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

DGPF

Band 22 **2013**



Vorträge

33. Wissenschaftlich-Technische Jahrestagung der DGPF

27. Februar – 1. März 2013
in Freiburg i. B.

Dreiländertagung D - A - CH

ISSN 0942-2870

Eckhardt Seyfert, Hrsg.

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



Band 22 **2013**

Vorträge

33. Wissenschaftlich-Technische
Jahrestagung der DGPF

27. Februar – 1. März 2013
in Freiburg i. B.

Dreiländertagung D -A - CH

ISSN 0942-2870

Eckhardt Seyfert, Hrsg.

ISSN 0942-2870

Publikationen der
Deutschen Gesellschaft für Photogrammetrie, Fernerkundung
und Geoinformation (DGPF) e.V.
Band 22 550 S., Potsdam 2013
Hrsg.: Eckhardt Seyfert

© Deutsche Gesellschaft für Photogrammetrie, Fernerkundung
und Geoinformation (DGPF) e.V.
Oldenburg 2008

Zu beziehen durch:

Geschäftsstelle der DGPF
c/o EFTAS GmbH
Oststraße 2-18
D-48145 Münster
Tel.: (0251) 133 070, Fax: (0251) 133 0733, E-Mail: klaus.komp@eftas.com

Bearbeitung:

Landesvermessung und Geobasisinformation Brandenburg
Heinrich-Mann-Allee 103, 14473 Potsdam
Tel.: (0331) 8844 -229, Fax: (0331) 8844-126

VOWORT

Die Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation lädt anlässlich ihrer 33. Wissenschaftlich-Technischen Jahrestagung vom 27. Februar bis 1. März 2013 nach Freiburg im Breisgau ein.

Diese Einladung erfolgt auch im Namen der mitveranstaltenden Schwestergesellschaften aus Österreich und der Schweiz, der Österreichischen Gesellschaft für Vermessung und Geoinformation sowie der Schweizerischen Gesellschaft für Photogrammetrie und Fernerkundung. Wir haben für die Tagung als gemeinsames Motto gewählt:

DREILÄNDERTAGUNG D – A – CH

Während das Thema sich aus der kollegialen und vertrauensvollen Kooperation der transnationalen wissenschaftlichen Verbindung der drei Gesellschaften fast anbietet, wollen wir mit der Wahl des diesjährigen Standortes auch bewusst unseren Altpräsidenten und unser Ehrenmitglied *Prof. Dr. Gerhard Hildebrandt* eine tief empfundene Referenz erweisen und ins Gedächtnis rufen, dass er lange vor der offiziellen Freigabe des amerikanischen CIR-Filmes in Vorstudien die dringende Nutzbarkeit dieses Instrumentariums für forstwissenschaftliche und allgemein vegetationskundliche Zwecke untersucht und propagiert hat.
Die Fortführung des Instituts unter der erfolgreichen Leitung von Frau Prof. Dr. Barbara Koch und das Engagement ihrer Mitarbeiter haben den Vorstand der DGPF ermutigt, den beiden Schwestergesellschaften den Standort Freiburg für die Dreiländertagung vorzuschlagen. Die vielfältigen Tutorien aus Forschung und Entwicklung sowie das breite Spektrum der Fachvorträge bieten den Teilnehmern die Möglichkeit sich eingehend fortzubilden.

In diesem Jahr wurden 100 Beiträge angemeldet von denen 58 auf der Tagungs-CD enthalten sind. Insgesamt haben 130 Autoren an den Beiträgen mitgearbeitet. Einundzwanzig von ihnen haben zu dieser Jahrestagung mehrfach publiziert.

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

Lokaler Gastgeber ist das Institut für Fernerkundung und Landschaftsinformationssysteme (FeLis). Wir danken der Institutsleitung für die unkomplizierte und freundliche Bereitstellung der Räume und besonders Frau Kollegin Barbara Koch und ihrem Team für die gründliche und engagierte Vorbereitung der Tagung und das interessante und vielfältige Begleitprogramm.

Prof. Dr. rer. nat. Thomas Kolbe	Dipl.-Ing. Gert Steinkellner	Prof. Francois Gervaix
Präsident DGPF	Präsident OVG	Präsident SGPF

INHALTSVERZEICHNIS

AUSWERTUNG VON FERNERKUNDUNGSDATEN

ADLER, P. & PETERS, S.: Automatische Klassifizierung von Fichten in digitalen 4-Kanal Luftbildern. ……. 9

BACH, E., FRÖHLICH, B., WALDE, I., HESE, S., SCHMULLIUS, C. & DENZLER, J.: Landbedeckungsklassifikation von Satellitenaufnahmen unter Verwendung von Kontextinformation mit minimaler menschlicher Interaktion. …… 17

BAYER, S., MIRI, M. & BUCHER, T.: Automatisierte Einzelbaumerfassung auf Basis von hochaufgelösten TrueOrtho-Mosaiken und photogrammetrisch abgeleiteten Oberflächenmodellen. …… 26

BECHTEL, B., ZAKŠEK, K. & BÖHNER, J.: Hochauflösende Datensätze urbaner Oberflächentemperaturen zur Beobachtung der städtischen Wärmeinsel. … 36

BOESCH, R. & GINZLER, C.: Automatische Berechnung einer nationalen Gehölzmaske. …… 46

DRAUSCHKE, M. & REIDELSTÜRZ, P.: Klassifikation von Fernerkundungsdaten zur Früherkennung von Borkenkäferschäden im Bayerischen Wald. …… 54

FANG, L., MAKSYMIUK, O., SCHMITT, M. & STILLA, U.: Improvement of Motion Estimation of the Taku Glacier using Spaceborne SAR Images. …… 62

GINZLER, C., MARTY, M. & BÜHLER, Y.: Grossflächige hochaufgelöste Schneehöhenkarten aus digitalen Stereoluftbildern. …… 71

GIOVENCO, M.M. & GIAQUINTO, G.: Comparison of Image Quality in Optical Radar Fused Image Sets. …… 79

GRENZDÖRFFER, G. & BOGDANOV, S.: UAS-basierte automatisierte Vogelzählung am Beispiel einer Sturmmöwenkolonie der Vogelschutzinsel Langenwerder. 89

HERDA, H. & BREUER, M.: Different Ways to Process UAV Imagery – Results of Comparative Benchmark Tests. …… 99

KATTENBORN, T.: Forest Change Assessment and Corresponding Driver Analysis in the Magdalena Department, Colombia (1985-2010). …… 108

MANTHEY, K.: About the impact of CCSDS image data compression on Image Quality. …… 116

MICHAELIS, R. & HESE, S.: Objektbasierte spektrale Analyse und Klassifikation von ölkontaminierter Landoberfläche in Westsibirien mit hochaufgelösten RapidEye Daten. …… 127

MIRI, M., BAYER, S. & BUCHER, T.: Evaluation of high resolution digital surface models for single tree extraction approaches in mixed forests. …… 137

NETZBAND, M.: Change detection by means of multitemporal satellite imagery for the Indian Mega-City Hyderabad. …… 146

NOVACK, T., MAKSYMIUK, O. & STILLA, U.: A Concept for Guiding the Learning of Conditional Random Fields for the Classification of Urban Areas in SAR Images. …… 152

SAČKOV, I. & KARDOŠ, M.: Point Clouds from Airborne Laser Scanning and Aerial Images for Forestry Application. …… 160

SAUERBIER, M.: 3D-Gebäudeaufnahme mittels UAV-basierter Luftbilder. …… 169

STERN, O.: Abbildende und nichtabbildende Geländespektrometrie zur Untersuchung von Stressphänomenen an Buchenpflanzen. …… 176

STRAUB, C., STEPPER, C., SEITZ, R. & WASER, L.T.: Möglichkeiten einer stratifizierten Schätzung von forstlichen Kenngrößen unter Verwendung von amtlichen Stereo-Luftbildern, Orthophotos und Laserscannerdaten. … 186

THONFELD, F. & MENZ, G.: Entwicklung eines robusten Veränderungsdetektionsverfahrens für räumlich hoch auflösende Satellitendaten. 194

BILDANALYSE UND BILDVERSTEHEN

BRANDENBURGER, W., DRAUSCHKE, M., NGUATEM, W. & MAYER, H.: Detektion von Fensterverdachungen und Gesims in Fassadenbildern und Punktwolken. .. 202
GRIWODZ, U. & BELLINGHOFF, M.: Traditionelle Luftbilder in 3D. 210
IWASZCZUK, D., HOEGNER, L. & STILLA, U.: Zuordnung von 3D Gebäudemodellen und IR-Videosequenzen mit Linienverfolgung. 220
MATTI, E. & NEBIKER, S.: Geometrie- und farbbasierte Punktwolkenklassifizierung von komplexen urbanen Laserscanningszenen mittels überwachter SOM-Klassifikation. 228
TRISTRAM, S. & REULKE, R.: 3D-Bildanalyse mit Scene Flow-Algorithmen. 237

