Random forests. Machine learning, 45(1), 5-32.
- FRÉNAY, B. & VERLEYSEN, M., 2014: Classification in the presence of label noise: a survey. IEEE transactions on neural networks and learning systems, **25**(5), 845-869.
- GERKE, M., 2014: 3D Semantic Labeling Contest. URL: http://www2.isprs.org/commissions/comm3/wg4/3d-semantic-labeling.html, last access 17.06.2018.
- GOODFELLOW, I., BENGIO, Y. & COURVILLE, A., 2016: Deep Learning. URL: http://www.deeplearningbook.org. MIT Press. ISBN: 978-0-262-03561-3.
- GRILLI, E., MENNA, F. & REMONDINO, F., 2017: A review of point clouds segmentation and classification algorithms. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(2/W3), 339-344.
- HEARST, M.A., DUMAIS, S.T., OSUNA, E., PLATT, J. & SCHOLKOPF, B., 1998: Support vector machines. IEEE Intelligent Systems and their Applications, **13**(4), 18-28.
- NIEMEYER, J., ROTTENSTEINER, F. & SÖRGEL, U., 2014: Contextual classification of Lidar data and building object detection in urban areas. ISPRS Journal of Photogrammetry and Remote Sensing, **87**, 152-165.
- OTEPKA, J., GHUFFAR, S., WALDHAUSER, C., HOCHREITER, R. & PFEIFER, N., 2013: Georeferenced point clouds: A survey of features and point cloud management. ISPRS International Journal of Geo-Information, 2(4), 1038-1065.
- QI, C.R., SU, H., MO, K. & GUIBAS, L.J., 2017a: PointNet: Deep learning on Point Sets for 3D Classification and Segmentation. IEEE Computer Vision and Pattern Recognition, 1(2), 4.
- QI, C.R., YI, L., SU, H. & GUIBAS, L.J., 2017b: PointNet++: Deep hierarchical feature learning on point sets in a metric space. Advances in Neural Information Processing Systems, 5105-5114.
- ROSENBLATT, F., 1958: The Perceptron: a probabilistic model for information storage and organization in the brain. Psychological Review, **65**(6), 386.
- TOPOSYS (WIEDENHÖFT, A. & VATSLID, S.G., 2014: Technischer Abschlussbericht LiDAR und RGB Land Vorarlberg.
- WEINMANN, M., JUTZI, B. & MALLET, C., 2013: Feature relevance assessment for the semantic interpretation of 3D point cloud data. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 5(W2).
- WEINMANN, M., JUTZI, B. & MALLET, C., 2014: Semantic 3D scene interpretation: a framework combining optimal neighborhood size selection with relevant features. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **2**(3), 181.
- YOUSEFHUSSIEN, M., KELBE, D.J., IENTILUCCI, E.J. & SALVAGGIO, C., 2018: A multi-scale fully convolutional network for semantic labeling of 3D point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, **143**, 191-204.

Mountain Permafrost Distribution Modeling – A Geomorphometry-Remote Sensing Approach for the Hohe Tauern National Park, Austria

SIMON KÖNIG¹, JOHANNES A. SCHULTZ^{1,2}, ANNA SCHOCH¹, JAN BLÖTHE¹, LOTHAR SCHROTT¹ & FRANK THONFELD^{1,3}

Abstract: Mountain permafrost is an important geomorphological component of alpine environments with many influences and high spatial heterogeneity, having a considerable hydrological and hazard-related relevance. Hence, modeling of its distribution is an important task, especially in the densely populated Alps where a large share of infrastructure is located in or below permafrost-prone areas. We present a modeling approach solely relying on remotely sensed data. Permafrost distribution is modeled for the Hohe Tauern National Park, Austria, using airborne laser altimetry, Sentinel-2 data and a published rock glacier inventory. Modeling is performed with two techniques: logistic regression and random forest regression, with 13 geomorphic and spectral parameters derived from the input data. Our results show that random forest regression is more capable of predicting the correct permafrost probability of rock glaciers, whereas logistic regression is more in accordance with a previous distribution model for the same area. Both methods produce promising results that may further be expanded, improved and applied in new areas, for which additional validation would be feasible.

1 Introduction

Permafrost can be defined as ground remaining at or below 0°C for at least two consecutive years (GRUBER & HAEBERLI 2009). In contrast to high-latitude permafrost, the existence and properties of alpine permafrost are essentially a result of high mountain topography. Alpine permafrost is therefore a highly complex phenomenon, influenced by a variety of environmental variables. These can roughly be subdivided into three categories: first, macroclimatic conditions such as latitude and large-scale wind patterns modified by complex mountain topography, determine overall energy and moisture input. Second, topoclimatic conditions, predominantly elevation, aspect and slope, locally modulate temperatures and energy input from solar radiation. Third, local ground properties, e.g. lithology, grain size and hydrological properties again alter topoclimatic conditions. Some environmental variables like snow cover properties are a feature of all three categories; and in general, influences interact with each other, resulting in a complex framework of variables. Ultimately, this leads to large ground temperature differences over short distances of up to 15°C over 1 km and results in a very heterogenous spatial permafrost distribution in mountain areas (GRUBER & HAEBERLI 2009).

¹ Geographisches Institut Universität Bonn, Meckenheimer Allee 166, D-53115 Bonn, E-Mail: simonkoenig@uni-bonn.de

² Ruhr-Universität Bochum, Geographisches Institut, Universitätsstraße 150, D-44801 Bochum

³ Zentrum für Fernerkundung der Landoberfläche, Universität Bonn, Genscherallee 3, D-53113 Bonn

Climate change is more pronounced in high mountains, inducing increased temperatures that result in permafrost degradation (RANGWALA & MILLER 2012; HARRIS et al. 2009). While mountain permafrost is associated with several hazards, the destabilization of steep rock walls is especially important in this context: the presence of permafrost stabilizes these in numerous ways, and its degradation can result in increased frequencies of mass movements of different magnitudes, as witnessed by increased rockfall activity in the Alps during the hot summer of 2003 (GRUBER & HAEBERLI 2007; GRUBER et al. 2004). Large scale slope failure related to permafrost degradation can have severe consequences: in 2002, a permafrost-containing rock wall above the *Kolka* glacier in the Greater Caucasus collapsed, fell on the glacier and induced a complex landslide with a volume of 10⁷ m³, causing more than 100 fatalities (HUGGEL et al. 2005). A similar event is reported for the Swiss Alps as well (PHILLIPS et al. 2017).

Hence, and because permafrost as a solely thermal phenomenon is not directly visible (GRUBER & HAEBERLI 2009), spatial distribution modeling remains a key challenge in the young field of mountain permafrost research. Since the 1970s, various approaches have been developed, from basic evidence-based broad rules to complex energy flux models. HAEBERLI published the very first approach in 1975, creating evidence-based altitudinal thresholds for possible and probable permafrost existence in a small area in Switzerland. The first GIS-based model was developed in the early 1990s (KELLER 1992). In recent years, spatial resolution and precision of remotely sensed data have increased, and higher computational power as well as advancements in the field of permafrost research allow for improved model capabilities (HAEBERLI et al. 2010). GRUBER developed a first global estimation based on climate and elevation data with a resolution of ~1 km in 2012. DELUIGI et al. (2017) used different machine learning techniques to model permafrost distribution based on in-situ permafrost evidence. In this context, this study aims to develop a new approach to model permafrost distribution in the Hohe Tauern National Park, Austria. Its purpose is to model the local probability of permafrost occurrence using a relatively simple, empiricalstatistical approach that solely relies on the analysis of remotely sensed data and does not require any field data for model calibration.

2 Study Area, Data and Methods

The *Hohe Tauern National Park* and its surroundings in the Austrian states of Carinthia, Salzburg and Tyrol was chosen as the study area (Fig. 1). The area accompanies the country's highest mountains and extends up to 3798 m above sea level. Permafrost distribution inside the area has been investigated by SCHROTT et al. (2012) based on extensive field surveys. The authors found that large parts of the park are likely to be underlain by permafrost, making the area most suitable for testing a different approach. The hazardous potential of mountain permafrost calls for the development of precise permafrost distribution assessments and continuous monitoring, as major infrastructure is located at high altitudes, including roads, reservoirs and touristic facilities like cable cars and skiing areas (SCHROTT et al. 2012).



Fig. 1: Study area map, including elevation, rock glaciers and their activity status, glaciers and borders of the Hohe Tauern National Park. Projection: *MGI GK Austria West*

While heavily glaciated, the region also accommodates more than 800 rock glaciers, i.e. ice-containing forms of debris that creep downslope under the force of gravity, due to deformation of ice and water-saturated material at their base as well as deformation of their active layer, i.e. the surficial permafrost layers that thaw during summer (HAEBERLI et al. 2006), following freeze-thaw cycles. Intact rock glaciers are commonly viewed as indicators for permafrost presence, whereas relict rock glaciers that do not longer contain ice, be it through climatic changes or a rock glacier extending below the local altitudinal permafrost limit, indicate permafrost absence. This evidence of permafrost abundance makes the study area suitable to test the capabilities of permafrost models (KELLERER-PIRKLBAUER et al. 2012; HAEBERLI et al. 2006).

Three data sets provided input for the model: We merged seven digital elevation models (DEMs) of the bordering districts and states to extract information on topographic parameters. Data was gathered via airborne laser scanning with a resolution of 10 m and provided by the local authorities. To acquire information on how snow cover and vegetation develop throughout the year, a total of eight Sentinel-2 scenes were acquired, four for the Western and four for the Eastern part of the study area. For both parts, two scenes represented the beginning and end of growing season, two additional embodied start and finish of snow season, respectively. The scenes were chosen to match the respective dates as closely possible while also exhibiting a cloud cover as low as possible. Sentinel-2 has a resolution of 10 m in all bands relevant for this model, which matches the DEMs. Data was provided at Level 1c processing stage, i.e. radiometrically corrected and geolocated (ESA 2015). Finally, ground truth of permafrost presence and absence were derived from an inventory of rock glaciers for the Eastern Alps (KELLERER-PIRKLBAUER et al. 2012). The inventory consists of polygons of any known rock glacier that were derived by manually analyzing DEM derivatives. No direct field data was included to create these polygons.

The complete model procedure is documented in Fig. 2. Prior to distribution modeling itself, a variety of pre-processing steps were performed. The DEMs were provided as seven individual files with different properties; they had to be reprojected and merged into one large data set. Local errors and gaps along the district borders that arose during this process were corrected manually by interpolation from neighboring cells. In order to achieve a consistent spectral database, the



Fig. 2: Flowchart of the conducted procedure. Colors represent used software: R (blue), ArcGIS/QGIS (orange), SNAP/Sen2Cor (green). Black boxes indicate multiple software. The model itself is indicated by the grey box.

Level 1c Sentinel-2 imagery was atmospherically corrected and thereby converted to Level 2a processing level. This was achieved using ESA's *Sen2Cor* algorithm (ESA 2018). Since the popular *Fmask* cloud removal tool (ZHU et al. 2015) did not score satisfying results, clouds were masked out by manually using visual interpretation. Mosaics of the Eastern/Western images were created for all four dates. Finally, the merged DEM and Sentinel-2 mosaics were clipped to the study area

extent. A total of 13 predictor variables were calculated from the inputs, serving as proxies for environmental conditions (Tab. 1). Activity status of rock glaciers was coded as a binary response variable -0 for relict and 1 for intact specimens. Including this, the final data set consisted of 14 variables and amounted to nearly 1 Billion data points. Data was extracted for all rock glaciers. This smaller data set of 600,000 observations of the 14 variables, i.e. 8,4 Million data points, served as input for the model.

Parameter	Relevance	Reference		
Elevation (raw DEM values)	Air & surface temperatures	DELUIGI et al. (2017)		
Aspect	Energy balance	ETZELMÜLLER et al. (2001)		
Slope	Energy balance, hydrology, snow distribution	Etzelmüller et al. (2001)		
Plan & profile curvature	Hydrology, snow distribution	DELUIGI et al. (2017)		
Topographic wetness index (TWI)	Hydrology	Etzelmüller et al. (2001)		
Local relief (defined here as _{Zmax} -z _{min})	Microclimate	Etzelmüller et al. (2001)		
Potential incoming solar radia- tion (PISR)	Energy balance	DELUIGI et al. (2017)		
Wind shelter index (WSI)	Snow distribution	WINSTRAL et al. (2002)		
Longitude & latitude	Climatic conditions	Arenson & Jakob (2010)		
Maximum Normalized Differ- ence Vegetation Index (NDVI) of all dates	Ground cover, active layer char- acteristics	DELUIGI et al. (2017)		
Maximum Normalized Differ- ence Snow Index (NDSI) of all dates	Snow cover characteristics & distribution	HALL et al. (2015)		

To compare the results of a classical statistical and a machine-learning regression procedure, two model approaches were incorporated. Both are commonly used in a variety of scientific questions, including geomorphological topics and especially mountain permafrost research (BRENNING & TROMBOTTO 2006; DELUIGI et al. 2017). *Logistic regression* (LR), just as linear regression, is a parametric technique. It makes use of a logistic function to assign probabilities between 0 and 1:

$$\log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \tag{1}$$

where p(X) is the probability to be estimated, $X_1, ..., X_p$ are the predictors and $\beta_0, ..., \beta_1$ are the coefficients for these predictors that are fit using the maximum-likelihood approach (JAMES et al. 2017, 132-135):

$$\ell(\beta_0, \beta_1, \dots, \beta_p) = \prod_{i: y_i = 1} p(x_i) \prod_{i': y_{i'} = 0} (1 - p(x_{i'})).$$
(2)

Random forests (RF) is a machine-learning method introduced by BREIMAN (2001), feasible for both classification and regression. The method builds on the concepts of decision trees and bagging. A decision tree applies recursive binary splitting of predictor space, resulting in a tree-like structure. Bagging relies on creating a multitude of trees, each only incorporating a random sample of the training observations to decrease model variance and reduce possible overfitting. RF extends this idea by using a different random predictor subset in each tree, resulting in decorrelated trees. The final probabilities are assigned by averaging over the whole number of trees, which in this study amounted to 500 (JAMES et al. 2017).

LR and RF models were fit and applied to the complete study area individually, and their results validated in two ways: firstly, the mean modeled probabilities of any rock glacier were computed and compared for the assumed binary values. Secondly, a comparison to SCHROTT et al. (2012) was drawn: the authors' model output is expressed as an index between 1 and 100 for all likely permafrost occurrences, corresponding to all probabilities above 0.5 for this study's results. A difference between both was drawn for each pixel, normalized by the sum of both. For the creation of final maps, values below 0.5 were disregarded to only include areas with higher probabilities.