FERNERKUNDUNG IN DER GEOLOGIE

BEHLING, R., ROESSNER, S., WETZEL, H.-U. & KAUFMANN, H.: Automated Multi-temporal Landslide Identification in Kyrgyzstan Using Satellite Remote Sensing Data. 244
FRANZKE, H. J. & WETZEL, H.-U.: Schlussfolgerungen aus der Anwendung der Geo-Fernerkundung in geowissenschaftlichen Projekten in SW-Deutschland. 256
WULF, G., POELCHAU, M. & KENKMANN, T.: Structural analysis of Martian impact craters using several remote sensing data. 263

HYPERSPEKTRALE FERNERKUNDUNG

KNAUER, U. & SEIFFERT, U.: An Approach for Automated Registration of Hyperspectral Images for Boresight Calibration. 271

OPTISCHE 3D-MESSTECHNIK

DARR, T., GÖTZ, C., TUTTAS, S., BECKER, T. & STILLA, U.: Genauigkeit der inversen Navigation bei hoher Passpunktsichtbarkeit zur Absolutpositionierung eines geführten Messsystems. 279
ELTNER, A.: Konzeption einer Methode zur quantitativen Erfassung von Bodenerosion durch integrierte Aufnahme mit einer Mikrodrohne und einem terrestrischen Laserscanner. 289
HANEL, A., GÖTZ, C., TUTTAS, S., BECKER, T. & STILLA, U.: Orientierungs-genauigkeit eines Multi-Kamerasystems bei Roboterbewegungen in einer Fahrzeugkarosserie. 298
HASTEDT, H. & LUHMANN, T.: Untersuchungen zum Genauigkeitspotential eines hybriden Kamerasystems aus RIM- und RGB-Kamera. 308
RUEß, D., LUBER, A. & REULKE, R.: Epipolar Geometry of Heterogeneous Stereo Camera Systems. 316
SCHLICHTING, A., BRENNER, C. & SCHÖN, S.: Bewertung von inertialen Messsystemen mittels Laserscannern und bekannter Landmarken. 326

3D-Stadtmodelle

Döring, I.: 3D GDI zur Unterstützung der Standortbewertung von Windenergieanlagen (WEA). 337
Eugster, H., Gottsmann, F., Käser, C., Nebiker, S. & Koch, R.: 3DRoadTV-Service – Georeferenzierter 3D-Bilddatendienst für die Infrastrukturpflege und -verwaltung. 344
Linkiewicz, M.: Extraktion von senkrechten Fassadenebenen aus 3D-Punktwolken von Schrägluftbildern. 354
Tuttas, S. & Stilla, U.: Rekonstruktion von Fenstern aus Schrägsicht-ALS Punktwolken zur Anreicherung von Gebäudemodellen. 363
Wieden, A. & Linkiewicz, M.: True- Oblique- Orthomosaik aus Schrägluftbildern zur Extraktion von 3D-Geoinformationen. 373

Radarfernerkundung und Flugzeuglaserscanning

Leiterer, R., Morsdorf, F. & Schaepman, M. E.: Characterization of forest understory using multi-temporal full-waveform airborne laser scanning. 384
Maksymiuk, O., Brenner, A. & Stilla, U.: Detection of Stationary Vehicles in Airborne Decimeter Resolution SAR Intensity Images using Morphological Attribute Filters. 391
Mücke, W., Schroiff, A., hollaus, M. & Pfeifer, N.: Erkennen von liegendem Totholz in full-waveform Airborne Laserscanning Daten. 401
Saur, G., Heinze, N. & Krüger, W.: Video-like exploitation of MiSAR sequences from small UAVs. 411
Schlund, M., von Poncet, F., Kuntz, S., Hoekman, D. H. & Schmullius, C.: TanDEM-X als Datengrundlage für Referenzkartierungen im Kontext von REDD. 418
Schmitt, M., Brück, A., Schönberger, J. & Stilla, U.: Potential of Airborne Single-Pass Millimeterwave InSAR Data for Individual Tree Recognition. ... 426
Yao, W., Krzystek, P. & Heurich, M.: A study on the detection of understory regenerations in temperate forest areas using airborne full-waveform LiDAR. 437

Sensoren und Plattformen

Cramer, M., Haala, N., Rothermel, M., Leinss, B. & Fritsch, D.: UAV-gestützte Datenerfassung für Anwendungen der Landesvermessung – das Hessigheim-Projekt. 450
Fritsch, D., Grimm, A., Kremer, J., Rothermel, M. & Wenzel, K.: Bilddatenerfassung mit einem Gyrocopter – Erste Erfahrungen zur „Photogrammetrie nach Bedarf“. 470
Hoegner, L., Weinmann, M., Jutzi, B., Hinz, S. & Stilla, U.: Co-registration of Time-of-Flight (TOF) camera generated 3d point clouds and thermal infrared images (IR). 481
Jacobsen, K.: Erstellung großflächiger Höhenmodelle und Vergleich mit anderen Daten. 489
Nebiker, S., Kamer, K. & Eugster, H.: Integrierte Georeferenzierung von luft- und bodengestützten Stereobilddaten im Straßenbereich – Strategien und Ergebnisse. 499

RUB, M., FARKAS, E. & WIEDEMANN, A.: Dynamische Erfassung linearer Infrastrukturen unter und über der Erde. 509
WIEDEN, A. & STEBNER, K.: Referenzorientierung für Bilddaten aus Mehrkopfkamerasystemen. 518

STANDARDISIERUNG UND QUALITÄTSSICHERUNG

REULKE, R.: DIN 18714-8 − Anforderungen an die Bildqualität. 526
WALDHÄUSL, P., ABART, G., HANKE, K., HIERMANSEDER, M., KÖNIG, H., MANSBERGER, R., NAVRATIL, G., SCHENNACH, G. & TWAROCH, C.: Sind das landesweite Netzwerk von Grenzen und die Grenzsteine ein Weltkulturerbe? 534

AUTORENVERZEICHNIS 543

Automatische Klassifizierung von Fichten in digitalen 4-Kanal Luftbildern

PETRA ADLER[1] & SEBASTIAN PETERS[1]

Für die fernerkundungsgestützte Baumartenerkennung stehen den Forstverwaltungen häufig Luftbilder zur Verfügung. Die Erkennung der Fichte hat aufgrund ihrer Bedeutung als wirtschaftliche Hauptbaumart und ihrer Anfälligkeit gegenüber den Auswirkungen des Klimawandels hohe Priorität. Die Fichte unterscheidet sich von den übrigen Baumarten deutlich in ihrer Struktur. In dieser Studie wurde deshalb neben spektralen Parametern die Verwendung von Texturparametern zur Detektion der Fichte getestet. Die Klassifizierung erfolgte objektbasiert. Nach der Segmentierung der Nadelbaumflächen wurde ein Generalisiertes Lineares Modell benutzt, um Fichtenobjekte zu erkennen. In allen drei untersuchten natürlichen Altersklassen konnte die Kombination von Textur und spektraler Information jeweils eine bessere Klassifikationsgenauigkeit im Vergleich zur rein spektralen Klassifikation erzielen.

1 Einleitung

Die Detektion der Fichte auf Landesebene ist eine Kernfrage für viele forstliche und ökologische Fragestellungen in Baden-Württemberg. Die automatische Differenzierung von Baumarten aus digitalen Luftbilddaten ist eine bereits seit längerem bestehende Fragestellung. Mit großer Zuverlässigkeit ist operational die Trennung der Baumartengruppen Laub-/Nadelbaum möglich. Die Trennung innerhalb dieser Gruppen ist noch auf einzelne Untersuchungen beschränkt und erreicht nur zum Teil befriedigende Genauigkeiten (WASER *et al.* 2011, HEINZEL & KOCH 2012). Die Bedeutung der Fichte als wirtschaftliche Hauptbaumart und ihre Anfälligkeit gegenüber den Auswirkungen des Klimawandels begründet den Bedarf einer möglichst vollständigen und räumlich expliziten Erfassung (SEITZ *et al*. 2010, KÖLLING *et al.* 2009.). In der vorliegenden Untersuchung wird deshalb die Trennbarkeit der Fichte zu anderen Nadelbäumen auf der Basis digitaler Luftbilder von Standardbefliegungen des Landesamtes für Geoinformation und Landentwicklung Baden-Württemberg (LGL) untersucht. Da die Fichte sich in ihrer Struktur deutlich von anderen Nadelbaumarten unterscheidet, werden neben spektralen Eigenschaften auch Texturparameter untersucht.

2 Studiengebiet

Das Studiengebiet befindet sich im bewegten Gelände im submontanen Bereich des Stadtwaldes Freiburg und umfasst eine Fläche von 33,5 ha. Das Alter der zehn Bestände, die laut Forsteinrichtungsdaten mit Stichtag 1.1.2011 innerhalb dieses Gebietes ausgewiesen sind, reicht von 0-130 Jahren.

1) Forstliche Versuchs- und Forschungsanstalt Baden-Württemberg, Wonnhaldestraße 4, 79100 Freiburg; E-Mail: Petra.Adler@Forst.bwl.de; Sebastian.Peters@Forst.bwl.de

Der Fichtenanteil liegt auf der gesamten Fläche bei 48 %, der Nadelbaumanteil bei 74 %. Unter den Nadelbaumarten kommen neben der Fichte noch Douglasie, Tanne und Lärche mit nennenswerten Anteilen vor.

3 Material & Methoden

Für die Untersuchung wurden digitale 4-Kanal Stereoluftbilder mit 20 cm Bodenauflösung und 8 bit radiometrischer Auflösung verwendet. Die Aufnahme erfolgte am 14. Juni 2009 mit einer UltraCamXP. Aus den Stereoluftbildern konnte unter Verwendung des Moduls *eATE* der Toolbox LPS in ERDAS IMAGINE via Image-Matching ein digitales Oberflächenmodell (DOM) abgeleitet werden. Die durchschnittliche Punktdichte der erzeugten 3D-Punktwolke entsprach 5,6 Punkte/m². Das DOM wurde über einen Medianfilter geglättet und anschließend für die Berechnung lagegenauer True-Orthophotos verwendet. Diese wurden zu einem Orthophotomosaik zusammengefügt.

3.1 Segmentierung

Die Segmentierung des Orthophotomosaiks erfolgte in mehreren iterativen Teilschritten. In einem ersten, pixelbasierten Schritt wurden mittels Bayes'scher Klassifikation anhand aller vier spektralen Kanäle Laub-Nadelholz Wahrscheinlichkeiten erzeugt. Durch ein Grenzwert- und Klumpungsverfahren wurden diese zu Nadelbaumobjekten zusammengefasst. Kleinere Lücken innerhalb der Nadelbaumobjekte wurden über Filteralgorithmen bereinigt. Die resultierenden Objekte variierten in ihrer Größe von Einzelkronen bis zu Baumgruppen.
In einem zweiten Schritt wurden die identifizierten Nadelbaumobjekte weiter unterteilt. Durch die Subtraktion eines digitalen Geländemodells (DGM) vom DOM wurde ein normalisiertes Oberflächenmodell gerechnet (nDSM). Über einen Wasserscheide-Algorithmus können aus dem invertierten nDSM Baumkronensegemente abgegrenzt werden (STRAUB 2003). Die Objekte aus der Wasserscheidentransformation waren dadurch gekennzeichnet, dass sie die gesamte Bildfläche inklusive der Schattenflächen abdeckten und weiterhin mehrere Bäume umfassen konnten. Diese Objekte der Wasserscheidentransformation wurden deshalb mit den Nadelbaumobjekten der Pixelklassifikation verschnitten. Anschließend wurden die lokalen Maxima des nDSM berechnet. Die lokalen Maxima entsprachen idealerweise dem jeweils höchsten Punkt innerhalb einer Baumkrone. Objekte des bis dato erfolgten Segmentierungsprozesses, die mehr als ein lokales Maximum beinhalteten, wurden anhand einer Euklidischen Zuordnung weiter unterteilt. In einem letzten Schritt wurden noch alle Flächen mit einer Mittelhöhe < 2 m und Sliverpolygone eliminiert. Abbildung 1 verdeutlicht den Segmentierungsprozess.

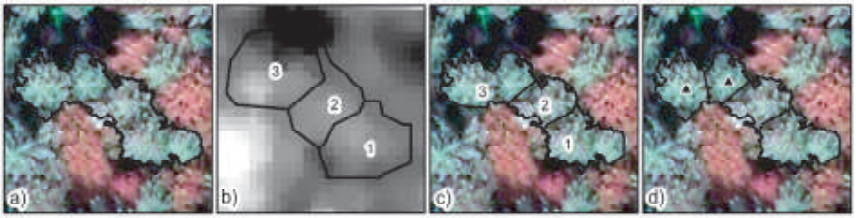


Abb. 1: schrittweise Segmentierung: a) Nadelbaumfläche nach Pixelklassifikation (Hintergrund Luftbild) b) Polygone nach Wasserscheidentransformation (Hintergrund nDSM) c) Segmente nach Verschneidung der Nadelbaumflächen mit Wasserscheidenpolygonen d) weitere Trennung von Segment 3 anhand lokaler Maxima (Dreiecke).

3.2 Klassifikation

Abhängig von ihrem Alter unterscheidet sich eine Baumart in ihren spektralen Eigenschaften und der Textur. Deshalb wurden nach der erfolgten Segmentierung alle Objekte in Anlehnung an einen Interpretationsschlüssel der ARBEITSGRUPPE FORSTLICHER LUFTBILDINTERPRETEN (AFL) (1998) aufgrund ihrer Höhenwerte in die drei natürlichen Altersklassen *Dickung*, *Stangenholz* und *Baumholz* eingeteilt. Dazu wurde das 0,9-Quantil ($Q_{0,9}$) aller nDSM Höhenwerte eines Segments verwendet. Alle Segmente mit einem $Q_{0,9} < 8$ m wurden in die Altersklasse *Dickung* gruppiert, Werte von 8 bis 20 m entsprachen der Altersklasse *Stangenholz* und ein Höhenwert > 20 m der Klasse *Baumholz* (Tab. 2).

3.2.1 Variablen

Als spektrale Variablen für die Klassifikation wurden neben den vier einzelnen Kanälen (R, G, B, NIR) sechs Verhältnisse verwendet (z.B. Rot/RGB bzw. NIR/CIR). Als Texturvariablen wurden zehn Texturparameter anhand einer Grauwertematrix berechnet (HARALICK 1979). Die Grauwertematrix (GLCM) wurde für jeden der vier Kanäle mit jeweils drei verschiedenen Fenstergrößen (3x3, 5x5 und 7x7) berechnet. Aus den 4 möglichen Berechnungsrichtungen der Texturen wurde jeweils ein Mittelwert gebildet. Für jedes Segment wurden *Minimum*, *Maximum*, *Spannweite*, *Mittelwert* und *Standardabweichung* aller Variablen ermittelt und als Parameter in der Modellbildung getestet. Tabelle 1 gibt einen Überblick über alle erzeugten Variablen.

Tab. 1: Übersicht über alle erzeugten Variablen zur Klassifikation der Segmente.

	Variable	
Spektrale Variablen	Einzelne Kanäle	R,G,B, NIR
	Verhältnisse	R/RGB, G/RGB, B/RGB R/CIR, G/CIR, NIR/CIR
Texturvariablen	Sum Average (SA), Entropy (ENT), Difference Entropy (DE), Sum Entropy (SE), Variance (VAR), Difference Variance (DV), Sum Variance (SV), Angular Second Moment (Uniformity) (ASM), Inverse Difference Moment (Homogeneity) (IDM), Contrast (CON)	

3.2.2 Generalisiertes Lineares Modell

Die Zielvariable Fichte/ sonstige Nadelbäume entspricht einer Binomialverteilung. Deshalb wurde die Klassifizierung zur Detektion der Fichte anhand eines Generalisierten Linearen Modells (GLM) basierend auf einer logistischen Regression durchgeführt (VENABLES & RIPLEY 2002, ORTIZ *et al.* 2012). Modellbildung, Variablenselektion und statistische Genauigkeitsanalysen wurden in einer *R* Umgebung unternommen.

Um eine Überanpassung der Modelle zu vermeiden, müssen wenige, aber robuste Variablen ausgewählt werden. Die Variablenselektion erfolgte schrittweise mit der Funktion *stepAIC*. Dazu wurde jeweils ein Modell mit weniger als 20 zufällig angeordneten Variablen gefittet und aus diesem die Variablen ausgesucht, die am meisten zur Verringerung des AIC (Akaike's Information Criterion; VENABLES & RIPLEY 2002) beitrugen. Mit allen ausgewählten Variablen wurde ein neues Modell gefittet und eine erneute Variablenselektion unternommen. Die danach verbliebenen Variablen wurden auf Korrelationen untersucht und weiter reduziert.

Die Qualität der Modelle wurde anhand von Konfusionsmatrizen und Kappa-Koeffizienten nach einer leave-one-out-Kreuzvalidierung (LOOCV) verglichen. Als Trainingssegmente für die Modellbildung dienten 301 stereoskopisch auf ihre Baumart interpretierte Segmente. Diese wurden zufällig ausgewählt und durch manuell ausgewählte Segmente ergänzt, um eine genügende Anzahl an Trainingskandidaten zu haben. Als Trainingsobjekte kamen nur Segmente zur Verwendung, die einem sicher interpretiertem Individuum entsprachen und keine Mischsegmente waren.

4 Ergebnisse & Diskussion

4.1 Segmentierung

Die Qualität der Segmentierung steht im Zusammenhang mit der Objekthöhe. Einzelbaumscharfe Segmente konnten in der natürlichen Altersklasse *Dickung* nur selten geschaffen werden. Ausnahmen bildeten einzelne Nadelbäume, die von Laubbäumen umgeben waren (und damit spektral abgrenzbar waren). Weniger dichte Aufwuchsverhältnisse ermöglichten eine Einzelbaumerkennung aufgrund der Topographie des DOMs. In dieser Gruppe gab es bedingt durch die Qualität des DOM aus Luftbildern auch viele Segmente, die substantielle Bodenflächen beinhalteten. Diese konnten weder spektral noch aufgrund ihrer mittleren Höhe als solche erkannt werden.

Als trade-off eines geringen Auslassungsfehlers bei der Bestimmung von Nadelbaumflächen in der Pixelklassifikation entstanden einige Laubbaumsegmente, die i.d.R. aus beschatteten Kronen und Kronenrändern bestanden. Diese kamen ebenfalls vermehrt in den jüngeren Altersklassen vor. Gleiches galt für kronenübergreifende Mischsegmente. Ab einer Höhe von 15 – 25 m erreichten die Segmente eine Qualität, die fast ausnahmslos dem Ziel einer einzelbaumscharfen Delinierung entsprach.