3 Results and Discussion

In general, both models produced similar results: permafrost probabilities greater than 0.5 are only abundant in the higher regions of the study area, surrounding the glacier covered peaks of the park's highest mountain groups (Fig. 3). A closer comparison reveals only subtle differences between both methods: LR predicts a more homogenous distribution with large areas of high probability in the inner and higher areas of the mountain groups, with sharply decreasing probabilities towards their margins, which appear to be 'fringy' (Fig. 3b). The total area of likely permafrost occurrence adds up to 13.5% of the study area. The permafrost bodies modeled by RF appear less homogenous and more differentiated. Permafrost areas are rather characterized by a slow and smooth increase in probability towards the highest peaks. In terms of area, RF models a slightly larger distribution: 14.8% is covered by permafrost. The similarity of both results is illustrated by their normalized difference as well: the mean for any pixel with an assigned probability greater than 0.5 amounts to just 0.016, i.e. on average, LR models a 1.6% higher probability.

Higher differences arise concerning the first validation method: RF is more capable of correctly predicting the binary values for rock glaciers. The mean difference between assumed and predicted value of all rock glaciers is 0.143 for LR and 0.017 for RF, while both methods perform better on relict rock glaciers (Fig. 4). The better performance of RF is also supported by qualitative assessment of the RF detail map, where intact rock glaciers are almost exclusively located inside high-probability areas and relict specimen are outside (Fig. 3c), a characteristic that is less detectable in the detailed LR map (Fig. 3b). The results produced by LR exhibit less deviance from the map by SCHROTT et al. (2012), with a difference that is 0.007 lower compared to RF. This however is fairly small compared to the overall deviances of both models, which amounts to 0.374 for RF and 0.367 for LR. I.e., distinctions in normalized difference to SCHROTT et al. (2012) are relatively high, with only slight differences between LR and RF.

Despite these differences in detail, model results are in general in accordance with the previous, field-based study by SCHROTT et al. (2012): permafrost is distributed over a relevant part of the national park's higher regions, and while both LR and RF show an increase of probability with

rising altitude, results are not that simple and reflect more complex environmental variable relationships. In general, the results attained by our study are in good agreement with earlier findings



Fig. 3: Permafrost distribution maps. Logistic regression (a, b) models more homogenous permafrost bodies with higher probabilities than random forest regression (c, d). Projection: *MGI GK Austria West*

(SCHROTT et al. 2012) and demonstrate the potential of using both machine learning and probabilistic approaches in the prediction of alpine permafrost. The similar approach by DELUIGI et al. (2017) used machine learning techniques as well, the authors calibrated their models with field data, however. Yet, both approaches elucidate the abilities of machine learning for mountain permafrost distribution modeling and in addition shows that remotely sensed data is a sensible input for this task.



Fig. 4: Stacked histograms of the difference between mean modeled and the assumed binary values for any rock glacier. As the x scales depict, logistic regression (a) results in higher deviances, compared to random forest regression (b). Red lines represent mean values for both.

While its renouncement of field data is a distinct advantage of this study's approach in terms of rapid low cost, high resolution mapping of mountain permafrost distribution as well as high transferability to other study areas, it also results in the disadvantage that it lacks truly independent test data for validation. Using the values of rock glaciers that were previously used for training to estimate model errors may be insufficient, as only the training error is calculated. In addition, rock glaciers and their information on permafrost status may not necessarily be representative for whole study areas, as they represent only small parts of the region and tend to exhibit lower temperatures and hence higher permafrost occurrence probabilities than their surroundings (HAEBERLI et al. 2006). Moreover, rock glaciers are a prominent, but not always present feature of periglacial environments, making the transfer to other regions difficult. Comparison of different model outputs is feasible in general and the results by SCHROTT et al. (2012) provide a valuable source of information, but are yet a definite validation source, as their model is also based on assumptions and may exhibit errors. Finally, the chosen predictors cannot completely depict actual influences on permafrost distribution, though they represent a broad range of environmental variables on all scales of influence. Hence, before application to other regions, the model should be validated with independent data and sensitivity studies may be conducted to test which parameters present actual conditions to the highest grade.

4 Conclusion

The approach presented in this study results in high-resolution maps of mountain permafrost abundance probabilities that were created using only remotely sensed data. It is data-driven and does not rely on expert input, nor field measurements. The existence of a rock glacier inventory is an advantage in this context, as deriving these inventories from airborne LiDAR is possible, but remains an elaborate and time-consuming task (KELLERER-PIRKLBAUER et al. 2012). Consequently, future extensions of this approach could possibly incorporate automatic rock glacier detection to automatize the whole process and make it more data-driven (BRENNING 2009). A challenge in this context is the distinction between intact and relict rock glaciers that today still requires expert knowledge (KELLERER-PIRKLBAUER et al. 2012). Additionally, an inclusion of additional remote sensing products may improve the model: especially, time series analyses of snow cover and vegetation could provide better estimates of their development. The use of radar data to include more sophisticated snow cover proxies as well as thermal infrared data for derivation of land surface temperatures are promising ideas in this context as well.

The high accuracies of random forest regression illustrate the benefits of using machine learning methods for environmental modeling. Hence, it appears natural to also test other techniques in this context, such as support vector machines. Notably, the use of deep learning algorithms which have often been used for a multitude of questions and have proven to be highly accurate, may especially benefit this approach (ZHU et al. 2017). All suggested techniques could also be used in continuous spatiotemporal monitoring of permafrost conditions, e.g. in precaution to the hazards described in the introduction. Combined with a continuously improved understanding of mountain permafrost in its whole, modeling of its distribution can represent an actual benefit for local communities.

5 Acknowledgements

We thank Jan-Christoph Otto for providing us with the data as published in SCHROTT et al. (2012). The rock glacier inventory (KELLERER-PIRKLBAUER et al. 2012; LIEB et al. 2012) is freely available at *https://doi.pangaea.de/10.1594/PANGAEA.869805* under Creative Commons Attribution 3.0. Outlines of glaciers in Figs. 1 & 3 (FISCHER et al. 2015) are available under Creative Commons Attribution 3.0 at *https://doi.pangaea.de/10.1594/PANGAEA.844985*. The DEMs as well as the outline of the national park borders can be found at *http://data.gv.at* under Creative Commons Attribution 3.0.

6 References

ARENSON, L. & JAKOB, M., 2010: A New GIS based Mountain Permafrost Distribution Model. Proceedings of the 6th Canadian Permafrost Conference, 452-458.

BREIMAN, L., 2001: Random Forests. Machine Learning, 45, 5-32.

BRENNING, A., 2009: Benchmarking classifiers to optimally integrate terrain analysis and multispectral remote sensing in automatic rock glacier detection. Remote Sensing of Environment, 113, 239-247.

- BRENNING, A. & TROMBOTTO, D., 2006: Logistic regression modeling of rock glacier and glacier distribution: Topographic and climatic controls in the semi-arid Andes. Geomorphology, 81, 141-154.
- DELUIGI, N., LAMBIEL, C. & KANEVSKI, M., 2017: Data-driven mapping of the potential mountain permafrost distribution. The Science of the total environment, **590-591**, 370-380.
- ETZELMÜLLER, B., ØDEGÅRD, R. S., BERTHLING, I. & SOLLID, J. L., 2001: Terrain parameters and remote sensing data in the analysis of permafrost distribution and periglacial processes: principles and examples from southern Norway. Permafrost and Periglacial Processes, **12**, 79-92.
- EUROPEAN SPACE AGENCY (ESA), 2015: Sentinel-2 User Handbook. https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook.
- EUROPEAN SPACE AGENCY (ESA), 2018: Sen2Cor Configuration and User Manual. http://step.esa.int/thirdparties/sen2cor/2.5.5/docs/S2-PDGS-MPC-L2A-SUM-V2.5.5_V2.pdf.
- FISCHER, A., SEISER, B., STOCKER-WALDHUBER, M. & ABERMANN, J., 2015: The Austrian Glacier Inventory GI 3, 2006, in ArcGIS (shapefile) format. https://doi.pangaea.de/10.1594/PAN-GAEA.844985.
- GRUBER, S., 2012: Derivation and analysis of a high-resolution estimate of global permafrost zonation. The Cryosphere, **6**, 221-233.
- GRUBER, S. & HAEBERLI, W., 2007: Permafrost in steep bedrock slopes and its temperature-related destabilization following climate change. Journal of Geophysical Research, **112**, 301.
- GRUBER, S. & HAEBERLI, W., 2009: Mountain Permafrost. Permafrost Soils, R. Margesin (ed.), Springer Verlag, Berlin/Heidelberg, 33-44.
- GRUBER, S., HOELZLE, M. & HAEBERLI, W., 2004: Permafrost thaw and destabilization of Alpine rock walls in the hot summer of 2003. Geophysical Research Letters, **31**, L13504.
- HAEBERLI, W., 1975: Untersuchungen zur Verbreitung von Permafrost zwischen Flüelapass und Piz Grialetsch (Graubünden). Mitteilung der Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie. Zürich.
- HAEBERLI, W., HALLET, B., ARENSON, L., ELCONIN, R., HUMLUM, O., KÄÄB, A., KAUFMANN, V., LADANYI, B., MATSUOKA, N., SPRINGMAN, S. & MÜHLL, D. V., 2006: Permafrost creep and rock glacier dynamics. Permafrost and Periglacial Processes, **17**, 189-214.
- HAEBERLI, W., NOETZLI, J., ARENSON, L., DELALOYE, R., GÄRTNER-ROER, I., GRUBER, S., ISAKSEN, K., KNEISEL, C., KRAUTBLATTER, M. & PHILLIPS, M., 2010: Mountain permafrost: development and challenges of a young research field. Journal of Glaciology, 56, 1043-1058.
- HALL, D. K., FREI, A. & DÉRY, S. J., 2015: Remote sensing of snow extent. Remote Sensing of the Cryosphere, M. Tedesco (ed.), Chichester, UK, 31-47.
- HARRIS, C., ARENSON, L. U., CHRISTIANSEN, H. H., ETZELMÜLLER, B., FRAUENFELDER, R., GRUBER, S., HAEBERLI, W., HAUCK, C., HÖLZLE, M., HUMLUM, O., ISAKSEN, K., KÄÄB, A., KERN-LÜTSCHG, M. A., LEHNING, M., MATSUOKA, N., MURTON, J. B., NÖTZLI, J., PHILLIPS, M., ROSS, N., SEPPÄLÄ, M., SPRINGMAN, S. M. & VONDER MÜHLL, D., 2009: Permafrost and climate in Europe: Monitoring and modelling thermal, geomorphological and geotechnical responses. Earth-Science Reviews, **92**, 117-171.

- HUGGEL, C., ZGRAGGEN-OSWALD, S., HAEBERLI, W., KÄÄB, A., POLKVOJ, A., GALUSHKIN, I. & EVANS, S. G., 2005: The 2002 rock/ice avalanche at Kolka/Karmadon, Russian Caucasus: assessment of extraordinary avalanche formation and mobility, and application of QuickBird satellite imagery. Natural Hazards and Earth System Science, 5, 173-187.
- JAMES, G., WITTEN, D., HASTIE, T. & TIBSHIRANI, R., 2017: An introduction to statistical learning. With applications in R. Springer texts in statistics. Springer Science + Business Media, New York, Heidelberg, Dordrecht, London.
- KELLER, F., 1992: Automated mapping of mountain permafrost using the program PERMAKART within the geographical information system ARC/INFO. Permafrost and Periglacial Processes, 3, 133-138.
- KELLERER-PIRKLBAUER, A., LIEB, G. K. & KLEINFERCHNER, H., 2012: A new Rock Glacier Inventory of the Eastern European Alps. Austrian Journal of Earth Sciences, **105**, 78-93.
- LIEB, G. K., KELLERER-PIRKLBAUER, A. & KLEINFERCHNER, H., 2012: Second rock glacier inventory (RGI2) of Central and Eastern Austria, link to Shapefile, supplement to: Kellerer-Pirklbauer, A., Lieb, G. K. & Kleinferchner, H., 2012: A new rock glacier inventory of the eastern European Alps. Austrian Journal of Earth Sciences, 105(2), 78-93. https://doi.pangaea.de/10.1594/PANGAEA.869805.
- PHILLIPS, M., WOLTER, A., LÜTHI, R., AMANN, F., KENNER, R. & BÜHLER, Y., 2017: Rock slope failure in a recently deglaciated permafrost rock wall at Piz Kesch (Eastern Swiss Alps), February 2014. Earth Surface Processes and Landforms, 42, 426-438.
- RANGWALA, I. & MILLER, J. R., 2012: Climate change in mountains: a review of elevation-dependent warming and its possible causes. Climatic Change, **114**, 527-547.
- SCHROTT, L., OTTO, J.-C. & KELLER, F., 2012: Modelling Alpine Permafrost Distribution in the Hohe Tauern Region, Austria. Austrian Journal of Earth Sciences, **105**, 169-183.
- WINSTRAL, A., ELDER, K. & DAVIS, R. E., 2002: Spatial snow modeling of wind-redistributed snow using terrain-based parameters. Journal of Hydrometeorology, **3**, 524-538.
- ZHU, X. X., TUIA, D., MOU, L., XIA, G.-S., ZHANG, L., XU, F. & FRAUNDORFER, F., 2017: Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. IEEE Geoscience and Remote Sensing Magazine, 5, 8-36.
- ZHU, Z., WANG, S. & WOODCOCK, C. E., 2015: Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. Remote Sensing of Environment, **159**, 269-277.

Development of large scale 3D point cloud processing modules – minimizing the number of input parameters through statistical modeling and optimization

MAXIMILIAN SPERLICH¹ & HOLGER WEINACKER¹

Abstract: The main objective of this thesis was to develop a pipeline of 3D point cloud processing modules, facilitating the extraction of training data for supervised classification tasks. The modular design enables the application of each module as standalone tools for filtering, segmentation or information retrieval.

Core component of most modules is a ground filtering algorithm, devised as a combination and further development of two well-established methods by WANG et al. (2015) and ZENG et al. (2016). The ground filter is thoroughly tested for robustness through a range of freely available benchmark datasets. Its performance is further compared to results produced by multiple methods tested on equal data.

In order to reduce the number of necessary parameters, an optimized general parameter set is produced through a regression model. The resulting default parameters prove to be applicable to a wide range of scenarios.

For further ground filtering improvements, a method for the regional segmentation of large scale outdoor point clouds is proposed, resulting in spatially coherent areas of similar terrain slope. Further characteristics, such as terrain roughness, are calculated for each segment. Optimized parameters are then calculated for the respective slope and roughness classes, thus enabling the implementation of automatic parameter adjustments into the ground filter.

As a measure for the viability of the accumulative results from all proposed methods, a supervised random forest classifier is trained from data, which has been preprocessed by each module. Hereby, the facilitation of manual training data extraction, as a result of the proposed methodology pipeline, is evaluated. The performance measures of the subsequent classification, serve as evidence for the effectiveness of the training data, as a cumulative representation of all presented modules.

1 Problem Statement

The scientific literature provides a wide array of options for point cloud classification. The effectiveness of supervised classification of 3D point clouds, based on 2D and 3D shape features derived from the local point neighborhood has been demonstrated on multiple occasions (BLOMLEY et al. 2016; DEMANTKÉ et al. 2012; JUTZI 2015; WEINMANN 2016). In many cases, the main limitation is a design focused on a specific task, based either on the applied sensor type or on the captured scene. However, improvements in the area of sensors enable the capture of large areas (> 10 ha), including varying landforms and land use types with point densities > 30 points/m². The technology for processing such data is becoming more affordable. Therefore, the need for a more generalized approach to large scale point cloud classification arises.

¹ Albert-Ludwigs-Universität Freiburg, Professur für Fernerkundung und Landschaftsinformationssysteme, Tennenbacherstr. 4, D-79106 Freiburg, E-Mail: maximilian.sperlich@felis.uni-freiburg.de

Furthermore, supervised classification requires the generation of training data. This is often a long and tedious process and great deal of time has to be spent, carefully picking the points by hand, in order to meet the following requirements for quality training:

- being as plentiful as possible while
- remaining randomly distributed,
- representing the actual class proportions and
- minimizing spatial autocorrelation.

The more extensive the data is, the more time has to be set aside for the generation of training data. Robust, unsupervised methods for segmenting the 3D points prior to training data selection facilitate the work flow. Here, methods for filtering the ground points (GP) are of utmost significance and numerous studies, especially progressive TIN densification (PTD) approaches, have produced valuable results and are under constant development (NIE et al. 2017; ZENG et al. 2016).

Further publications also address the problem of retrieving information from the remaining nonground points (NGP), in order to extract either building or vegetation point clusters (TÓVÁRI 2006; WANG & TSENG 2014). However, RAMIYA et al. (2016) point out, that the issue of automatically separating tree from building clusters, when both cluster classes are to be returned, is scarcely covered in literature. Again, the methods found in literature are often bound to specific landforms or land uses and the results also rely on the fine tuning of multiple parameters.

Hence, any supervised classification task of large scale 3D points would benefit from a robust ground filtering algorithm, unaffected by varying data attributes like terrain slope or roughness. Furthermore, additional subsequent methods for processing the remaining NGP are also beneficial. Most unsupervised methods studied in the course of this thesis, rely on the fine tuning of multiple input parameters.

The presented study aims at solving all afore mentioned problems.

2 Methods

Figure 1 displays the proposed methodology pipeline passing through every developed module, with the aim of facilitating training data extraction and generating additional features for the improvement of supervised 3D point cloud classification.

2.1 Low outlier filter

Outlying points below the true ground are removed, as they could otherwise cause problems with the ground filter. This is achieved through iteratively executing a simplified version of the ground filter, outlined in section 2.5.1, while inverting the z-axis after every step.

2.2 Statistical model and parameter optimization

A regression model, as part of the ground filter, is trained in advance with a wide range of training data. Based on the features resulting from the regional segmentation module, parameters for the ground filter and the cluster segmentation within the semi-supervised classification / segmentation module are optimized.



Fig. 1: Overview of the classification pipeline with the various processing modules grouped into preliminary processing modules and the main classification algorithm. Raw 3D point data is the input and semantically classified 3D point data the output of the pipeline.

2.3 2D/3D point feature calculation

Based on circular (2D), spherical (3D) and gridded (3D) local point neighborhood of each point, 2D and 3D shape features are derived, mainly calculated from respective eigenvalues. While the resulting features mainly serve as input for the supervised random forest classification module, features like linearity, planarity or scattering are independently informative, justifying a separate module for the feature calculation.

2.4 Regional segmentation

The filtered point cloud is segmented into regions of coherent terrain slope for which additional features, like terrain roughness or the proportion of probable vegetation and artificial structures, are calculated.

2.5 Semi-supervised classification / segmentation

With a preceding regional segmentation, each previously segmented region is processed individually. With the parameters adjusted through the regression models, in a first step, the GP are separated from the NGP. In a second step, the NGP are spatially clustered and every cluster is searched for planarly organized points as an indicator of unnatural objects. The resulting categories are: ground, probable building cluster, probable vegetation cluster, and unclassified NGP.

2.5.1 Ground filter

The ground filter consists of two main algorithms. In the TIN densification, the lowest point in each grid cell serves as the initial TIN vertices. Based on the triangle characteristics, the regression model predicts the tolerated height difference for each triangle. Points with a vertical TIN distance

below the calculated threshold are potential GP in the next densification iteration. Once a target grid size is reached, the remaining GP are extracted through an entropy based ground filter. The switch to an entropy based ground filter substantially increases processing speeds with point clouds of high point density. Various optimized ground filter parameters are optionally set, based on preceding slope region segmentation.

2.5.2 Cluster segmentation

This module processes only NGP. Spatial clusters are grouped through Density-based spatial clustering of applications with noise (DBSCAN) and filtered by size. Planar segments are detected and serve as a measure for building / vegetation probability. Clusters with a probability between 20% and 70% are split into segment and non-segment points. Points within close proximity of the segments are added to the resulting building cluster. The remaining points result to a vegetation cluster. The planar segments are further merged into individual building clusters.

2.6 Supervised random forest classification

The results of the preceding modules simplify the training data selection for a supervised classification. The supervised classification method used here is a random forest approach based on 2D/3D shape features of the local point neighborhood. In addition to the mandatory shape features, the resulting point attributes of preprocessing steps improve the classification results.

3 Results

For all 10 terrestrial laser scanning (TLS) reference datasets from BRODU & LAGUE (2012), none of which included outliers, the outlier filter adequately did not flag any points as outliers. While a high ratio of the present outliers is detected in many datasets, in some cases the filter misses all outliers. In others, often cases without any outliers present, the filter falsely classifies many points as outliers. In only two out of the 16 airborne laser scanning (ALS) benchmark datasets more than 30% of the outliers were missed. In six cases all outliers were detected. The main weaknesses of the outlier filter are false positives in areas with jump edges or regions with few true ground points. Applying the outlier filter to aerial based photogrammetric point cloud data demonstrates an advantage of this method over methods such as statistical outlier removal. Aside from isolated low outliers, also outlier clusters are correctly flagged, resulting in unprecedented ground filtering results.

Based on visual examination, the results from the regional segmentation module capture the varying terrain characteristics very well. Problems only arise in cases where erroneous ground filtering results, based on the default parameter set, precede. Manual adjustment of said parameters significantly improves the regional segmentation outcome.

The main reference for the ground filtering results is the benchmark data from SITHOLE & VOSSELMAN (2003), on which a wide range of methods have been tested. With a mean total error of 4.90 and a mean kappa-coefficient of 84.38 the proposed ground filtering algorithm is only outperformed by two out of the 14 compared methods, proposed by CHEN et al. (2013, 2016), displaying total errors of 4.11 and 3.03 and kappa-coefficients of 86.27 and 89.44. The proposed method without preceding regional segmentation is thus already able to produce results similar to the leading methods, and additional regional segmentation cannot improve the results for the

benchmark datasets. Tested on data including regions of very steep terrain, the ground filter performance is much better, if the steep regions are segmented and processed individually with automatically adjusted parameters. None of the compared methods featured data of similarly steep terrain. It is also vital to notice, that the benchmark data was recorded in 2003 and does not represent the point density attainable with modern sensors. It is conceivable that, given new benchmark data, the proposed method displays a significant advantage in processing speed.

A visual representation of the results of a supervised classification is displayed in figure 2. Aided by all previously generated results, manual training data selection is accomplished within under an hour, extensively covering the entire area of the data. The mean performance measures result to an Accuracy of 0.98, a Precision of 0.75, a Recall of 0.64 and a F1 Score of 0.67. The relatively low values of Precision, Recall and F1 Score can here be attributed to the classifiers inability to capture the cable running from the top left of the image to the bottom right corner. The F1 Scores for the individual classes aside from the 'cable' class all lie above 0.84 with F1 Scores for 'ground' and 'high vegetation' at 0.98.

Aside from being valuable individual tools, the pipeline of processing modules as a whole can greatly facilitate training data selection, whenever a point cloud is to be manually subset into individual classes for supervised classification.



Fig. 2: Supervised classification: unclassified – red, ground – brown, low vegetation – light green, high vegetation – dark green, building – blue, vehicle – yellow. The additionally trained class cable is not visible in the depicted image.

4 Bibliography

- BLOMLEY, R., JUTZI, B. & WEINMANN, M., 2016: Classification of airborne laser scanning data using geometric multi-scale features and different neighbourhood types The ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, III-3, 69-176.
- BRODU, N. & LAGUE, D., 2012: 3d terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. ISPRS Journal of Photogrammetry and Remote Sensing, **68**(1), 121-134.
- CHEN, C., LI, Y., LI, W. & DAI, H., 2013: A multiresolution hierarchical classification algorithm for filtering airborne lidar data. ISPRS Journal of Photogrammetry and Remote Sensing, **82**, 1-9.

- CHEN, C., LI, Y., ZHAO, N., GUO, J. & LIU, G. (2016). A fast and robust interpolation filter for airborne lidar point clouds, PLoS ONE, **12**(5), 1-20.
- DEMANTKÉ, J., MALLET, C., DAVID, N. & VALLET, B., 2012: Dimensionality based scale selection in 3d lidar point clouds. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **38**(5), 97-102.
- JUTZI, B., 2015: Methoden zur automatischen Szenencharakterisierung basierend auf aktiven optischen Sensoren für die Photogrammetrie und Fernerkundung. Habilitation, Karlsruher Institut für Technologie. https://hal.archives-ouvertes.fr/tel-01513489/document.
- NIE, S., WANG, C., DONG, P., XI, X., LUO, S. & QIN, H., 2017: A revised progressive tin densification for filtering airborne lidar data, Measurement: Journal of the International Measurement Confederation, 104, 70-77.
- RAMIYA, A. M., NIDAMANURI, R. R. & KRISHNAN, R., 2016: Segmentation based building detection approach from lidar point cloud. The Egyptian Journal of Remote Sensing and Space Sciences, 1-7.
- ROTTENSTEINER, F., SOHN, G., GERKE, M. & WEGNER, J. D., 2013: ISPRS test project on urban classification and 3d building reconstruction. ISPRS - Commission III – Photogrammetric Computer Vision and Image Analysis Working Group III / 4 - 3D Scene Analysis, 1-16.
- SITHOLE, G. & VOSSELMAN, G., 2003: Report: ISPRS comparison of filters. ISPRS commission III, working group 3.
- Tóvári, D.; 2006: Segmentation Based Classification of Airborne Laser Scanner Data. Doctoral dissertation, PhD Thesis, Universität Karlsruhe. https://d-nb.info/1003662897/34.
- WANG, M. & TSENG, Y.-H. (2014). Extraction of surface features from lidar point clouds using incremental segmentation strategy. ISPRS Journal of Photogrammetry and Remote Sensing, 19(1), 49-62.
- WANG, H., WANG, S., CHEN, Q., JIN, W. & SUN, M., 2015: An improved filter of progressive tin densification for lidar point cloud data. Wuhan University Journal of Natural Sciences, 20(4), 362-368.
- WEINMANN, M., 2016: Reconstruction and analysis of 3D scenes: From irregularly distributed 3D points to object classes. Springer International Publishing, Cham. ISBN: 978-3-319-29246-5.
- ZENG, Z., WAN, J. & LIU, H., 2016: An entropy-based filtering approach for airborne laser scanning data, Infrared Physics & Technology, 75, 87-92.

Klassifikation hochaufgelöster LiDAR- und MVS-Punktwolken zu Monitoringzwecken

MICHAEL KÖLLE¹, DOMINIK LAUPHEIMER¹ & NORBERT HAALA¹

Zusammenfassung: Die erreichbare Genauigkeit bei der Erfassung raumbezogener Daten durch Methoden der Photogrammetrie sowie des Airborne Laserscannings (ALS) von Unmanned-Aerial-Vehicles (UAVs) ermöglicht mittlerweile auch Anwendungen wie das Monitoring von Bauwerksdeformationen oder Bodensenkungen. Unerlässliche Voraussetzung hierfür ist die Identifizierung und Elimination zeitlich veränderlicher Bereiche wie Vegetation aus den erfassten Punktwolken, um die Bestimmung der Deformationen auf sinnvolle Bereiche zu fokussieren. Für eine dementsprechend erforderliche Klassifikation werden sowohl geeignete geometrische Merkmale aus einer LiDAR- bzw. Multi-View-Stereo (MVS)-Punktwolke als auch radiometrische Merkmale aus den zugehörigen Bilddaten abgeleitet. Diese werden anschließend neben dem klassischen Random-Forest-Klassifikator (RF) auch vergleichend einem auf Merkmalen basierenden Multi-Branch 1D-Convolutional Neural Network (CNN) dargeboten.

1 Einleitung

Sowohl Airborne Laserscanning (ALS)- als auch Multi-View-Stereo (MVS)-Daten, die von Unmanned-Aerial-Vehicle (UAV)-Plattformen aus erfasst werden, haben mittlerweile ein Genauigkeitspotential nahe dem Millimeterbereich erreicht. Dies macht Anwendungen wie die Erfassung von Deformationen von Bauwerken oder geologischer Setzungen des Geländes möglich, und zwar mit deutlich größerer Flächenleistung als die klassische Ingenieurgeodäsie. Dies wird am Beispiel des mit der Bundesanstalt für Gewässerkunde (BfG) durchgeführten Projekts Hessigheim (CRA-MER et al. 2018) demonstriert, wobei dieses Gebiet nahezu zeitgleich mit einem photogrammetrischen Kamerasystem und einem LiDAR-System beflogen worden ist. Damit kann auf zwei Punktwolken unterschiedlicher Charakteristika zurückgegriffen werden. Eine weitere Besonderheit dieser Daten liegt in der extrem hohen räumlichen Dichte der aufgenommenen Punkte im LiDAR-Fall (ca. 800 Punkte/m²) bzw. der mit *Agisoft PhotoScan* (AGISOFT 2018) prozessierten Punkte im MVS-Fall (ca. 8000 Punkte/m²).

Im Hinblick auf potentielle Höhenänderungen aufgrund von möglichen Setzungen des Geländes soll zunächst in den beiden Datensätzen der ersten Epoche automatisch zwischen festen, nichtvertikalen Flächen, d.h. *Geeignet für Monitoring*, und dynamischen Flächen, d.h. *Ungeeignet für Monitoring*, unterschieden werden. Neben einer binären Klassifikation erfolgt zusätzlich eine solche, die diese beiden Klassen im Sinne einer Mehrklassenklassifikation jeweils weiter unterteilt. Damit ist eine universell anwendbare semantische Segmentierung gegeben, welche im Folgenden im Fokus steht. Die Klassifikation erfolgt mittels des klassischen Random-Forest-Klassifikators

¹ Universität Stuttgart, Institut für Photogrammetrie, Geschwister-Scholl-Str. 24D,

D-70174 Stuttgart, E-Mail: [michael.koelle, dominik.laupheimer, norbert.haala]@ifp.uni-stuttgart.de

(RF) unter Verwendung geometrischer sowie radiometrischer Merkmale, wobei letztere im Li-DAR-Fall speziell abgeleitet werden. Da sich Convolutional Neural Networks (CNNs) immer mehr als erfolgreiches Werkzeug für die Klassifikation von Punktwolken erweisen (z.B. ZHAO et al. 2018), wird auch eine Klassifikation mittels des Ansatzes eines Modellwissen in Form von Merkmalen nutzenden Multi-Branch 1D-CNN nach GEORGE et al. (2018) angewandt.