Tab. 2: Zusammenfassung der Segmentierung und Trainingssegmente nach natürlichen Altersklassen.

Natürliche Altersklasse	Höhe ($Q_{0,9}$)	N Segmente	Ø Fläche (m²)	Ø Höhe ($Q_{0,9}$)	Trainingssegmente					
					Fi	Dgl	Ta	Lä	Lb	∅
Baumholz	> 20 m	2.472	32,3 (± 28,8)	27,5 (± 5,6)	49	40	28	34	0	151
Stangenholz	8-20 m	804	20,0 (± 20,1)	16,1 (± 3,1)	35	30	19	8	2	94
Dickung	< 8 m	116	8,1 (± 6,1)	4,8 (± 1,7)	0	15	14	1	26	56
∅		3.392	28,6	24,0	84	85	61	43	28	301

Fi = Fichte, Dgl = Douglasie, Ta = Tanne, Lä = Lärche, Lb = Laubbaum

4.2 Klassifizierung

4.2.1 Dickung

Nur 3,4 % aller Segmente fielen in die natürliche Altersklasse *Dickung*. Aufgrund der Qualität der Segmente konnten nur 20 eindeutige Nadelbaumsegmente identifiziert werden, darunter kein Fichtensegment. Die Klassifikation beschränkte sich daher auf eine Trennung der Segmente in Nadelbaum- bzw. Laubbaumobjekte. Die Klassifikationsgenauigkeit war gering. Die Gesamtgenauigkeit von 70,4 % bei Verwendung dreier spektraler Variablen (Mittelwerte des nahen Infrarots, des blauen Kanals und des Verhältnisses Grün/RGB) konnte durch die Einbindung einer Texturvariable (Mittelwert von ASM/blau/7x7) leicht auf 75,9 % verbessert werden (Tab. 3). Die Kappa Koeffizienten lagen nach LANDIS & KOCH (1977) mit 0,4 bzw. 0,51 im moderaten Bereich. Die Produzentengenauigkeit für Nadelbaumsegmente lag unter 70%. Bei der Bewertung dieser Ergebnisse ist zu berücksichtigen, dass hier solche Segmente klassifiziert wurden, die nach pixelbasierter Bayes'scher Klassifikation als Nadelbaumsegmente eingestuft wurden.

Tab. 3: Konfusionsmatrix des besten Modells in der Altersklasse *Dickung* (nach LOOCV).

Vorhergesagt (Klassifikation)	Beobachtet (Referenz)		
	LB	NB	UA (%)
LB	24	8	75,0
NB	5	17	77,3
PA (%)	82,8	68,0	75,9

PA Produzentengenauigkeit (producer's accuracy), UA Nutzergenauigkeit (user's accuracy), Gesamtgenauigkeit = untere rechte Zelle

4.2.2 Stangenholz

Knapp ein Viertel der Segmente wurde der natürlichen Altersklasse *Stangenholz* zugeordnet. In dieser wurde eine Unterscheidung der Fichte von sonstigen Nadelbäumen anhand von 92 Trainingssegmenten vorgenommen. Laubbaumsegmente wurden nicht berücksichtigt. Das beste Modell mit ausschließlich spektralen Variablen konnte eine Gesamtgenauigkeit von 76,7 % bei einem Kappa von 0,49 erreichen. Die Produzentengenauigkeit, die den Auslassungsfehler bei der

Detektion der Fichte beschreibt, lag nur bei 62,9 %. In der Kombination der drei ausgewählten spektralen Variablen mit Texturvariablen (Mittelwert von Rot/5x5/CON und Minimum von Blau/5x5/SE) konnte die Produzentengenauigkeit bei der Erkennung der Fichte auf 77,1 % angehoben werden. Die Gesamtgenauigkeit stieg auf 84,4 %, der Kappa Wert auf 0,67. Das beste Modell mit ausschließlich Texturvariablen erzielte nur minimal geringere Genauigkeiten als das kombinierte Modell aus spektraler Information und Texturinformation. Die Produzentengenauigkeit bei der Fichte blieb dabei gleich (siehe Tab. 4).

4.2.3 Baumholz

Der überwiegende Teil der Segmente ließ sich der natürlichen Altersklasse *Baumholz* zuordnen. In dieser wurde das beste rein auf spektralen Variablen beruhende Modell mit den Mittelwerten des Verhältnisses von Rot zu RGB bzw. Rot zu CIR sowie der Standardabweichung von Blau/RGB gebildet. Die Gesamtgenauigkeit lag mit 82,8 % höher als bei dem spektralen Modell der Gruppe *Stangenholz*, der Auslassungsfehler bei der Fichte lag jedoch weiterhin unter 70 %. Eine deutliche Steigerung des Kappa Koeffizienten von 0,59 auf 0,74 sowie eine Produzentengenauigkeit bei der Erkennung der Fichte von 79,6 % konnte unter Verwendung der spektralen Variablen mit zusätzlichen Texturvariablen (Mittelwerte von NIR/7x7/SV und Rot/5x5/CON sowie Minimum von NIR/3x3/CON) erreicht werden. Ähnlich wie beim *Stangenholz* erreichte das beste Modell mit ausschließlich Texturvariablen nur minimal geringere Genauigkeiten im Vergleich zum kombinierten Modell (siehe Tab. 4).

Tab. 4: Genauigkeitsanalyse der besten Modelle für *Baumholz* und *Stangenholz* (nach LOOCV) .

	Variablen		PA	UA			
Altersklasse	Gruppe	N	Fichte	Fichte	OA	Kappa	AIC
Baumholz	nur spektral	3	67,3%	76,7%	82,8%	0,59	134
	spektral & Textur	6	79,6%	84,8%	88,7%	0,74	99
	nur Textur	3	77,6%	84,4%	88,1%	0,72	107
Stangenholz	nur spektral	3	62,9%	73,3%	76,7%	0,49	109
	spektral & Textur	5	77,1%	81,8%	84,4%	0,67	93
	nur Textur	3	77,1%	79,4%	83,3%	0,65	101

PA Produzentengenauigkeit (producer‘s accuracy), UA Nutzergenauigkeit (user‘s accuracy) OA Gesamtgenauigkeit (overall accuracy)

Neben den oben genannten Texturparametern können noch weitere, via *stepAIC* ausgewählte Texturvariablen in der Klassifikation eingesetzt werden, ohne deutlich schlechtere Genauigkeiten zu erzielen. Da die gleichen Texturen für alle vier Kanäle der Luftbilder mit jeweils drei verschiedenen Fenstergrößen berechnet wurden, sind viele dieser identifizierten Variablen miteinander korreliert. Zu den geeignetsten Texturen zur Erkennung der Fichte zählten übergreifend über die Altersklassen *Baumholz* und *Stangenholz* Contrast (CON) und Difference Variance (DV). Im *Baumholz* wurden außerdem Homogeneity (IDM), Sum Variance (SV) und Variance (VAR) häufig ausgewählt, im *Stangenholz* Entropy, Sum Entropy und Uniformity (ASM). Texturparameter des grünen Kanals schienen weniger zur Unterscheidung beizutragen. Unter den drei getesteten Fenstergrößen konnte sich übergreifend über alle Texturen und Kanäle keine als vorrangig geeignet herauskristallisieren.

Eine Vergleichbarkeit mit anderen Studien zur Baumartenklassifikation unter Einbeziehung von Texturparametern ist aufgrund unterschiedlicher Regionen, der Anzahl der zu unterscheidenden Klassen respektive Baumarten, der gewählten Klassifikation sowie der benutzten Sensoren schwierig. Von den Haralick Texturen konnten HEINZEL & KOCH (2012) Contrast und Homogeneity als zwei der 14 wichtigsten Variablen zur Trennung von vier Baumarten identifizieren. KANDA *et al.* (2004) konnten ebenfalls Unterschiede im Parameter Homogeneity zwischen drei Baumarten erkennen und in Kombination mit zwei spektralen Variablen Klassifikationsgenauigkeiten von 80-90% erzielen. KULIKOVA *et al.* (2007) konnten hohe Genauigkeiten mit der Kombination von spektralen Variablen und Form- und Texturparametern bei der Unterscheidung von vier borealen Baumarten erzielen.

4.3 Übertragung

Die Genauigkeitsüberprüfungen aus der Kreuzvalidierung beziehen sich auf einzelbaumscharfe Segmente. Die Übertragbarkeit der Modelle auf die gesamte Fläche, und damit auch auf Mischsegmente, wurde stereoskopisch anhand von 198 zufällig ausgewählten Segmenten der Altersklassen *Stangenholz* und *Baumholz* überprüft. Acht dieser Segmente waren beschattete Laubbaumränder. Bei den übrigen 190 Segmenten lag die Produzentengenauigkeit bei der Fichte bei 67,6 %, der Kappa bei 0,62 (Tab. 5). Die geringere Genauigkeit im Vergleich zu der Kreuzvalidierung (Tab. 4) entsteht durch Mischsegmente und Segmente, die nur einen Teil der Krone ausmachen.