2 Semantische Interpretation der Punktwolken

Der im Rahmen dieser Arbeit vorgestellte Ansatz zur semantischen Interpretation der Punktwolken gliedert sich in drei Schritte. Zunächst werden entsprechende Referenzdaten für das Interessensgebiet erzeugt, indem auf Grundlage einer geeigneten Klassenfestlegung ein manuelles Labeln der Daten erfolgt (Kap. 2.1). Nachfolgend gilt es, die für die Klassifikation notwendigen geometrischen und radiometrischen Merkmale für die einzelnen Punktprimitive zu extrahieren (Kap. 2.2). Darauf aufbauend können die einzelnen Datensätze mittels RF und Multi-Branch 1D-CNN klassifiziert werden (Kap. 2.3).

2.1 Klassendefinition und Ableitung der Referenzdaten

Für das Training und die Evaluierung eines Klassifikators müssen jeweils gelabelte, in Trainings-, Validierungs- und Testdaten untergliederte Referenzdaten zur Verfügung stehen. Diese werden für die geplante Anwendung in Anlehnung an NIEMEYER et al. (2014) jeweils detailliert in 11 Klassen unterteilt, welche Abb. 1 entnommen werden können. Deren Vertreter sollen anschließend durch einen Klassifikator automatisch identifiziert werden.



Abb. 1: Ausschnitt der gelabelten LiDAR-Punktwolke (0=*Stromleitung*, 1=*Niedrige Vegetation*, 2=*Versiegelte Fläche*, 3=*Fahrzeug*, 4=*Sonstige Ausstattung*, 5=*Dach*, 6=*Fassade*, 7=*Strauch/Hecke*, 8=*Baum*, 9=*Offener Boden/Geröll*, 10=*Vertikale Fläche*).

Für den verfolgten Anwendungszweck des Monitorings bezüglich Senkungen des Geländes ist bereits eine binäre Klassifikation in *Geeignet für Monitoring* und *Ungeeignet für Monitoring* ausreichend. In der Klasse *Geeignet für Monitoring* können lediglich die Klassen *Versiegelte Fläche*

und *Dach* als fest annehmbare, nicht-vertikale Flächen vereinigt werden und in das angestrebte Deformationsmodell eingehen.

2.2 Merkmalsextraktion

Die zur beabsichtigten Klassifikation der Punktprimitive zu extrahierenden Merkmale müssen im Falle von Punktwolken in der Regel im Zusammenhang mit umgebenden Primitiven bestimmt werden. Als geeignete Nachbarschaftsdefinitionen werden sowohl eine Kugel- als auch eine Zylindernachbarschaft im Sinne einer Multiskalenanalyse, jeweils mit Radien von 1, 2, 3 und 5 m angesetzt (WEINMANN et al. 2018), innerhalb derer die einzelnen Merkmale berechnet werden. Um hierbei den Rechenaufwand für das Auffinden der Nachbarpunkte im Hinblick auf die vorliegenden dichten Punktwolken zu minimieren, erfolgt die Nachbarsuche mit zunehmendem Suchradius in einer stärker ausgedünnten Punktwolke (HACKEL et al. 2016). Bevor nun im Folgenden auf die verwendeten Merkmale eingegangen wird, sei auf Abb. 2 verwiesen. Hierin erfolgt eine Visualisierung der anhand des RF-Klassifikators ausgewerteten Relevanz der einzelnen Merkmale für dessen Entscheidungen.

Die definierten Nachbarschaften werden insbesondere zur Ableitung geometrischer Merkmale benötigt. Hierzu zählen eigenwertbasierte Merkmale (BECKER et al. 2017), welche die charakteristische Punktverteilung innerhalb der Nachbarschaft beschreiben (jeweils die ersten 8 Säulen einer Nachbarschaft sowie Q_{ES_2} in Abb. 2). Diese werden durch weitere, ebenfalls die Punktverteilung berücksichtigenden Merkmale, jedoch nicht-eigenwertbasierter Natur, wie etwa der Punktdichte (jeweils Säule 9 in Abb. 2) und dem Verhältnis der Punktanzahl zwischen Kugel- und Zylindernachbarschaft gleicher Radien (Q_N), ergänzt. Als extrem aussagekräftig haben sich zusätzlich höhenbasierte Merkmale herausgestellt, wozu die maximale Höhendifferenz und die Varianz der Höhen (CHEHATA et al. 2009) in einer entsprechenden Nachbarschaft zählen (jeweils Säulen 10 und 11 in Abb. 2). Daneben trägt insbesondere das Merkmal der Höhe Δz_{rel} eines Punktes über Grund zu einer eindeutigen Diskrimination der Punktprimitive bei. Weitere geometrische Merkmale können durch die Bestimmung einer robust geschätzten lokalen Ebene in jedem einzelnen Punkt abgeleitet werden (jeweils die letzten 5 Säulen in Abb. 2). Hierzu zählt der Neigungswinkel der Ebene sowie die Varianz desselben in einer gegebenen Nachbarschaft. Ferner wird neben dem Maß der Vertikalität der geschätzten Ebene zusätzlich der Abstand D des jeweiligen Berechnungspunktes von der ausgleichenden Ebene bestimmt. Als besonders effektives Merkmal wird auf Grundlage dieser Distanzen eine robust bestimmte Standardabweichung (Median of Absolute Differences) nach

$$\sigma_{MAD} = \mathrm{med}[\boldsymbol{D}_N - \mathrm{med}(\boldsymbol{D}_N)]$$

in der Umgebung berechnet, welche zur Detektion von Vegetationsflächen dient.

Um neben geometrischen auch radiometrische Merkmale der Punktprimitive, und zwar nicht nur für die MVS-Punktwolke, sondern auch für den LiDAR-Datensatz, zu gewinnen, wurden zeitgleich mit der LiDAR-Punktwolke Schrägluftbilder aufgezeichnet. Im Sinne einer integrierten Prozessierung wird die aufgezeichnete LiDAR-Punktwolke mittels einer aus diesen Aufnahmen erzeugten photogrammetrischen Punktwolke durch Interpolation aus dieser koloriert. Damit können auch für die LiDAR-Punktwolke Farbmerkmale, wie etwa die Farbkomponenten des HSV- Farbraums und deren Mittelung (Säulen 1-3 bzw. 4-27 von \mathbf{x}_{Farbe} in Abb. 2) in den verwendeten Nachbarschaften (BECKER et al. 2017), als äußerst zweckdienliche Merkmale genutzt werden. Zudem wird aus der aus Nadiraufnahmen prozessierten MVS-Punktwolke ein Orthophoto generiert, welches mittels *Segnet* (BADRINARAYANAN et al. 2017) segmentiert wird. Die entsprechenden Label lassen sich dann als weiteres Merkmal orthogonal auf beide Punktwolken projizieren (letzte Säule von \mathbf{x}_{Farbe} in Abb. 2). Im Falle der LiDAR-Daten kann zusätzlich neben der materialabhängigen Reflektanz des ausgesandten Laserpulses auch das Pulsverhältnis im Rahmen der Merkmalsberechnung verwendet werden (\mathbf{x}_{LiDAR}). Abschließend werden im LiDAR-Fall alle 167 Merkmale bzw. im MVS-Fall 165 Merkmale in einem normierten Merkmalsvektor zusammengefasst.



Abb. 2: Relevanz der Merkmale ausgewertet anhand des RF-Klassifikators für den LiDAR-Datensatz bezüglich des Mehrklassenfalls. $\mathbf{x}_{K1} - \mathbf{x}_{K5}$ und $\mathbf{x}_{Z1} - \mathbf{x}_{Z5}$ bezeichnen die für die Kugel- bzw. Zylindernachbarschaften evaluierten geometrischen Merkmale entsprechend verwendeter ansteigender Radien. Die übrigen Merkmalsbezeichnungen sowie entsprechende Erläuterungen sind Kap. 2.2 zu entnehmen.

Für die praktische Merkmalsberechnung wird auf die Software *OPALS* (PFEIFER et al. 2014) zurückgegriffen. Hauptgrund hierfür ist, dass *OPALS* eine effiziente Datenstruktur nutzt und auch nach Möglichkeit sämtliche Vorgänge parallelisiert werden, was eine hohe Prozessierungsgeschwindigkeit zur Folge hat. Daher ist *OPALS* insbesondere für die Verarbeitung großer Datenmengen, wie sie auch im Rahmen dieser Arbeit vorliegen, prädestiniert. Dennoch hat sich die Prozessierung der originalen hochaufgelösten Punktwolken auf einem Rechner mittlerer Ausstattung als extrem zeitintensiv herausgestellt. Daher wurde untersucht, inwieweit sich die Punktdichte auf die Klassifikationsgüte eines RF-Klassifikators auswirkt, indem für unterschiedlich stark unterabgetastete Punktwolken jeweils die Merkmalsberechnung und eine entsprechende Klassifikation durchgeführt wurden. Als Gütemaß wird hierfür auf den Kappa-Koeffizienten κ zurückgegriffen, welcher in Abhängigkeit von der Punktdichte in Abb. 3 visualisiert ist. Hieraus ist ersichtlich, dass



Abb. 3: Einfluss der Punktdichte auf die Klassifikation ausgewertet für die LiDAR-Punktwolke mittels des RF-Klassifikators anhand des Kappa-Koeffizienten *κ*.

mit abnehmendem Punktabstand bzw. mit zunehmender Punktdichte die Klassifikationsgenauigkeit nahezu linear zunimmt. Bezogen auf die Rechenzeit der Merkmale ergibt sich bei Erhöhung der Punktdichte allerdings ein exponentieller Anstieg. Daher wird als Kompromiss zwischen Rechendauer einerseits und Klassifikationsgenauigkeit andererseits eine Unterabtastung auf 30 cm Punktabstand gewählt. Weiterhin zeigt Abb. 3 auf, dass die Rechendauer des RF-Klassifikators im Vergleich zur Dauer der Merkmalsberechnung zwar nahezu vernachlässigbar ist, jedoch auch diese mit zunehmender Punktdichte ansteigt.

2.3 Klassifikation mittels Random Forest und Multi-Branch 1D-CNN

Die Klassifikation erfolgt sowohl für den LiDAR- als auch für den MVS-Datensatz zunächst durch den klassischen RF-Klassifikator (BREIMAN 2001). Vergleichend wird für den LiDAR-Datensatz zusätzlich der unkonventionelle Ansatz eines Multi-Branch 1D-CNN angewandt, welcher analog zum RF Modellwissen in Form von Merkmalen berücksichtigt. Die Architektur desselben wurde GEORGE et al. (2018) nachempfunden, jedoch für den Zweck dieser Arbeit entsprechend angepasst und ist in Abb. 4 visualisiert. Ein wesentlicher Vorteil eines solchen aus Faltungslayern aufgebauten Netzes gegenüber einer aus Fully-Connected-Layern bestehenden Architektur, mit welcher Beziehungen zwischen benachbarten, aber unkorrelierten Pixeln bzw. Merkmalen verhindert werden können, liegt insbesondere in einer deutlichen Reduktion des Trainingsaufwandes.

Da CNNs hauptsächlich auf die Verwendung im Bildbereich ausgelegt sind, könnte der für jeden Punkt vorliegende eindimensionale Merkmalsvektor jeweils in ein zweidimensionales Bild umgeformt werden und in dieser Form in das CNN eingehen. Um jedoch dabei eben solche durch die Faltungslayer implizierte Beziehungen zwischen benachbarten Merkmalen zu minimieren, schlagen GEORGE et al. (2018) die Nutzung eben jenes eindimensionalen CNNs vor. Die Grundidee der Verwendung mehrerer Äste in der Architektur dieses Multi-Branch 1D-CNN liegt im unabhängigen Training der Teilmerkmalsvektoren unterschiedlicher Skalen in den gegebenen Ästen. Die



Abb. 4: Netzarchitektur des Multi-Branch 1D-CNN (nach GEORGE et al. (2018)). Der ursprüngliche LiDAR-Merkmalsvektor wird auf Grundlage des jeweiligen Nachbarschaftsradius in mehrere Teilmerkmalsvektoren zerlegt und jeweils durch die punktweise berechenbaren Merkmale ergänzt. Diese gehen in mehrere aus Faltungslayern, Batch-Normalisierungen, Leaky-ReLU-Aktivierungsfunktionen und Max-Pooling-Layern bestehende Branches ein. Deren Ergebnis wird anschließend zusammengefasst. Die nun deutlich reduzierte Datenmenge kann dann in einen Fully-Connected-Layer einfließen, um unter nochmaliger Nutzung eines solchen Layers das Klassenlabel zu prädizieren.

prozessierten Teilmerkmalsvektoren werden erst anschließend zur Übergabe an die Fully-Connected-Layer und zum Erhalt der Klassenlabel kombiniert. Ein solcher Teilmerkmalsvektor enthält dabei neben den Merkmalen, die speziell für eine Nachbarschaft des entsprechenden Radius evaluiert worden sind, auch jeweils diejenigen Merkmale, die unabhängig von einer Nachbarschaft sind und punktweise berechnet werden können.

2.4 Analyse der Ergebnisse

Für den LiDAR-Datensatz konnte unter Nutzung des RF-Klassifikators mittels optimierter Hyperparameter (300 Entscheidungsbäume der Tiefe 18, minimale Sampleanzahl von 7 für einen neuen Entscheidungstest) und integrierter Merkmalsberechnung mit einer Overall Accuracy (OA) von 86,80% (vgl. Tab. 1) das beste Klassifikationsergebnis erzielt werden. Die zugehörige, bezüglich der Referenzdaten normierte Konfusionsmatrix ist in Abb. 5 dargestellt, wobei im Folgenden auf die größten Fehlzuordnungen exemplarisch eingegangen werden soll.