Tab. 5: Genauigkeitsanalyse der Übertragung der Modelle auf die gesamte Fläche anhand von 190 zufällig ausgewählten Segmenten der Klassen *Baumholz* und *Stangenholz*.

Vorhergesagt (Klassifikation)	Beobachtet				
	Fichte		sNB		
	Einzel	Misch	Einzel	Misch	UA
Fichte	36	10	6	3	83,6%
sNB	15	7	96	17	83,7%
PA	67,6%		92,6%		83,7%

Misch bedeutet kronenübergreifende Segmente, die von der jeweiligen Klasse dominiert werden. sNB sonstige Nadelbäume, PA Produzentengenauigkeit (producer's accuracy), UA Nutzergenauigkeit (user's accuracy), Gesamtgenauigkeit = untere rechte Zelle

5 Zusammenfassung & Ausblick

Eine einzelbaumscharfe Segmentierung als Grundlage für eine objekt-basierte Klassifikation von Baumarten ist in den vertikal und horizontal differenzierten Wäldern Mitteleuropas eine Herausforderung. Insbesondere gilt dies für Segmentierungsprozesse, die rein auf Stereoluftbildern beruhen. Eine Erhöhung der Punktdichte beim Image-Matching könnte die Segmentierung verbessern. Die Erkennung der Fichte konnte unter Verwendung von Texturparametern deutlich verbessert werden. Dies gilt sowohl für die älteren Bäume der natürlichen Altersklasse *Baumholz* als auch jüngere Baumindividuen aus dem *Stangenholz*. Eine automatische Unterscheidung von Baumarten der Altersklasse *Dickung* scheiterte sowohl an der

Qualität der erzeugten Segmente als auch der eindeutigen visuellen Interpretierbarkeit. Da die Fichte mit ihren klaren radialen Strukturen unter den hiesigen Baumarten bezüglich der Texturinformation am meisten Unterscheidungspotential aufweist, sind die Aussagen zur Eignung der Textur bei der Baumartenerkennung nicht eins zu eins auf andere Baumarten übertragbar. Zu testen wären zur Erkennung der Fichte beziehungsweise Trennung von weiteren Baumarten noch solche Texturparameter, die aus transformierten Spektralinformationen (z.B. Vegetationsindizes, IHS) erzeugt werden.

6 Literaturverzeichnis

ARBEITSGRUPPE FORSTLICHER LUFTBILDINTERPRETEN (AFL) 1998: Luftbild-Interpretationsschlüssel. Bestimmung der natürlichen Altersklasse von Waldbeständen im Color-Infrarot-Luftbild (CIR-Luftbild). *LÖBF-Mitteilungen*, **23** (1), S. 45-50.

HARALICK, R.M. 1979: Statistical and structural approaches to texture. *Proceedings of the IEEE*, **33** (5), S. 786-804.

HEINZEL, J. & KOCH, B. 2012: Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation. *International Journal of Applied Earth Observation and Geoinformation*, **18**, S. 101-110.

KANDA, F., KUBO, M. & MURAMOTO, K. 2004: Watershed Segmentation and Classification of Tree Species Using High Resolution Forest Imagery. *Proceedings of Geoscience and Remote Sensing Symposium, 2004. IEEE International*, **6**, S. 3822-3825.

KÖLLING, C., KNOKE, T., SCHALL, P. & AMMER, C. 2009: Überlegungen zum Risiko des Fichtenanbaus in Deutschland vor dem Hintergrund des Klimawandels. *forstarchiv*, **80** (5), S. 42-54.

KULIKOVA, M.S., MANI, M., SRIVASTAVA, A., DESCOMBES, X. & ZERUBIA, J. 2007: Tree species classification using radiometry, textures and shape based features. *15th European Signal Processing Conference, Poznan, Poland EURASIP*, S. 1595–1599.

LANDIS, J.R. & KOCH, G.G. 1977: The Measurement of Observer Agreement for Categorical Data. *Biometrics*, **33** (1), S. 159-174.

ORTIZ, S.M., BREIDENBACH, J., KNUTH, R. & KÄNDLER, G. 2012: The Influence of DEM Quality on Mapping Accuracy of Coniferous- and Deciduous-Dominated Forest Using TerraSAR-X Images. *Remote Sensing*, **4** (3), S. 661-681.

SEITZ, R., TROYCKE, A., REBHAN, P. & GRUBERT, B. 2010: Semi-automatic Assessment of Norway Spruce (Picea abies) in modern digital aerial Photographs. *ISPRS TC VII Symposium – 100 Years ISPRS, Vienna, Austria, July 5–7*, S. 517–522.

STRAUB, B.M. 2003: Automatische Extraktion von Bäumen aus Fernerkundungsdaten. Dissertation. Universität Hannover. 97 S.

VENABLES, W.N. & RIPLEY, B.D. 2002: Modern applied statistics with S. Springer, New York, 495 S.

WASER, L.T., GINZLER, C., KUECHLER, M., BALTSAVIAS E. & HURNI, L. 2011: Semi-automatic classification of tree species in different forest ecosystems by spectral and geometric variables derived from Airborne Digital Sensor (ADS40) and RC30 data. *Remote Sensing of Environment*, **115** (1), S. 76–85.

Landbedeckungsklassifikation von Satellitenaufnahmen unter Verwendung von Kontextinformation mit minimaler menschlicher Interaktion

ERIC BACH[1], BJÖRN FRÖHLICH[1,3], IRENE WALDE[2,3], SÖREN HESE[2,3], CHRISTIANE SCHMULLIUS[2,3] & JOACHIM DENZLER[1,3]

Zusammenfassung: Dieser Beitrag stellt ein Verfahren zur automatischen Klassifikation von Satellitenaufnahmen in verschiedene, vorher definierte, Landbedeckungsklassen vor. Das vorliegende Verfahren ermöglicht eine voll automatische Segmentierung und Klassifikation bei minimaler Nutzerinteraktion und wenigen Trainingsgebieten. Dazu werden Techniken der semantischen Segmentierung verwendet, die bereits in anderen Aufgabenstellungen der Bildverarbeitung, wie zum Beispiel in der Fassadenerkennung, große Erfolge aufweisen können. Bei dieser Methode werden neben lokalen Merkmalen auch kontextuelle Bezüge von Klassen untereinander berücksichtigt. Unter Zuhilfenahme von Kontextinformationen werden die Segmentierungsergebnisse signifikant verbessert. Das Verfahren ist flexibel und kann mit beliebig vielen Bildkanälen und deren Kombinationen arbeiten. Weiterhin ist es für verschiedene Szenarien einsetzbar und Anwendungen in anderen Themenbereichen mit räumlich geringer aufgelösten Daten sind möglich. Dennoch liegt der Schwerpunkt dieser Arbeit auf hochaufgelösten Satellitendaten im urbanen Bereich.

1 Einleitung

Die Anfänge der Landbedeckungsklassifikation aus Luftbildern liegen ca. 70 Jahre zurück (ANDERSON et al. 1976). Seither werden Luft- und Satellitenbilder verwendet, um großräumig und berührungsfrei Landbedeckung zu extrahieren. DI GREGORIO (2005) definiert Landbedeckung als beobachtete (bio)physikalische Schicht der Erdoberfläche. Ob als Grundlagedaten für Veränderungsdetektion oder zur Ableitung relevanter Planungs- und Modellierungsparameter, Landbedeckung ist eine essentielle Information, um Themen wie Klimawandel, Biodiversität, Ressourcenmanagement, Ökosystemeinwirkungen, Wohnqualität, Landnutzung oder Katastrophenmanagement zu analysieren und zu visualisieren (HEROLD et al. 2008, HÜTTICH et al. 2011, WALDE et al. 2012). Manuelle Kartierung/Digitalisierung oder Vermessung bedeuten einen hohen finanziellen, personellen und zeitlichen Aufwand. Deshalb werden automatische Prozessierungsmöglichkeiten auf Basis von flächendeckend verfügbaren Fernerkundungsdaten genutzt und stetig verbessert. Räumlich sehr hochaufgelöste Satellitenbilder (QuickBird, Ikonos, Worldview) ermöglichen die Ableitung der heterogenen Vielfalt urbaner Landbedeckung. Durch die räumlich sehr hoch aufgelösten Satellitenbilder wurden OBIA-Methoden (Object Based Image Analysis) entwickelt (BENZ et al. 2004, HAY & CASTILLA 2008, BLASCHKE 2010), die in urbanen Analysen den pixelbasierten Verfahren

[1] Lehrstuhl für digitale Bildverarbeitung, Friedrich-Schiller-Universität Jena
[2] Lehrstuhl für Fernerkundung, Friedrich-Schiller-Universität Jena
[3] Graduiertenschule für Bildverarbeitung und Bildinterpretation, ProExzellenz Thüringen
eMail: {eric.bach,bjoern.froehlich,irene.walde,soeren.hese,c.schmullius,joachim.denzler}@uni-jena.de

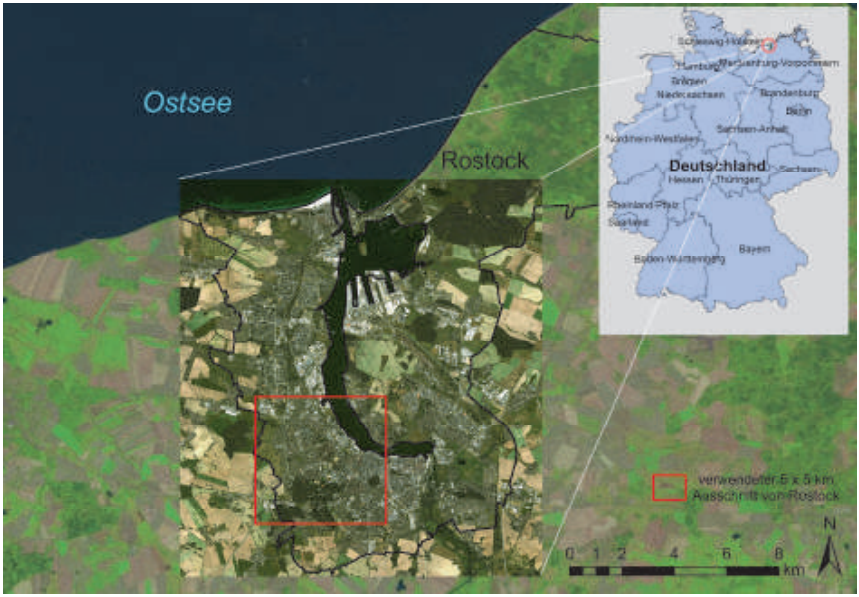


Abb. 1: QuickBird Szene (©DigitalGlobe, Inc., 2011) mit überlagerten Stadtgrenzen von Rostock und Kennzeichnung des verwendeten Ausschnitts. Hintergrund: Landsat 7 Szene (zur Verfügung gestellt vom U.S. Geological Survey, 1999).

vorzuziehen sind (MYINT et al. 2011). Pixelbasierte Verfahren betrachten ausschließlich spektrale Eigenschaften während objektbasierte Methoden, neben spektralen Eigenschaften auch Form, Textur oder Nachbarschaftsmerkmale in den Klassifikationsprozess einbeziehen.

In dieser Arbeit wird ein Verfahren zur automatischen semantischen Segmentierung und Klassifikation vorgestellt, dass mit Hilfe von minimaler Nutzerinteraktion und wenigen Trainingsgebieten urbane Landbedeckung aus hochaufgelösten Satellitenbildern extrahiert. Als Verfahren werden die so genannten Iterative Context Forests von FRÖHLICH et al. (2012a, b) verwendet. Bei dieser Methode werden neben lokalen Merkmalen auch kontextuelle Bezüge von Klassen untereinander berücksichtigt. So ist in Industriegebieten die Wahrscheinlichkeit der Nachbarschaft von Gebäuden und versiegelten Flächen wie z.B. Parkplätzen höher als bei Gebäuden in Gartenanlagen. Unter Zuhilfenahme von Kontextinformationen werden die Segmentierungsergebnisse signifikant verbessert. Das Verfahren ist flexibel und kann mit beliebig vielen Bildkanälen und deren Kombinationen arbeiten. Während eines Trainingsschrittes werden die optimalen Merkmale für eine Klasse aus einem großen Merkmalspool selbstständig ausgewählt, was der hohen spektralen Form- und Struktur-Variabilität der Klassen Rechnung trägt. Als Merkmale werden bewährte Maße aus der Bildverarbeitung verwendet, wie zum Beispiel Rechteckmerkmale aus der Personendetektion.

Die Iterative Context Forests wurden für Aufgabenstellungen der Bildverarbeitung wie z.B. Fassadenerkennung entwickelt und bezüglich Fernerkundungsdaten adaptiert.
Der Aufbau dieser Arbeit ist wie folgt strukturiert: Kapitel 2 beschreibt das Untersuchungsgebiet und das vorhandene Datenmaterial. In Kapitel 3 wird das Verfahren zur semantischen Segmentierung und die Adaptionen für Fernerkundungsdaten erläutert. Die Ergebnisse werden in Kapitel 4 dargelegt und diskutiert. Das Fazit und ein kurzer Ausblick schließen diesen Beitrag ab.

2 Untersuchungsgebiet und Datenbeschreibung

Das Untersuchungsgebiet Rostock, eine norddeutsche Hansestadt mit mehr als 200.000 Einwohnern und einer Fläche von ca. 181 km² (MECKLENBURG-VORPOMMERN STATISTISCHES AMT), liegt an der Warnow-Mündung in die Ostsee. Eine wolkenfreie QuickBird Szene vom September 2009 mit vier Kanälen (Blau, Grün, Rot, Nahes Infrarot) ist verfügbar (Abb. 1). Ein 25 km² großer Ausschnitt mit einer Bodenpixelauflösung von 2 m wurde verwendet, um die Methode zu testen. Die Szene wurde als Ortho-Ready Standard (OR2A) Produkt (CHENG et al. 2003) geliefert und ist auf eine mittlere Geländehöhe projiziert. Das Satellitenbild wurde bezüglich atmosphärischer Effekte korrigiert und mit Hilfe von Passpunkten und einem Geländemodell orthorektifiziert. Zusätzlich liegen LiDAR Daten für das Untersuchungsgebiet von 2006 vor. Um ein normalisiertes Oberflächenmodell (nDSM) zu erhalten, wurde das Geländemodell vom Oberflächenmodell subtrahiert. Damit stehen relative Objekthöhen als Rasterdatensatz mit einer Auflösung von zwei Metern zur Verfügung.

3 Semantische Segmentierung

Die Bezeichnung semantische Segmentierung fasst verschiedene Verfahren zur pixelgenauen Annotation von Bildern zusammen ohne auf ein spezielles Anwendungsszenario beschränkt zu sein.
Als Segmentierung wird der Prozess zur Einteilung eines Bildes in disjunkte Pixelmengen, sogenannte Regionen bezeichnet. Pixel innerhalb einer Region sind bezüglich eines für alle Regionen gleichen Homogenitätskriterium (z.B. der Farbe) ähnlich. Ziel der Segmentierung ist es, das Bild in eine einfacher analysierbare Form zu überführen, oder eine Darstellung zu finden, die auf die für eine konkrete Anwendung wesentlichen Bestandteile reduziert ist. Die klassische Segmentierung ist ein unüberwachtes Verfahren, was bedeutet, dass die Pixel zwar in Regionen zusammengefasst sind, diesen Regionen aber keine Bedeutung zugeordnet wird. Weiterhin bedeutet dies auch, dass es nicht zwangsläufig einen Trainingsschritt gibt. Regionen enthalten im Idealfall genau einen wesentlichen Bestandteil eines Bildes. Diese Bestandteile werden in dieser Arbeit als Objekte bezeichnet. Da der Idealfall häufig nicht erreicht wird, kann zwischen Über- und Untersegmentierung unterschieden werden. Während bei der Untersegmentierung verschiedene Objekte in einer Region zusammengefasst werden, kann bei der Übersegmentierung ein einzelnes Objekt aus mehreren Regionen bestehen.

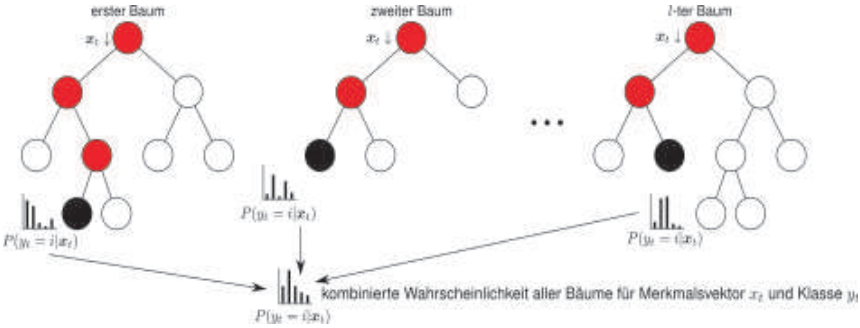


Abb. 2: Zufälliger Entscheidungswald aus *l* Bäumen. Für einen beispielhaften neuen Merkmalsvektor x_t zeigen die roten Knoten den Weg des Merkmalvektors an und die schwarzen die resultierenden Blätter.

Bei der semantischen Segmentierung wird jeder Region zusätzlich eine semantische Bedeutung zugeordnet. Dies ist üblicherweise ein Klassenname aus einer Menge von vorher definierten und trainierten Klassen. Die Auswahl der Klassen hängt von dem gewähltem Szenario und der Zielstellung ab. So kann z.B. die Satellitenaufnahme eines Landes betrachtet werden, wobei die gewählten Klassen Stadt oder Wald sein könnten. Wird hingegen die Aufnahme einer einzelnen Stadt betrachtet, dann wären sinnvolle Klassen beispielsweise Dach oder versiegelte Fläche.
In diesem Abschnitt werden die Iterative Context Forests (ICF) von FRÖHLICH et al. (2012a, b) vorgestellt. Anschließend werden die Unterschiede und Adaptionen im Vergleich zur Originalarbeit gezeigt.

3.1 Notwendige Grundlagen

Merkmalsvektoren sind eine abstrakte Ansammlung von Merkmalen, die ein Objekt oder den Teil eines Objektes beschreiben. So ist z. B. der durchschnittliche Farbwert innerhalb eines Objektes eine einfache Ansammlung von Merkmalen. Um komplexere Strukturen beschreiben zu können, sind neben Farbe auch Textur und Form wichtige Eigenschaften, die durch Merkmale beschrieben werden sollten.