Der mit 40% höchste Grad der Verwechslung ergibt sich durch die Zuordnung von offenem Boden/Geröll zu *Niedrige Vegetation*. Begründet ist dieser Effekt durch ähnliche relative Höhen und nahezu identische Geometrien. Auch das Hinzuziehen von Farbinformationen kann hier, bedingt durch Überschneidungen in der bräunlichen Farbgebung zwischen offenem Boden und Bereichen dünner Begrasung, keine Klassifikationsverbesserung bewirken. Verstärkt wird dieser Effekt durch den Zeitpunkt der Befliegung Mitte März, da zu dieser Jahreszeit Vegetationsflächen noch nicht durch ihre typische grüne Farbe gekennzeichnet sind. Ferner werden 26% der Punkte, welche tatsächlich Fahrzeuge repräsentieren, fälschlicherweise als *Sonstige Ausstattung* interpretiert. Verursacht wird dies insbesondere durch die Vielfalt innerhalb dieser Klasse, welche eine Fülle unterschiedlicher Farben und Geometrien beinhaltet. Dies erschwert eine Abgrenzung zur Klasse *Fahrzeug*, welche ebenfalls durch große Farbvariationen geprägt ist und deren Vertreter auch bezüglich der relativen Höhen denjenigen der Klasse *Sonstige Ausstattung* ähneln. Weitere Verwechslungen aufgrund jeweils ähnlicher Geometrien und Farbgebung treten erwartungsgemäß zwischen *Fassade* und *Vertikale Fläche* sowie zwischen *Strauch/Hecke* und *Baum* auf.

Das erhaltene Klassifikationsergebnis wird zusätzlich in Abb. 6 graphisch den Referenzdaten gegenübergestellt, woraus eine weitgehende Übereinstimmung erkennbar ist. Abweichungen fallen insbesondere anhand der Baustelle im unteren Bereich auf, wobei sich hier das Klassifikationsergebnis in Teilen von der tatsächlichen Klasse *Offener Boden/Geröll* unterscheidet. Weiterhin zeigen sich auch Fehlzuordnungen von Zäunen, welche oft fälschlicherweise als *Strauch/Hecke* klassifiziert werden. Dies ist insofern nachvollziehbar, als beispielsweise eine Weinrebe, welche per Definition als *Strauch/Hecke* gilt, oft nur schwerlich von einem leicht überwucherten Zaun zu trennen ist.

Dieses Klassifikationsergebnis fällt insgesamt trotz höherer Klassenanzahl und damit anspruchsvollerer Klassifikationsaufgabe besser aus als das beste Ergebnis für den ISPRS Benchmark Datensatz Vaihingen/Enz (ZHAO et al. 2018). Da für diesen jedoch keine gleichzeitig erfassten Bilddaten vorliegen, kann eine Anwendung der vorgestellten Methodik auf diesen Datensatz zum direkten Vergleich nicht erfolgen.

		Konfusionsmatrix, normiert										
Klassen aus Klassifikation	Stromleitung -	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Niedrige Vegetation -	0.00	0.96	0.08	0.03	0.07	0.00	0.02	0.07	0.00	0.40	0.02
	Versiegelte Fläche -	0.00	0.01	0.86	0.04	0.01	0.03	0.01	0.00	0.00	0.15	0.05
	Fahrzeug -	0.00	0.00	0.00	0.51	0.01	0.00	0.00	0.00	0.00	0.00	0.00
	Sonstige Ausstattung -	0.00	0.00	0.00	0.26	0.35	0.00	0.03	0.01	0.00	0.00	0.02
	Dach -	0.00	0.00	0.02	0.02	0.08	0.92	0.05	0.00	0.00	0.00	0.01
	Fassade -	0.00	0.00	0.01	0.02	0.14	0.02	0.84	0.01	0.00	0.00	0.16
	Strauch/Hecke -	0.00	0.01	0.00	0.07	0.17	0.00	0.01	0.70	0.00	0.00	0.00
	Baum -	0.00	0.01	0.00	0.04	0.11	0.03	0.03	0.21	0.99	0.00	0.00
	Offener Boden/Geröll -	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.44	0.01
	Vertikale Fläche -	0.00	0.00	0.00	0.01	0.04	0.00	0.02	0.00	0.00	0.00	0.74
		Stromleitung -	Niedrige Vegetation -	Versiegelte Fläche -	Fahrzeug -	Sonstige Ausstattung -	Dach -	Fassade -	- Strauch/Hecke -	Baum -	Offener Boden/Geröll -	Vertikale Fläche -
		Klassen aus Referenzdaten										

Abb. 5: Darstellung des Klassifikationsergebnisses des RF-Klassifikators für den LiDAR-Datensatz bezüglich des Mehrklassenfalls anhand der normierten Konfusionsmatrix.



Abb. 6: Vergleich des Klassifikationsergebnisses des RF-Klassifikators für den LiDAR-Datensatz (unten) mit Referenzdaten (oben) bzgl. des Mehrklassenfalls. (0=Stromleitung, 1=Niedrige Vegetation, 2=Versiegelte Fläche, 3=Fahrzeug, 4=Sonstige Ausstattung, 5=Dach, 6=Fassade, 7=Strauch/Hecke, 8=Baum, 9=Offener Boden/Geröll, 10=Vertikale Fläche). Insbesondere die Klassen Dach, Versiegelte Fläche, Baum und Stromleitung werden detektiert. Die Klasse Offener Boden/Geröll wird oft nicht oder nur in Teilen erkannt. Eine alternative Klassifikation mittels des in 2.3 vorgestellten Ansatzes des Multi-Branch 1D-CNN erreicht nur eine geringfügig schlechtere OA von 84,67% (vgl. Tab. 1).

	Randor	n Forest	Multi-Branch 1D-CNN		
Klassenbezeichnung	PA [%]	UA [%]	PA [%]	UA [%]	
Stromleitung	100	100	98,10(-1,90)	96,27(-3,73)	
Niedrige Vegetation	96,37	85,00	94,22(-2,15)	85,01(+0,01)	
Versiegelte Fläche	86,26	90,03	78,31(-7,95)	89,21(-0,79)	
Fahrzeug	51,31	88,39	64,87(+13,56)	66,55(-21,84)	
Sonstige Ausstattung	35,12	68,39	43,54(+8,42)	55,98(-12,41)	
Dach	91,60	87,49	92,23(+0,63)	81,50(-5,99)	
Fassade	83,89	81,59	80,71(-3,18)	80,97(-0,62)	
Strauch/Hecke	69,54	69,56	73,99(+4,45)	60,58(-8,98)	
Baum	99,15	92,69	97,06(-2,09)	95,63(+2,94)	
Offener Boden/Geröll	44,48	79,73	37,51(-6,97)	65,74(-13,99)	
Vertikale Fläche	73,74	85,70	73,16(-0,58)	76,28(-9,42)	
OA [%]	86	,80	84,67(-2,13)		

Tab. 1: Vergleich der Genauigkeitsmaße des RF-Klassifikators mit dem Multi-Branch 1D-CNN-Klassifikator für den LiDAR-Datensatz bzgl. des Mehrklassenfalls (PA=Producer's Accuracy, UA=User's Accuracy).

Die Ursache für das beim Multi-Branch 1D-CNN bessere Ergebnis für die Producer's Accuracy (PA) insbesondere in Bezug auf die Klassen *Fahrzeug*, *Sonstige Ausstattung* und *Strauch/Hecke* (vgl. Tab. 1) liegt darin, dass an die Loss-Funktion eine Gewichtung bezüglich der einzelnen Klassen angebracht wird, mit dem Ziel, schwach vertretene Klassen im Rahmen des Trainings stärker zu berücksichtigen. Konkret wird die Gewichtung für jede Klasse aus der Inversen der Wurzel aus den relativen Häufigkeiten der Klassen in den Trainingsdaten abgeleitet.

Die Klassifikation des MVS-Datensatzes mittels RF fällt mit einer OA von 81,60% etwas schlechter aus, da diese Punktwolke aufgrund von Verdeckungen und erschwertem Matching in Vegetationsbereichen von einer im Vergleich zum LiDAR-Datensatz unvollständigeren Geometrie gekennzeichnet ist.

Für die zu Monitoringzwecken ausreichende binäre Klassifikation konnte durch erneutes Trainieren auf binären Referenzdaten für den LiDAR-Datensatz mittels RF mit einer OA von 95,71% das beabsichtigte Ziel der Detektion von festen, nicht-vertikalen Flächen erreicht werden. Die so aufgefundenen Flächen gehen mit einer Zuverlässigkeit von 94,23% in das Deformationsmodell ein.

3 Fazit & Ausblick

Im Rahmen der Klassifikation hat sich die gleichzeitige Verfügbarkeit von Bild- und LiDAR-Daten für eine integrierte Merkmalsextraktion als vorteilhaft herausgestellt, sodass das Ziel der Detektion von für Monitoringzwecke geeigneten Flächen erreicht werden konnte. Als Klassifikator mit bereits sehr guter Genauigkeit hat sich zudem das Multi-Branch 1D-CNN erwiesen, wenngleich dieses vom RF noch übertroffen wird. Nachteil des CNN-Klassifikators ist jedoch das für diesen Klassifikatortyp typische Erfordernis einer großen Menge an Trainingsdaten, welche durch die vorliegenden hochaufgelösten Punktwolken zwar gestellt werden, mittels derzeit verfügbarer Implementierungen und Hardware jedoch noch nicht ohne erheblichen Zeitaufwand ausgenutzt werden kann. Wenn dieses Problem noch bewältigt werden kann, hat dieser Klassifikator das Potential, den genutzten RF im Hinblick auf die Genauigkeit zu übertreffen. In Bezug auf das im Rahmen dieser Arbeit vorgestellte Monitoringprojekt soll zudem künftig untersucht werden, in-wieweit der auf den Daten der ersten Epoche basierende Klassifikator auf Daten anderer Epochen abweichender Jahreszeiten anwendbar ist.

4 Literaturverzeichnis

- AGISOFT PHOTOSCAN, 2018: PhotoScan photogrammetric processing of digital images and 3D spatial data generation. Website. (http://www.agisoft.com/; letzter Zugriff 27.11.2018)
- BADRINARAYANAN, V., KENDALL, A. & CIPPOLLA, R., 2017: SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), 2481-2495.
- BECKER, C., HÄNI, N., ROSINSKAYA, E., D'ANGELO, E. & STRECHA, C., 2017: Classification of Aerial Photogrammetric 3D Point Clouds. ISPRS Ann. Photogr. Rem. Sens. Spat. Inf. Sci., IV-1/W1, 3-10.
- BREIMAN, L., 2001: Random Forests. Journal of Machine Learning, 45(1), 5-32.
- CHEHATA, N., GUO, L. & MALLET, C., 2009: Airborne LiDAR feature selection for urban classification using random forests. Int. Arch. Photogr. Rem. Sens. and Spat. Inf. Sci., **38**(3/W8), 207-212.
- CRAMER, M., HAALA, N., LAUPHEIMER, D., MANDLBURGER, G. & HAVEL, P., 2018: Ultra-High Precision UAV-based LiDAR and Dense Image Matching. Int. Arch. Photogr. Rem. Sens. Spat. Inf. Sci., **42**(1), 115-120.
- GEORGE, D., XIE, X. & TAM, G., 2018: 3D Mesh Segmentation via Multi-branch 1D Convolutional Neural Networks. Graphical Models, **96**, 1-10.
- HACKEL, T., WEGNER, J. & SCHINDLER, K., 2016: Fast semantic segmentation of 3D point clouds with strongly varying density. ISPRS Ann. Photogr. Rem. Sens. Spat. Inf. Sci., III-3, 177-184.
- NIEMEYER, J., ROTTENSTEINER, F. & SÖRGEL, U., 2014: Contextual classification of LiDAR data and building object detection in urban areas. ISPRS Jour. Photogr. Rem. Sens., 87, 152-165.
- PFEIFER, N., MANDLBURGER, G., OTEPKA, J. & KAREL, W., 2014: OPALS A framework for Airborne Laser Scanning data analysis. Comp. Environ. Urban Sys., **45**, 125-136.
- WEINMANN, M., BLOMLEY, R., WEINMANN, M. & JUTZI, B., 2018: Investigations on the Potential of Binary and Multi-class Classification for Object Extraction from Airborne Laser Scanning Point Clouds. Tagungsband der 38. Wissensch.-Tech. Jahrestagung der DGPF, 27, 408-421.
- ZHAO, R., PANG, M. & WANG, J., 2018: Classifying airborne LiDAR point clouds via deep features learned by a multi-scale convolutional neural network. Int. Jour. Geo. Inf. Sci., 32(5), 960-979.

Semantic Data Cubes Utilising Free and Open-Access EO-Data to Generate Spatially-Explicit Evidence for Environmental Monitoring: Applied Use-Case in Syria Based on Sentinel-2 Data

HANNAH AUGUSTIN¹, MARTIN SUDMANNS¹, DIRK TIEDE¹ & ANDREA BARALDI²

Abstract: Collections of free and open-access Earth observation (EO) data with global coverage are growing with increasingly higher spatial resolutions and temporal frequency. They are one of the few globally consistent data sources available for generating information in support of international initiatives. However, these data require automated workflows for handling, processing and analysis, including methods to convert data into valid information.

A proof-of-concept implementation of a semantic EO data cube is presented using Open Data Cube technology to generate ad-hoc reproducible, scalable, repeatable and spatiallyexplicit information as indicators to support monitoring international environmental initiatives. Sentinel-2 data for the study area in north-western Syria (30,000km²) are incorporated daily, including automatically generated generic semantic enrichment. As of December 2018, this encompasses over 800 scenes (~480GB unprocessed Sentinel-2 data). A semantic query resulting in a normalised vegetation index of occurrence over 3 years for the entire study area is demonstrated as a proof-of-concept example.

1 Context

Technological advances have driven many changes in Earth observation (EO) from space, including new, innovative sensors and ways to handle, store and process rapidly growing data archives. In 1972, the launch of the Landsat programme began what is now the longest record of Earth's status and dynamics from space. Opening the archive to public access in 2008 set the stage for free and open EO data (WULDER et al. 2012). The Sentinel satellites from the EU's EO programme, Copernicus, have provided data since 2014. This has increased the spatial resolution and temporal frequency of freely and openly available EO data from a variety of sensors, including radar and multi-spectral instruments. Often referred to as big Earth data, the challenges this type of data poses researchers are not merely technological, in terms of data storage, access and processing, but rather in developing methods to distill *information* from this wealth of data.

Free and open, high-resolution EO data are ideal sources of evidence for generating meaningful information products to support decision-makers in an international context. They provide consistent global coverage, independent of political or other human-imposed borders, offering potential for large-scale, multi-temporal and persistent monitoring and analysis, especially with continued data acquisition for years to come (DRUSCH et al. 2012). The data are constantly gathered without requiring tasking or direct acquisition costs for those utilising them, and their high

¹ Department of Geoinformatics -- Z_GIS, University of Salzburg, Schillerstr. 30,

A-5020 Salzburg, E-Mail: [hannah.augustin, martin.sudmanns, dirk.tiede]@sbg.ac.at

² Italian Space Agency (ASI), Via del Politecnico, I-00133 Rome RM, E-Mail: andrea6311@gmail.com

temporal frequency could improve timeliness of actionable information for global initiatives given automated information extraction processes (e.g. automated-prior-knowledge based or machine learning classification procedures).