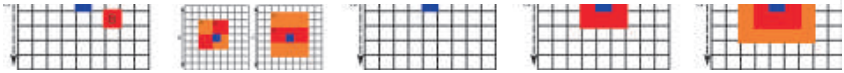
Das Klassifikationsproblem bezeichnet in der Mustererkennung die Zuordnung eines Klassennamen zu einem Merkmalsvektor. Um dies zu erreichen, müssen bereits vorher beschriftete Merkmalsvektoren zur Verfügung stehen, die die Klassen repräsentativ beschreiben. Ein Klassifikator versucht das Problem aus den Trainingsdaten während eines Trainingsschrittes zu abstrahieren, um neuen, unbekannten Merkmalsvektoren eine Klasse zuordnen zu können.

3.2 Iterative Context Forests

Iterative Context Forests (ICF) sind eine spezielle Art von Klassifikator, die auf Entscheidungswäldern (engl.: Random Decision Forests, RDF) (BREIMAN, 2001) basieren, die wiederum ein Ensemble von Entscheidungsbäumen (engl.: Decision Tree, DT) darstellen.

(e) Differenz zweier

Abb. 3: Merkmalsextraktion auf den Kanälen: Merkmale werden in einem Fenster der Größe *d* um einen aktuellen Pixel (blauer Pixel) extrahiert. In Abhängigkeit vom gewählten Merkmalstypen werden entweder einzelne Pixel oder ganze Regionen betrachtet. Jeder Parameter (z.B. Größe und Position eines Rechtecks) wird zufällig gewählt.



3.2.1 Entscheidungsbäume

Entscheidungsbäume (DUDA et al., 2000, Kapitel 8.2) sind ein schneller und einfacher Weg, um das Klassifikationsproblem zu lösen. Die Menge an Trainingsdaten wird durch simple Entscheidungen (z.B. ist der Wert im Grünkanal größer 127 oder nicht) in zwei Teilmengen aufgespalten. Diese Teilmengen werden mit anderen, ebenso einfachen, binären Entscheidungen weiter aufgeteilt, bis jede Teilmenge nur noch aus Merkmalsvektoren einer Klasse besteht. Dadurch entsteht eine Baumstruktur, bei der die Teilmengen mit nur noch einer Klasse als Blatt bezeichnet werden und alle anderen als innerer Knoten. Dieser Baum wird von einem neuen Merkmalsvektor mit unbekannter Klasse wiederum durchlaufen, bis dieser Vektor in einem Blatt endet. Die resultierende Klasse des unbekannten Merkmalvektors ist dieselbe, wie allen Trainingsmerkmalsvektoren in dem Blatt zugeordnet wurde. Als Trennfunktion einer Menge in zwei Teilmengen wird jeweils die beste Aufteilung durch erschöpfende Suche in den Trainingsdaten bezüglich der Maximierung der Kullback-Leibler-Entropie gewählt.

3.2.2 Zufällige Entscheidungswälder

In der Praxis hat sich gezeigt, dass Entscheidungsbäume sich stark an die Trainingsdaten anpassen (Überanpassung). Gerade bei stark verrauschten Daten kann sich ein Baum soweit aufteilen, dass viele Blätter nur noch einen Merkmalsvektor enthalten. Um das zu verhindern hat BREIMAN (2001) ein Verfahren vorgestellt, dass durch verschiedene zufällige Selektionen eine Überanpassung verhindert. Als erstes wird nicht mehr nur ein Baum trainiert, sondern mehrere. Jeder dieser Bäume bekommt eine andere zufällige Teilmenge an Trainingsdaten zum Trainieren. Weiterhin wird nicht solange trainiert, bis nur noch Merkmalsvektoren einer Klasse in einem Knoten vorkommen. Es gibt stattdessen verschiedene Abbruchkriterien: eine maximale Tiefe des Baumes darf nicht überschritten werden, eine minimale Anzahl von Trainingsbeispielen in einem Blatt darf nicht unterschritten werden und eine minimale Entropie darf nicht unterschritten werden. Somit entsteht eine a-posteriori Wahrscheinlichkeit aus der Verteilung der Klassen der verbliebenen Merkmalsvektoren in einem Blatt pro Baum. Ein neuer Merkmalsvektor durchschreitet alle Bäume und die wahrscheinlichste Klasse aus allen Bäumen wird dem unbekanntem Merkmalsvektor zugeordnet (siehe Abb. 2).

3.2.3 Ausnutzen von mehreren Millionen Merkmalen

Das vorgestellte Verfahren basiert auf der Extraktion von vielen verschiedenen Charakteristika aus einem Eingabebild. Neben den einzelnen Eingabekanälen können zusätzliche Kanäle, wie z.B. das Gradientenbild berechnet werden. Auf jedem dieser Kanäle und auf Kombinationen von diesen können pro Bildpunkt in einer Nachbarschaft *d* diverse Merkmale berechnet werden. So beispielsweise die Differenz zweier zufällig gewählter Pixel relativ zum aktuellen Pixel oder der mittlere Wert eines zufällig gewählten Bereiches relativ zum aktuellen Pixel (siehe Abb. 3).

3.2.4 Autokontextmerkmale

Der Hauptunterschied zu einem klassischen Entscheidungswald ist die Verwendung von sich verändernden Merkmalen während des Trainingsprozesses. Die Bäume müssen dafür zwangsläufig Ebene für Ebene gelernt werden. Nach jeder Ebene werden die Wahrscheinlichkeiten für jeden Pixel und für jede Klasse als zusätzlicher Kanal der möglichen Merkmalsauswahl hinzugefügt. Wenn davon ausgegangen wird, dass das vorherige Ergebnis mit hoher Wahrscheinlichkeit richtig ist, kann dadurch Kontextwissen aus der Nachbarschaft extrahiert und verwendet werden.

3.3 Adaptionen für Fernerkundungsprobleme

Das vorgestellte Verfahren wurde bisher auf Datensätze angewandt, die Hausfassaden abgebildet haben (FRÖHLICH et al., 2012a, b). Dieses Szenario unterscheidet sich von den Ansprüchen der Fernerkundung. Bei dem in dieser Arbeit verwendeten 5 x 5 km großen Ausschnitt von Rostock sind die abgebildeten Objekte durch die hohe Gebietsabdeckung und Auflösung kleiner, als beispielsweise ein Fenster einer Hauswand bei einer Fassadenaufnahme. Um dieser Charakteristik gerecht zu werden, wurde die Fenstergröße *d* reduziert (siehe Abschnitt 3.2.3 und Abb. 3). Des Weiteren sind einige der verwendeten Merkmalskanäle aus FRÖHLICH et al. (2012b) nicht sinnvoll auf Fernerkundungsdaten anwendbar. Dazu gehört der geometrische Kontext (HOIEM et al., 2005). Stattdessen kommen als zusätzliche Merkmale für die klassische Bildverarbeitung untypische Kanäle wie nahes Infrarot und LiDAR nDSM zum Einsatz. Durch den flexiblen Algorithmus ist die Nutzung von beliebig vielen Merkmalskanälen möglich. Aus diesem Grund kann als zusätzlicher Eingabekanal der „Normalized Differenced Vegetation Index“ (NDVI) genutzt werden. Dieser Index, berechnet aus rotem und nahem Infrarot Kanal, ermöglicht eine gute Abgrenzung von Vegetation und versiegelten Flächen.

4 Ergebnisse und Diskussion

Zum Testen wurden einzelne Trainingsgebiete markiert. Auf den restlichen Bereichen wurde das vorgestellte Verfahren getestet und für zufällig ausgewählte 75 Punkte je Klasse validiert. Jede Klassifikation der unterschiedlichen Kanalkombinationen wurde mit zufällig ausgewählten Merkmalen zehnmal wiederholt und die Ergebnisse gemittelt. Die Landbedeckungsklassen Gebäude, Versiegelt, Boden, Baum, Gras und Wasser wurden extrahiert.

Alle Tests wurden mit der Fenstergröße von zehn Pixeln für nicht kontextuelle Merkmale und 40 Pixeln für alle kontextuellen Merkmale durchgeführt. Diese Werte haben sich in vorangestellten Testläufen als optimal herausgestellt.

Bei dem Test des vorgestellten Verfahrens auf dem QuickBird Datensatz (Abb. 4a) ergab sich eine relativ geringe producer accuracy (Herstellergenauigkeit) und user accuracy (Nutzergenauigkeit) für die Klassen Gebäude, Gras, sowie versiegelte Flächen. Unter Einbeziehung der relativen Objekthöhen aus dem LiDAR nDSM konnte die producer accuracy für die Klassen Versiegelt und Gebäude verbessert werden (Abb. 4b). Die user accuracy der Klassen Gebäude, Gras und versiegelte Flächen wurde ebenfalls signifikant verbessert. Durch die Hinzunahme des NDVI (Abb. 4c) verbessert sich die producer und user accuracy für die Klasse Gras. Nimmt man, zusätzlich zu dem QuickBird Datensatz, nDSM und NDVI hinzu (Abb. 4d), verbessert sich die Erkennungsrate nochmals leicht.