Semantic enrichment (i.e. generating meaningful information) is necessary for turning data into understandable information products. Optical EO data (e.g. Sentinel-2) cannot directly measure most objects, processes or events on Earth because digital numbers are not a standard unit and many different surfaces can be represented by similar values. Indicator extraction is one way to translate this data into meaningful information. Automatic spectral categorisation (i.e. preliminary classification) is one existing transferable method for initial, generic semantic enrichment that can be used to generate indicators (BARALDI et al. 2010).

Automated workflows are necessary for handling the Sentinel-2 mission's expected 3.4TB of daily data *volume* (EUROPEAN SPACE AGENCY 2017). The data have a relatively high *velocity* due to global coverage, on average every five days at the equator, and quite a *variety* in terms of consistency and quality levels (e.g. cloud coverage differs depending on the spatio-temporal location) (SOILLE et al. 2018).

The work presented here is an example of an automated, reproducible framework for handling and analysing big EO data. It demonstrates the benefits of automated, knowledge-based, generic semantic enrichment as the basis for multi-temporal, spatial, semantic queries to produce diverse, transferable, semi-automated indicators for a variety of domains, including environmental monitoring.

2 Applied Use-Case

The north-western Syrian border to Turkey was chosen as the use-case location based on existing EO-based research (TIEDE et al. 2014), low average cloud cover in the Sentinel-2 archive, and the currently on-going conflict, which makes other methods of data acquisition challenging or impossible. The area is comprised of three Sentinel-2 granules covering over 30,000km² (lati-tudes 36.01°-37.05°N; longitudes 35.67°-39.11°E). Data included is from 28 June 2015 until the time of writing (19 December 2018; see Fig. 1) and incorporates any new data available in the Copernicus Open Access Data Hub on a daily basis. Two relative orbits cover this area resulting in temporally denser data availability where they overlap (Fig. 2).

2.1 Proof-of-Concept Implementation

The workflow implemented here focuses on automation and big data. It encompasses automated downloads from the Copernicus Open Access Hub, re-formatting Sentinel-2 data for processing, preliminarily classification with the Satellite Image Automatic Mapper (SIAMTM) (i.e. generating multiple information layers), indexing Sentinel-2 scenes and information layers into an implementation of the ODC and ingesting information layers (Fig. 3). This process runs automatically every day for each of the three study area Sentinel-2 granules (37SBA, 37SCA, 37SDA). The result is daily incorporation of the most recently available Sentinel-2 data ready for analysis that can include semantic queries. Queries are facilitated using Jupyter notebooks by accessing the ingested information layers produced by SIAMTM via the ODC's python API. At the time or

writing, over 800 scenes (ca. 480GB unprocessed Sentinel-2 data) and their information layers are ingested and can be queried.



*all unprocessed Sentinel-2 scenes, information layers and reformatted stack





Fig. 2: Study area based on 3 overlapping granules (e.g. 37SBA, 37SCA, 37SDA provided in UTM Zone 37N) with overlapping Sentinel-2 relative acquisition orbits (modified from AUGUSTIN et al., 2018)



Fig. 3: Automated workflow overview from download to data cube incorporation (AUGUSTIN et al., 2018)

Spectral-based image pre-classification, as implemented by SIAM[™], divides the feature space of a multi-spectral image into semantic semi-concepts using a knowledge-based approach, which is in contrast to data-driven approaches (e.g. supervised classification) (BARALDI et al. 2010; BAR-ALDI 2011; BARALDI 2018). This could also be called descriptive colour naming, because the resulting semi-categories refer to similar pixels in terms of the multi-spectral information a pixel can offer. Assuming scenes are calibrated to a minimum of top-of-atmosphere reflectance, semiconcepts generated by SIAM[™] are comparable and transferable between multiple images and optical sensors without any additional user-defined parameterisation (i.e. fully automatic). This is the only software used in the workflow that is closed-source.

All other software is open-source. The computing environment used for this implementation is a Red Hat Enterprise Linux 7 virtual machine with 16 virtual CPU at 2.5GHz clocking, 31GB RAM and 3TB of generic, all-use storage. The entire workflow is implemented using python in reproducible virtual environments using bash scripts run as automated cronjobs in Linux. The ODC technology is implemented with a PostgreSQL backend. Refer to AUGUSTIN et al. (2018) for more detailed information about the technical implementation.

2.2 Discussion and Example

The automated workflow stops at incorporation into the ODC implementation, resulting in a semantic EO data cube. This data cube can then be semi-automatically queried using existing Jupyter notebooks (i.e. existing blocks of interactive code), to construct various queries utilising the ODC's python API.

Concepts of reproducibility were strong drivers behind this implementation. The reproducibility of information extraction in this case is highly linked to the level of automation of information production, which is quite high. The methods and results of any query ought to be reproducible, given access to the same Sentinel-2 data, versions of the SIAMTM and ODC software, python computing environments, code and queries. Incorporating all Sentinel-2 data available for an

area including information layers generated through pre-classification in an implementation of the ODC can currently be considered as providing the data in an analysis-ready way. Due to the fully automated preparation of data from acquisition to ODC ingestion, this implementation can be considered highly repeatable. Given the automation developed in this work, a copy of this semantic data cube could be rebuilt in a similar computing environment in an estimated 5 days, assuming at least four parallel SIAMTM processes and stable downloads.

One example of a semantic query is for a normalised index of observed semi-concepts over time, e.g. for vegetation-like semi-concepts after having filtered out cloud-like semi-concepts (Fig. 4). As seen in Figure 5, such a query was conducted for nearly 3 years of data (28 June 2015 until 22 June 2018) over the entire spatial extent of the implementation (i.e. over 30,000km² with 10m pixels) (Fig. 5c and 5e). In addition, a query for the number of scenes available per pixel was conducted (Fig. 5a), as well as calculating the number of "clean pixels" by excluding cloud-like, snow-like, and unknown semi-concepts (i.e. relatively high reflectance in multiple spectral bands) (Fig. 5b). Each of these three massive queries took around 4 hours to complete in the given computing environment. Keep in mind, each query accessed the information layers belonging to 591 independent Sentinel-2 scenes. Figure 5d displays free and open OpenStreetMap (OSM) data to demonstrate the plausibility of the results, whereby multiple canals, buildings and even the border are visible in the absence of vegetation-like observations, and visible field structures coincide with the normalised vegetation occurrence output.

For such a massive query in terms of space, time and data volume, the number of scenes and clean pixels helps contextualise the heterogeneity of data underlying the results. What is lacking is a metric or analysis of distribution through time, which, for example, would make seasonal cloud-cover differences over large areas visible. Additionally, the simplified "clean pixel" calculation tends to exclude buildings (e.g. city of Aleppo), shallow water bodies (e.g. saline lake, Jabbūl, in the south), and other highly reflective, ambiguous surfaces. This could be overcome with more complex rules to take temporal (in)stability of semi-concepts into account, or their spatial neighbours (e.g. moving towards object detection and object-based methods in space and time).



Fig. 4: This demonstrates how vegetation-like semi-concept occurrence aggregated over time is calculated. Green represents various vegetation-like semi-concepts, light-grey cloud-like, and pink/magenta anything else (e.g. bare soil). The clean stack contains observations excluding cloud-like semi-concepts following semantic-enrichment. The query selection contains only vegetation-like observations. Author's illustration



Query of Entire Study Area: 2015-06-28 until 2018-06-22

Fig. 5: Multiple semantic queries of the same spatio-temporal extent were conducted: (a) number of available scenes; (b) number of pixels excluding cloud-like, snow-like and unknown semi-concepts; (c) normalised vegetation index based on vegetation-like semi-concepts. A closer look of (c) identified with a red inset square is visible in (e). Existing OSM data for the same area and scale as (e) is visible in (d). Author's illustration

3 Looking Forward

The contribution and innovation demonstrated by this proof-of-concept implementation is the automated set-up of a semantic data cube. The semantic data cube contrasts to most other existing data cube implementations because it stores information and data together in an analysis-ready way, allowing ad-hoc multi-temporal, spatial and semantic queries, while supporting reproducible results. The generic, application-independent semantic enrichment utilised enables queries and EO-based indicator extraction for many thematic tasks. The semantic semi-concepts can be thought of as transferable, reproducible, sensor-agnostic building blocks for conducting further analysis. Given solid documentation on methods applied to generate output, reproducible results and repeatable analysis ought to be possible since the information layers (i.e. basis for semantic queries and analysis) continue to exist in the data cube and are stable concepts. This could be particularly relevant for supporting global initiatives (e.g. UN's sustainable development goals (UNITED NATIONS 2015b), Sendai Framework for Disaster and Risk Reduction (UNITED NATIONS 2015b), because information is based on data collected independent of political borders and in a constant, relatively unbiased way.

However, the work presented here is not only a technical implementation towards developing indicators, but also an initial exploration of some of the challenges faced when working with dense time-series of EO data over larger spatial areas and timespans. These challenges directly result from qualities that are characteristic of big Earth data, i.e. their volume, velocity, and variety, but also varied methods of information generation and heterogeneity in underlying data and assumptions. This heterogeneity can encompass the distribution, variation, variability and uncertainty in space and time for all multi-temporal EO analysis and archives, especially when covering relatively large spatial extents.

This is a moment in time with unprecedented processing capabilities and free and open data availability. It is increasingly important to have meaningful, comprehensive and standardised methods to characterise and visualise uneven spatio-temporal distribution and coverage, uncertainty and variability as well as variation in data quality (e.g. cloud coverage) for different big Earth data sources and archives. Just because the data can possibly be considered unbiased in their regular, global acquisition does not mean that information generated from them are unbiased or will be used and understood in ways that are not misleading to different audiences. The more data that is incorporated in analysis and information generation, the more important it becomes to characterise the underlying data in terms of difference in quality and the previously mentioned characteristics.

Further research will include: (1) increasingly expressive or comprehensive rule-sets for queries taking spatial (e.g. neighbours, texture, objects), or temporal context (e.g. before and after, (dis)continuity, patterns like phenology) into account; (2) developing reliable indicators tested for agreement with existing sources that may be relevant to existing international initiatives; and (3) exploring methods and metrics to better assess the distribution, variation, variability and uncertainty inherent to dense, multi-temporal EO analysis and archives.

4 Acknowledgements

This work was supervised by Dr. Dirk Tiede and Martin Sudmanns, MSc, and financially supported by the Austrian Federal Ministry of Transport, Innovation and Technology (BMVIT) under the program "ICT of the Future" within the project SemEO (contract number: 855467). Continued work will be financially supported within the scope of the Sen2Cube.at project (contract number: 866016) funded by the Austrian Research Promotion Agency (FFG) under the Austrian Space Applications Programme (ASAP 14). None of this would have been possible without the contribution and support of Dr. Andrea Baraldi and access to his software, SIAMTM. A free and openly available PDF copy of the complete master thesis this contribution is based on is available at https://github.com/augustinh22/msc_markdown/ in the output directory.

5 References

- AUGUSTIN, H., SUDMANNS, M., TIEDE, D. & BARALDI, A., 2018: A Semantic Earth Observation Data Cube for Monitoring Environmental Changes during the Syrian Conflict. GI_Forum 2018, 1, 214-227.
- BARALDI, A., 2018: Satellite Image Automatic MapperTM System and Products Description. Retrieved 29 January 2018, from http://siam.andreabaraldi.com/content/Documentation/SIAM Report BACRES v1.18.pdf
- 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 Transactions on Geoscience and Remote Sensing, 48(3), 1299-1325.
- BARALDI, A., 2011: Satellite Image Automatic MapperTM (SIAMTM) A Turnkey Software Executable for Automatic Near Real-Time Multi-Sensor Multi-Resolution Spectral Rule-Based Preliminary Classification of Spaceborne Multi-Spectral Images. Recent Patents on Space Technology, 1(2), 81-106.
- DRUSCH, M., DEL BELLO, U., CARLIER, S., COLIN, O., FERNANDEZ, V., GASCON, F., ... BARGELLINI, P., 2012: Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. Remote Sensing of Environment, **120**, 25-36.
- EUROPEAN SPACE AGENCY, 2017: Sentinel High Level Operations Plan (HLOP): COPE-S1OP-EOPG-PL-15-0020. Retrieved from https://earth.esa.int/documents/247904/685154/Sentinel High Level Operations Plan
- SOILLE, P., BURGER, A., DE MARCHI, D., KEMPENEERS, P., RODRIGUEZ, D., SYRRIS, V. & VA-SILEV, V., 2018: A versatile data-intensive computing platform for information retrieval from big geospatial data. Future Generation Computer Systems, **81**, 30-40.
- TIEDE, D., LÜTHJE, F. & BARALDI, A., 2014: Automatic post-classification land cover change detection in Landsat images: Analysis of changes in agricultural areas during the Syrian crisis. Publikationen der Deut. Gesellschaft für Photogrammetrie, Fernerkundung u. Geoinformation e.V., Band 23, Beitrag 191.
- UNITED NATIONS, 2015a: Resolution 69/283: Sendai Framework for Disaster Risk Reduction 2015–2030. United Nations General Assembly. Retrieved from https://undocs.org/A/RES/69/283
- UNITED NATIONS, 2015b: Resolution 70/1: Transforming our world: the 2030 agenda for sustainable development. United Nations General Assembly. Retrieved from https://undocs.org/A/RES/70/1
- WULDER, M. A., MASEK, J. G., COHEN, W. B., LOVELAND, T. R. & WOODCOCK, C. E., 2012: Opening the archive: How free data has enabled the science and monitoring promise of Landsat. Remote Sensing of Environment, 122, 2-10.

Untersuchung von Orientierungs- und Matchingverfahren für die hochgenaue 3D-Oberflächenerfassung von Schweißnähten mit einem mobilen Kamerasystem

SIMON NIETIEDT¹, OLIVER KAHMEN¹ & THOMAS LUHMANN¹

Zusammenfassung: Im Folgenden werden die Ergebnisse einer Masterarbeit vorgestellt, die im Rahmen des Forschungsprojektes "Entwicklung eines kompakten Prototyps zur hochgenauen 3D-Oberflächenmessung unter Wasser" durchgeführt wurde. Ziel der Arbeit war es, verschiedene Orientierungs- und Matchingverfahren für die hochgenaue Erfassung von Schweißnähten zu untersuchen. Hierfür werden ausschließlich synthetische Daten verwendet, sodass die Erkenntnisse zusätzlich auf weitere Anwendungen übertragbar sind und der jeweilige Einfluss einzelner Parameter gezielt simuliert werden kann. Die durchgeführten Untersuchungen zeigen auf, dass die aktuell verwendete Aufnahmekonfiguration modifiziert werden muss, wenn der gesamte Bildverband anhand der untersuchten Orientierungsverfahren mit der erforderlichen Genauigkeit orientiert werden soll. Zudem eignen sich beide der untersuchten Matchingverfahren prinzipiell für die hochgenaue 3D-Oberflächenerfassung, da die erforderliche Rekonstruktionsgüte erreicht wird.

1 Motivation

Das Schweißen hat in den letzten Jahrzehnten durch die ansteigenden technologischen Anforderungen an Bedeutung gewonnen. Besonders im Bereich risikobehafteter Unterwasserbauwerke, wie beispielsweise Pipelines, müssen Schweißnähte hohen Qualitätsanforderungen genügen, die durch aktuelle Prüfverfahren nur schwer zu verifizieren sind. Deshalb steigt nach NEILL (2016) die Nachfrage an objektiven Prüfverfahren, die kleine Defekte zuverlässig aufdecken können. Hierfür haben bereits EKKEL et al. (2015) ein Prüfverfahren entwickelt, das unter Wasser eingesetzt werden kann und die Oberfläche mithilfe einer projizierten Laserlinie erfasst. Zudem wird aktuell am Institut für Angewandte Photogrammetrie und Geoinformatik der Jade Hochschule (IAPG) im Rahmen des Projektes "Entwicklung eines kompakten Prototyps zur hochgenauen 3D-Oberflächenmessung unter Wasser" ein neues Prüfverfahren entwickelt. Das Projekt wird durch den Europäischen Fonds für regionale Entwicklung (EFRE) gefördert. Das Messkonzept der Neuentwicklung sieht vor, die Schweißnahtoberflächen mithilfe von Matchingverfahren hochgenau zu rekonstruieren. Die praxisnahe Verifizierung des Konzeptes ist jedoch aufgrund der speziellen Umweltbedingungen und den hohen Genauigkeits- und Auflösungsanforderungen von 30 µm derzeit noch nicht realisierbar und mit sehr hohem praktischen Aufwand verbunden. Deshalb wurde in der zugrundeliegenden Masterarbeit ein Konzept erarbeitet, um Orientierungsund Matchingverfahren anhand von synthetischen Daten möglichst realitätsnah zu untersuchen.

¹ Jade Hochschule, Institut für Angewandte Photogrammetrie und Geoinformatik, Ofener Straße 16/19, D-26121 Oldenburg, E-Mail: [simon.nietiedt, oliver.kahmen, thomas.luhmann]@jade-hs.de

2 Konzeption

Das Konzept, welches aktuell am IAPG entwickelt wird, ist schematisch in Abbildung 1 dargestellt. Hierbei kann das mobile Kamerasystem mithilfe von Messmarken, die um die Schweißnaht herum angeordnet sind, orientiert werden. Die verwendeten Bilder der einzelnen Kamerapositionen tasten die Oberfläche sequentiell ab, wodurch die gesamte Schweißnaht in überlappenden Bildbereichen vorliegt. Zwei Einzelbilder können als Stereosystem aufgefasst werden, sodass die Rekonstruktion der Schweißnahtoberfläche durch Vorwärtsschnitte erfolgen kann. Die benötigten korrespondierenden Pixel können mit einem Matchingverfahren bestimmt werden.



Abb. 1: Schematische Darstellung des Aufnahmekonzeptes für die Erfassung von Schweißnähten

Die Genauigkeit, Vollständigkeit und Auflösung der Rekonstruktion hängt im Wesentlichen von der Aufnahmekonfiguration, den Einflüssen der Umwelt sowie von den verwendeten Orientierungs- und Matchingverfahren ab. In diesem Beitrag wird die Untersuchung der Aufnahmekonfiguration im Zusammenhang mit den Orientierungs- und Matchingverfahren in drei Abschnitte gegliedert.

2.1 Untersuchung der Matchingverfahren

Als Matchingverfahren werden das Semi-Global Matching (SGM) und Least-Squares Matching (LSM) untersucht. Beim LSM handelt es sich um ein Zuordnungsverfahren, welches die optimale Zuordnung über eine nichtlineare Ausgleichung nach kleinsten Quadraten ermittelt. Hierbei wird nach GRÜN (1985) als funktionales Modell eine Affintransformation verwendet, die um zwei radiometrischen Parameter erweitert wurde. Beim SGM nach HIRSCHMÜLLER (2005) wird dagegen folgende Energiefunktion minimiert, welche durch die optimalen Disparitäten D beschrieben wird.

$$E(D) = \sum_{p} \left(C(p, D_p) + \sum_{q \in N_p} P_1 T[|D_p - D_q| = 1] + \sum_{q \in N_p} P_2 T[|D_p - D_q| > 1] \right)$$

Um diese Disparitäten zu berechnen, wird ein beliebiges Ähnlichkeitsmaß $C(p, D_p)$ eingesetzt. Zusätzlich werden die Strafterme P_1 und P_2 verwendet, um den Einfluss der benachbarten Pixel N_p zu steuern.

Für die Untersuchung von Matchingverfahren haben unter anderem SCHARSTEIN & SZELISKI (2002, 2003) und SCHARSTEIN et al. (2014) unterschiedliche Datensätze von verschiedenen Objekten veröffentlicht. Diese Datensätze konnten jedoch nicht im Rahmen dieser Untersuchungen verwendet werden, da die Aufnahmekonfiguration weitestgehend unbekannt ist und die Objektgeometrien nicht mit der gefordertern Genauigkeit und Auflösung vorliegen. Daher wurden zwei synthetische Testfelder in Form eines rotationssymmetrischen Sinusoids entwickelt, die sich durch zehn unterschiedliche Ringe auszeichnen. Während beim ersten Testfeld, welches in Abbildung 2 dargestellt ist, die Höhe der Ringe variiert, unterscheiden sich die Ringe beim zweiten Testfeld durch ihre Breite. Die Wahl der Textur basiert auf einem Texturmerkmal, dass vorab hinsichtlich der Nutzbarkeit zur Beschreibung des Zuordnungsniveaus untersucht wird. Hierfür wird unter anderem das Leistungsspektrum untersucht, dass nach POHLE (2004) feine Strukturen durch hohe Frequenzen beschreibt. Des Weiteren wird der Grauwertgradient analysiert, der nach SCHNEIDER (1991) als Beschreibungsmerkmal des Zuordnungsniveaus verwendet werden kann.



Abb. 2: Testfeld mit vertikaler Variation. Links: Draufsicht mit farblich markiertem Radius. Rechts: Vertikaler Querschnitt des farblich markierten Radius mit den verwendeten Höhen

Das Leistungsspektrum der verwendeten Textur zeichnet sich durch hohe Amplituden von niedrigen und hohen Frequenzen aus, da nach den Erkenntnissen der Voruntersuchungen somit eine hohe Zuordnung erwartet werden kann. Für die Untersuchung zum Einfluss der Fenstergröße (Bereich der Bildinformation) und beim SGM zusätzlich der Strafterme P_1 und P_2 , werden von beiden Testfeldern jeweils zwei fehlerfreie Normalbilder erzeugt, die einen Aufnahmeabstand von 50 mm und eine Basislänge von 20 mm aufweisen. Mithilfe dieser Bilddaten werden anschließend beide Verfahren mit denselben Fenstergrößen und Straftermen angewendet. Die nachfolgende Analyse der ermittelten Disparitäten erfolgt im Objektraum, weshalb die Objektkoordinaten durch Vorwärtsschnitte berechnet werden. Hierfür werden zusätzlich die fehlerfrei vorliegenden Orientierungsparameter des Stereopaares verwendet. Die resultierende Punktwolke, welche somit nur von der Güte des Matchings abhängig ist, wird anhand der fehlerfrei vorliegenden Referenz analysiert und bewertet.

2.2 Untersuchung der Orientierungsverfahren

Die Orientierungsverfahren werden anhand eines synthetischen Testfeldes untersucht, das aus Messmarken unterschiedlicher Höhe besteht. Von dem in Abbildung 3 dargestellten Testfeld werden insgesamt sieben Aufnahmen in Form eines Bildstreifens generiert, die jeweils einen Abstand von 10 mm besitzen. In jedem synthetischen Bild, dessen innere und äußere Orientierung fehlerfrei bekannt ist, liegen mindestens vier fehlerfreie Passpunkte und weitere Verknüpfungspunkte vor, die formatfüllend und nach dem Anwendungsfall (Abbildung 1) angeordnet sind. Diese Bilder werden verwendet, um den Einfluss der inneren Orientierung und der Bildmessgenauigkeit auf die Güte der äußeren Orientierung zu untersuchen. Hierfür wurden vorab für verschiedene Basislängen die zulässigen Abweichungen der äußeren Orientierung ermittelt, bei denen die mittels Vorwärtsschnitt berechneten Koordinaten um maximal 30 µm von den Referenzwerten abweichen.



Passpunktverteilung

Bildverbandes

Abb. 3: Testfeld für die Untersuchung der Orientierungsverfahren. Links: Verwendete Passpunkte (rot), Bereich der Verknüpfungspunkte (blau) und der Schweißnaht (grau). Rechts: Abgebildeter Bereich der äußeren Bilder (farbliche Quadrate) und Mittelpunkte der einzelnen Bilder

Im ersten Versuchsblock wird das Genauigkeitsniveau der Orientierungsverfahren *Bündelausgleichung* und *Rückwärtsschnitt* anhand vorgegebener Unsicherheiten der inneren Orientierung und der Bildmessung mithilfe einer Monte-Carlo Simulation (MCS) ermittelt. Diese ist nach HASTEDT (2004) aufgebaut und besteht im Rahmen dieser Arbeit aus 10.000 Simulationsschritten. Die benötigten normalverteilten Unsicherheiten richten sich nach den Ergebnissen einer real durchgeführten Testfeldkalibrierung einer Industriekamera der Firma Basler. Eine solche Kamera ist im realen System verbaut, wodurch die Kalibrierergebnisse als realistische Grundlage für die Untersuchungen dienen.

Um das Genauigkeitsniveau der einzelnen Parameter der inneren Orientierung und der Bildmessung zu ermitteln, werden ebenfalls der simulierte Bildverband und eine MCS verwendet. Hierfür wird bei der Simulation der Orientierungsverfahren jeweils nur ein Parameter modifiziert. Die modifizierte Standardabweichung wird iterativ solange angepasst, bis die vorab ermittelten zulässigen Variationen der resultierenden Abweichungen entsprechen. Die so bestimmten Standardabweichungen spiegeln somit das erforderliche Genauigkeitsniveau der Einflussfaktoren wider. Dreiländertagung der DGPF, der OVG und der SGPF in Wien, Österreich – Publikationen der DGPF, Band 28, 2019

Anwendung des Messkonzeptes 2.3

Die Untersuchungen werden durch die Anwendung des Auswertekonzeptes abgeschlossen. Hierfür werden beide Orientierungs- und Matchingverfahren miteinander kombiniert und an einer Schweißnaht erprobt, die als CAD-Modell vorliegt. Um die Qualität des Systems realistisch zu quantifizieren, wurde die fehlerfreie innere Orientierung der Kamera im Rahmen der vorgegebenen Standardabweichungen verrauscht. Jedoch kann aufgrund des Leistungsspektrums der zugrundeliegenden Schweißnahtoberfläche erwartet werden, dass die Zuordnungsqualität gering ausfällt, weshalb neben einer realitätsnahen zusätzlich eine optimierte Textur verwendet wird, die in Abbildung 4 zu sehen ist. Diese beruht auf den Erkenntnissen der Voruntersuchungen Texturmerkmalen, sodass aufgrund des Leistungsspektrums von den eine hohe Zuordnungsqualität erwartet werden kann. Die resultierenden Punktwolken werden anschließend mit dem vorliegenden CAD-Modell verglichen und bewertet.



Optimierte Textur

Abb. 4: Texturen einer realen und der synthetischen Schweißnaht

3 Ergebnisse

Matchingverfahren 3.1

Die Untersuchung der Matchingverfahren zeigt, dass mit beiden Methoden ein hohes Rekonstruktionsniveau der Testfelder erreicht werden kann, das allerdings von den Parametereinstellungen abhängig ist. So können horizontale und vertikale Strukturen von bis zu 80 µm mit der geforderten Genauigkeit rekonstruiert werden.



Abb. 5: Rekonstruktionsergebnis des Testfeldes mit horizontaler Variation mittels SGM und einer Fenstergröße von 3x3 Pixel. Links: Rekonstruiertes Testfeld. Rechts: Dazugehöriger Querschnitt des Testfeldradius

Zudem kann beobachtet werden, dass das Rauschen beim LSM vergleichsweise höher ausfällt als beim SGM. Dies ist auf die fehlende Berücksichtigung der Nachbarschaftsbeziehungen zurückzuführen. In Abbildung 5 ist die Rekonstruktion des Testfeldes zu erkennen, wenn eine Fenstergröße von 3x3 Pixel verwendet wird und die Strafterme P_1 und P_2 jeweils dem sechsfachen der quadrierten Fenstergröße entsprechen. Es kann beobachtet werden, dass die Rekonstruktion der äußeren Ringe eine hohe Qualität aufweisen. Jedoch ist anhand des abgebildeten Querschnittes zu erkennen, dass ebene Flächen durch das Messrauschen gestört und feine Strukturen geglättet werden. In Abbildung 6 wird dagegen eine Fenstergröße von 15x15 Pixel verwendet, wobei die Strafterme ebenfalls dem sechsfachen der quadrierten Fenstergröße entsprechen. Auch bei dieser Rekonstruktion weisen die äußeren Ringe eine hohe Qualität auf, wobei die Flanken der Ringe durch unterschiedliche Plateaus repräsentiert werden. Zudem kann beobachtet werden, dass deutlich mehr Ringe nicht korrekt rekonstruiert werden. Dieses ist auf den Glättungseffekt der Fenstergröße zurückzuführen, der gleichzeitig das Zuordnungsrauschen dämpft. Besonders anhand der Querschnitte ist die Glättung feiner Strukturen zu erkennen. Dieser Effekt wird beim SGM zusätzlich verstärkt, wenn der Strafterm P_1 deutlich größer als die angesetzte quadrierte Fenstergröße ist. Stattdessen hat der Strafterm P2 kaum Einfluss. Jedoch führen Fenstergrößen größer als 15x15 Pixel in Kombination mit Straftermen, die Werte größer als das zehnfache der quadrierten Fenstergröße aufweisen, dazu, dass die Wahrscheinlichkeit von Fehlzuordnungen zunimmt.



Abb. 6: Rekonstruktionsergebnis des Testfeldes mit horizontaler Variation mittels SGM und einer Fenstergröße von 15x15 Pixel. Links: Rekonstruiertes Testfeld. Rechts: Dazugehöriger Querschnitt des Testfeldradius

3.2 Orientierungsverfahren

Bei den Untersuchungen der Orientierungsverfahren zeigt sich, dass mit beiden Verfahren das angestrebte Genauigkeitsniveau mit kurzen Basislängen nicht erreicht wird. Besonders die hohen Genauigkeitsanforderungen der Translation in X-Richtung (Basislänge) und der Rotation um die Y-Achse (Konvergenzwinkel) führen dazu, dass Basislängen ab 20 mm für die weiteren Berechnungsschritte verwendet werden sollten. Ein Vergleich der zulässigen Standardabweichungen der inneren Orientierung zum realen Kalibrierdatensatz hat ergeben, dass vermutlich die Standardabweichung der radial-symmetrischen Verzeichnung zu den kritischen Abweichungen führt. In Abbildung 7 ist exemplarisch das Genauigkeitsniveau des Konvergenzwinkels von den Bildern des Bildverbandes dargestellt, wenn die Orientierung anhand der beiden Methoden berechnet wird. Hierbei kann bei beiden Verfahren beobachtet werden, dass nur die mittleren Bilder des Bildstreifens die erforderliche Orientierungsgenauigkeit erreichen. Diese auftretende Systematik ist vermutlich auf die Passpunktverteilung zurückzuführen, da die Messmarken bei den Randbildern nicht auf den gesamten Aufnahmebereich homogen verteilt sind. Da eine homogene Passpunktverteilung in der realen Anwendung nur bedingt umsetzbar ist, sollten weitere Möglichkeiten wie der Einsatz eines Linearschlitten oder von ICP-Verfahren näher analysiert werden.



Abb. 7: Resultierende Orientierungsabweichungen und die zulässige Variation der Rotation um Y aus der MCS.

3.3 Anwendung an der Schweißnaht

Bei der Anwendung der untersuchten Orientierungs- und Matchingverfahren an einer synthetischen Schweißnaht konnten kaum Unterschiede zwischen den resultierenden Punktwolken werden. Jedoch hat die verwendete Textur einen Einfluss auf festgestellt das Rekonstruktionsergebnis. So ist in Abbildung 8 (links) zum einen die resultierende Punktwolke der Schweißnaht zu erkennen, wenn die realitätsnahe Textur verwendet wird. Des Weiteren ist in Abbildung 8 (rechts) die rekonstruierte Schweißnaht dargestellt, wenn die optimierte Textur verwendet wird. Während aus der optimierten Textur eine dichte homogene Punktwolke resultiert, sind bei der realitätsnahen Textur nur einzelne zusammenhängende Bereiche zu erkennen. Zudem weist sie größere Abweichungen zu der Referenz auf. Bei der optimierten Textur liegen die Abweichungen weitestgehend unterhalb von 30 µm und erreichen somit das geforderte Genauigkeitsniveau. Jedoch ist das Aufbringen einer optimierten Textur aufgrund von speziellen Umweltbedingungen und den hohen Genauigkeits- und Auflösungsanforderungen schwierig, weshalb die reale Oberflächentextur verwendet werden sollte. Um dennoch eine korrekte und vollständige Punktwolke zu erreichen, sollte die Ähnlichkeit zwischen den Bildern des Stereopaares erhöht werden, was vermutlich nur durch eine verkürzte Basis zu realisieren ist. Damit steigt jedoch die Empfindlichkeit des Messsystems gegenüber Orientierungsunsicherheiten. Durch die Ergebnisse dieser Arbeit kann somit weiterhin zu einer strukturierten Beleuchtung geraten werden, die eine künstliche Textur erzeugt. Diese minimiert den Einfluss der unterschiedlichen Perspektive, wodurch eine große Basis und eine hohe Zuordnung der künstlichen Textur, dem strukturierten Licht, vermutlich ermöglicht wird.



Abb. 8: Resultierende Punktwolken mittels SGM und Bündelausgleichung. Alle Angaben in Millimeter

4 Fazit & Ausblick

Bei der in diesem Beitrag diskutierten Masterarbeit wurden zwei Matching- und Orientierungsverfahren im Hinblick auf die hochgenaue Erfassung von Schweißnahtoberflächen untersucht. Hierfür wurden aufgrund des hohen Kosten- und Zeitaufwandes zur Erzeugung realer Daten, vor allem aber aufgrund der individuellen Evaluationsmöglichkeit einzelner Systemparameter, ausschließlich synthetische Bilddaten verwendet. Durch das Konzept der Evaluierung wird somit die Trennung des Einflusses einzelner Parameter ermöglicht. So zeigte sich, dass die Konfiguration des Bildverbandes sowie die Kalibrierung der inneren Orientierung optimiert werden sollten, damit die gesamte Schweißnaht hochgenau rekonstruiert wird. Eine Möglichkeit wäre das Aufnehmen von weiteren überlappenden Bildstreifen, wodurch vermutlich die Stabilität des gesamten Bildverbandes steigt und weitere Bilder für die hochgenaue Rekonstruktion verwendet werden können. Die Matchingverfahren SGM und LSM können prinzipiell für die hochgenaue Rekonstruktion von feinen Strukturen verwendet werden. Der Einfluss der Textur ist offensichtlich, jedoch auf Basis der Untersuchungen nur schwer zu quantifizieren, weshalb weitere Analysen notwendig sind. Bestätigen sich die Ergebnisse dieses Beitrages, können weitere Matchingverfahren wie das objektbasierte Semi-Global Matching (OSGM) nach BETHMANN & LUHMANN (2015) oder Multi-View Stereoverfahren (MVS) untersucht werden. Zudem können für reale Situationen weitere Parameter angebracht werden, um beispielsweise die Abschattung von Störobjekten oder radiometrische Besonderheiten zu simulieren. Des Weiteren sollten strukturierte Beleuchtungen in Betracht gezogen werden, um den Kompromiss aus Basislänge und Zuordnungsfähigkeit messtechnisch optimiert zu gestalten. Zusätzlich erlaubt der Aufbau der synthetischen Testfelder eine objektive, reproduzierbare und umfassende Evaluation ganzer Messsysteme für hochgenaue Anwendungen.

Dreiländertagung der DGPF, der OVG und der SGPF in Wien, Österreich – Publikationen der DGPF, Band 28, 2019

5 Literaturverzeichnis

- BETHMANN, F. & LUHMANN, T., 2015. Semi-Global Matching in Object Space. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **40**(3/W2), 23-30.
- EKKEL, T., SCHMIK, J., LUHMANN, T. & HASTEDT, H., 2015: Precise laser-based optical 3D Measurement of welding seams under water. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **40**(5/W5), 117-122.
- GRÜN, A., 1985: Adaptive kleinste Quadrate Korrelation und geometrische Zusatzinformationen. Vermessung, Photogrammetrie, Kulturtechnik, **9**, 309-312.
- HASTEDT, H., 2004: Monte-Carlo Simulation in Close Range Photogrammetry. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **34**, 18-23.
- HIRSCHMÜLLER, H., 2005: Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, **2**, 437-444.
- NEILL, A. M., 2016: The application of 3D reconstruction by stereo vision for the purpose of assessing weld quality. PhD Thesis, Department of Mechanical Engineering, Colorado School of Mines.
- POHLE, R., 2004. Computerunterstützte Bildanalyse zur Auswertung medizinischer Bilddaten. Habilitationsschrift, Fakultät für Informatik der Otto-von-Guericke-Universität, Magdeburg.
- SCHARSTEIN, D., HIRSCHMÜLLER, H., KITAJIMA, Y., KRATHWOHL, G., NESIC, N., WANG, X. & WESTLING, P., 2014: High-Resolution Stereo Datasets with Subpixel-Accurate Ground Truth. German Conference on Pattern Recognition, 31-42.
- SCHARSTEIN, D. & SZELISKI, R., 2002: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision, 47(1/2/3), 7-42.
- SCHARSTEIN, D. & SZELISKI, R., 2004: High-accuracy stereo depth maps using structured light. IEEE Comouter Society Confernce on Computer Vision and Pattern Recognition, 1, 195-202.
- SCHNEIDER, C.-T., 1991. Objektgestützte Mehrbildzuordnung. Dissertation, Deutsche Geodätische Kommission, München, Reihe C, Nr. 506.

AUTORENVERZEICHNIS

Α

Adler, Petra Anders, Karl-Heinrich	510 310
AUFLEGER, MARKUS	7,537
	702
	044
BACH, HEIKE	214 188
BANNEHR LUTZ	100
BARALIN, LOTZ	702
BARAN, RAMONA	
BAUER, THOMAS	256
BAUMGARTNER, KATHARINA	537
BECKER, ROLF	616
Benger, Wener	418
BLASER, STEFAN	458
BLÖTHE, JAN	675
BOERNER, RICHARD	548
BRENNER, CLAUS	68
BRUGGISSER, MORITZ	498
BRUNN, ANSGAR	J, 054
C	
CARVALHAIS, NUNO	203
CHEN, LIN	3,375
	21
	472
	1
	000
	303
DISCHER, SOREN	. 104 /18
DOBLER, WOLFGANG	226
DÖLLNER JÜRGEN	154
Dorigo. Wouter	203
DORNINGER, PETER	606
DOROZYNSKI, MAREIKE	387
F	
ERIACHER MELANIE	310
Essl. Laura	288
Eysn, Lothar	15
F	
• Falkner Johannes	15
FORKEL MATTHIAS	
,	

G

GANZ, SELINA	510
GINZLER, CHRISTIAN	
GLÄßER, CORNELIA	
GORKOVCHUK. DENYS	
GRAF, LUKAS	

Η

n	
HAALA, NORBERT	472, 692
HANEL, ALEXANDER	54
Hastedt, Heidi	27
HAVEL, PATRICK	472
HEIPKE, CHRISTIAN	363, 375
HODAM, HENRYK	
Hoegner, Ludwig	548
Hollaus, Markus	498
HOLZER, NICOLAI	188
HÜBNER, PATRICK	44
HUPFER, MICHAEL	558
I	
Immitzer, Markus	256, 623
INNERHOFER, DANIEL	347
J	
JÄGER, REINER	
JEPPING, ANNIKA	27
JORDAN, DENIS	647
Jung, András	196
JUNG, TÍMEA	196
JÜRGENS, CARSTEN	241
к	
Käber, Yannek	510
KAHMEN, OLIVER	
KAMPEN, MAX	623
Kang, Junhua	
KAPAJ, ARMAND	1
Karel, Wilfried	123
KAUFMANN, VIKTOR	572, 597
KAUSCH, ISABELLA	
Keller, Boglárka	
KLAR, ROBERT	537
KLINGLER, ANDREAS	301
KNÖTTNER, JULIA	654
Kölle, Michael	
König, Simon	675
Kovács, László	
KRAUSE, STUART	167

L

LACK, NATALIE	
LADSTÄDTER, RICHARD	
LADSTÄTTER, STEFAN	175
LANDGRAF, STEVEN	
LAUPHEIMER, DOMINIK	472, 692
LEDERBAUER, STEFAN	623
LEHNER, HUBERT	443
LEIMER, WOLFGANG	418
LESKOVAR, JUTTA	
LINDNER, CLAUDIA	
LÖCKER, KLAUS	
LOGHIN, ANA-MARIA	123
LUHMANN, THOMAS	
LULEY, PATRICK	

Μ

Maas, Hans-Gerd	318, 334,	407
MADER, DAVID	334,	407
MAGYAR, MARIANNA		196
MANDLBURGER, GOTTFRIED 318, 347,	443, 472,	663
MANSBERGER, REINFRIED		256
Marty, Mauro		400
MAYER, HEINZ		175
MAYR, PETER		310
MEYER, ADRIAN		647
Meyer, Theresa		. 83
MIKL, THOMAS		597
Müller, Christina		241
MULSOW, CHRISTIAN		318
MUND, JAN-PETER	167,	623
MUSTAFIĆ, SEAD		527
Ν		
NEBIKER, STEPHAN		458
NEUBAUER. WOLEGANG		234
NIFTIEDT. SIMON		711
•		
		044
		241
UTEPKA, JOHANNES		123
Ρ		
PAULUS, GERNOT		310
PÉTERFALVI, NÓRA		196
Peters, Torben		. 68
PFEIFER, NORBERT	123,	498
PFEIL, ISABELLA		606
PFENNIGBAUER, MARTIN		428
PFENNINGER, STEPHAN		. 54
PIECHL, THOMAS		226
Pöchtrager, Markus		123
Pötsch, Erich M.		301
PRZYBILLA, HEINZ-JÜRGEN		483

Q

R

RÁK, RENÁTA	
Ressl, Camillo	
RICHTER, KATJA	
RIENOW, ANDREAS	
RITTER, MARCEL	
ROSENBAUM, DOMINIK	
ROTTENSTEINER, FRANZ	

S

0	
SANDEN, TARU	288
SANDERS, TANJA	167
SCHARDT, MATHIAS	527
SCHAUMBERGER, ANDREAS	301
SCHMOHL, STEFAN	111
Schneeberger, Rudi	310
SCHOCH, ANNA	675
SCHROTT, LOTHAR	675
SCHULTZ, JOHANNES	241, 675
SCHULZE, MARTINA	558
Schwarz, Roland	428
SEIER, GERNOT	587
SELG, FABIAN	241
SILLINGER, FANNI	196
SÖRGEL, UWE	111
SPERLICH, MAXIMILIAN	686
SPIEGEL, ADELHEID	
STEINBACH, ECKEHARD	54
STEINBACHER, FRANK	
STILLA, UWE	54, 97, 548
STRER, MAXIMILIAN	167
Sudi, Prajwala	54
SUDMANNS, MARTIN	
Sulzer, Wolfgang	572, 587
SZALAY, KORNÉL	
т	
	004

Taylor, Timothy	234
TESSEMA, LEALEM SEYOUM	97
Teucher, Mike	282
THONFELD, FRANK	675
THONICKE, KIRSTEN	203
Tiede, Dirk	702
Товіая, Налк	214
TRINKS, IMMO	234
U	
Ulrich, Christoph	558

V

Vohland, Michael Vreugdenhil, Mariette Vuolo, Francesco	
W	
WAGNER, WOLFGANG	606
WALLNER, MARIO	234
Wang, Di	498
Waser, Lars T	400
WECHT, MATTHIAS	587
WEIDNER, UWE	138
WEINACKER, HOLGER	686
WEINMANN, MARTIN	44, 138
WEIRATHER, MIRA	640
Weiß, Robert	407
WESTFELD, PATRICK	334, 407
WICH, VALENTIN	270
WILKENING, JAN	
WINIWARTER, LUKAS	
WISLER, DOMINIK	458
WITTICH, DENNIS	
WOLF, JOHANNES	
WOLFE, JASON	
WURSTHORN, SVEN	44
X	
Υ	
Z	~
	375

PROGRAMMKOMITEE

Abmayr, Thomas Atzberger, Clement Bareth, Georg Brunn, Ansgar Gläßer, Cornelia Gold, Wolfgang Gülch, Eberhard Haala, Norbert Haunert, Jan-Henrik Heipke, Christian Höfle, Bernhard Immitzer, Markus Jacobsen, Karsten Jung, András Kaufmann, Viktor Kersten, Thomas Kohlhofer, Gerald Krauß, Herbert Krzystek, Peter Leidenmühler, Sarah Mandlburger, Gottfried Mansberger, Reinfried Nebiker, Stephan Otepka, Johannes Petzold, Bettina Przybilla, Heinz-Jürgen Ressl, Camillo Ressl, Christine Schneider, Danilo Weinmann, Martin Wolff, Kirsten