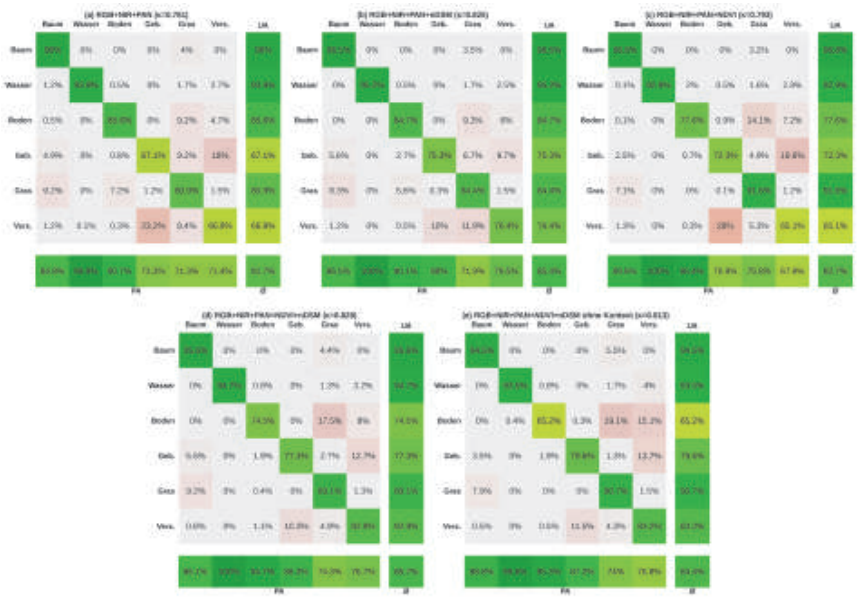


Abb. 4: Ergebnis unter Verwendung der QuickBird-Kanäle (RGB, NIR, PAN) (a-e) und zusätzlicher Verwendung des LiDAR nDSM (b,d,e), des NDVI (c-e) und ohne Kontextinformation (e). UA: user accuracy, PA: producer accuracy, ø: durchschnittliche Erkennungsrate.

In Abb. 4e sind die Ergebnisse der besten Methode (QuickBird Daten, NDVI und nDSM Kanäle) ohne die Verwendung der Autokontextmerkmale dargestellt. Die Erkennungsraten und die producer accuracy sind für viele Klassen schlechter, wie in Abb. 4d unter Verwendung der Kontextinformationen. Deutlich wird der Einfluss von Kontextinformationen bei Betrachtung der Fehlerrate von Boden in Abb. 4e. Ohne Kontext entscheidet sich der Klassifikator oft für versiegelte Fläche anstatt Boden. Mit Kontext hingegen wird Boden vom Klassifikator sicherer erkannt (vgl. Abb. 4d). Eine mögliche Erklärung ist, dass die Bodengebiete weniger im Stadtgebiet sondern im Umland vorkommen. Dadurch ist die Nachbarschaft von Boden zu Gras/Baum wahrscheinlicher als die zu Gebäuden. Unter Einbezug dieser Nachbarschaftsinformationen wird Boden mit geringerer Fehlerrate klassifiziert (Abb. 4d).
Das beste Klassifikationsergebnis (Nutzen von Kontextinformationen auf QuickBird Daten, nDSM und NDVI) ist in Abb. 5 dargestellt.

5 Fazit und Ausblick

In dieser Arbeit wurde ein aktuelles Verfahren der Bildverarbeitung für die Semantische Segmentierung vorgestellt. Weiterhin wurde veranschaulicht, wie dieses Verfahren für die Landbedeckungsklassifikation verwendet werden kann. In den Experimenten konnte gezeigt

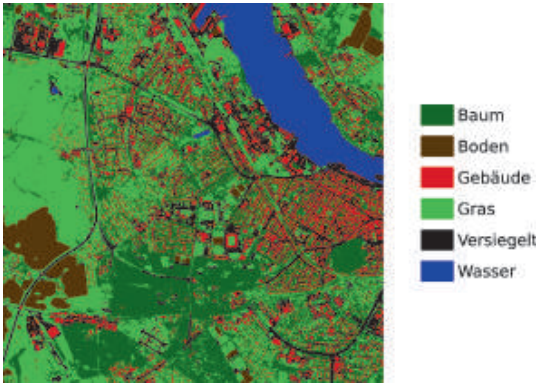


Abb. 5: Klassifikationsergebnis des 5 x 5km Ausschnittes von Rostock.

werden, dass die Verwendung von LiDAR nDSM, NDVI sowie von Kontextinformationen für diese Problemstellung das Ergebnis signifikant verbessern.
Für die zukünftige Arbeit ist die Verwendung einer unüberwachten Segmentierung geplant, um die Genauigkeit, insbesondere an den Klassenrändern, zu verbessern. Eine weitere Verbesserung der Erkennungsraten und des Detaillierungsgrades ist zu erwarten, wenn die pangeschärften Bilder mit einer räumlichen Auflösung von 60 cm verwendet werden. Zusätzlich soll die Performanz bezüglich der Analyse der gesamten hochaufgelösten QuickBird Szene (13,8 x 15,5 km) untersucht werden.

Danksagung

Diese Arbeit wurde von der ProExzellenz Initiative des Thüringer Ministeriums für Bildung, Wissenschaft und Kultur (TMBWK) unterstützt (Förderkennzeichen: PE309-2). Wir möchten der Stadt Rostock für die Bereitstellung der LiDAR-Daten danken.

6 Literaturverzeichnis

ANDERSON, J. R.; HARDY, E. E.; ROACH, J. T.; WITMER, R. E., 1976: A Land Use and Land Cover Classification System for Use with Remote Sensor Data. United States Government Printing Office, Washington. U.S. Geological Survey Circular, **671**.

BENZ, U. C.; HOFMANN, P.; WILLHAUCK, G.; LINGENFELDER, I.; MARKUS, H., 2004: Multi-Resolution, Object-Oriented Fuzzy Analysis Of Remote Sensing Data For Gis-Ready Information, ISPRS Journal of Photogrammetry & Remote Sensing, **58**, S. 239–258.

BLASCHKE, T., 2010: Object Based Image Analysis For Remote Sensing. ISPRS Journal of Photogrammetry and Remote Sensing, **65** (1), S. 2–16.

BREIMAN, L.,2001: Random Forests. Machine Learning, Kluwer Academic Publishers, **45**, S. 5–32.

CHENG, P.; TOUTIN, T.; ZHANG, Y.; WOOD, M., 2003: QuickBird – Geometric Correction, Path and Block Processing and Data Fusion. Earth Observation Magazine, **12** (3), S. 24–28.

DI GREGORIO, A., 2005: Land cover classification system software version 2. Based on the orig. software version 1. rev. Rome. Environment and natural resources series Geo-spatial data and information, **8**.

DUDA, R. O.; HART, P. E. & STORK, D. G., 2000: Pattern Classification (2nd Edition), Wiley-Interscience.

FRÖHLICH, B.; RODNER, E. & DENZLER, J., 2012a: As Time Goes By — Anytime Semantic Segmentation with Iterative Context Forests, Symposium of the German Association for Pattern Recognition (DAGM), S. 1–10.

FRÖHLICH, B.; RODNER, E. & DENZLER, J., 2012b: Semantic Segmentation with Millions of Features: Integrating Multiple Cues in a Combined Random Forest Approach", Asian Conference on Computer Vision (ACCV).

HAY, G. J.; CASTILLA, G., 2008: Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. Object-Based Image Analysis. Spatial Concepts For Knowledge-Driven Remote Sensing Applications. Berlin, London: Springer, S. 75–89.

HEROLD, M.; WOODCOCK, C.E.; LOVELAND, T.R.; TOWNSHEND, J.; BRADY, M.; STEENMANS, C.; SCHMULLIUS, C.C ., 2008: Land-Cover Observations as Part of a Global Earth Observation System of Systems (GEOSS): Progress, Activities, and Prospects. IEEE Systems Journal, **2** (3), SS. 414–423.

HOIEM, D.; EFROS, A. A. & HEBERT, M., 2005: Geometric Context from a Single Image, International Conference of Computer Vision (ICCV), IEEE, **1**, 654–661.

HÜTTICH, C.; HEROLD, M.; WEGMANN, M.; CORD, A.; STROHBACH, B.; SCHMULLIUS, C.; DECH, S., 2011: Assessing effects of temporal compositing and varying observation periods for large-area land-cover mapping in semi-arid ecosystems: Implications for global monitoring. Remote Sensing of Environment, **115** (10), S. 2445–2459.

MECKLENBURG-VORPOMMERN STATISTISCHES AMT. [Online]. http://www.statistik-mv.de.

MYINT, S. W.; GOBER, P.; BRAZEL, A.; GROSSMAN-CLARKE, S.; WENG, Q., 2011: Per-Pixel Vs. Object-Based Classification Of Urban Land Cover Extraction Using High Spatial Resolution Imagery. Remote Sensing of Environment, **115** (5), S. 1145–1161.

VIOLA, P. & JONES, M., 2001: Rapid Object Detection Using A Boosted Cascade Of Simple Features, Computer Vision and Pattern Recognition, IEEE (CVPR), **1**, 511–518.

WALDE, I.; HESE, S.; BERGER, C.; SCHMULLIUS, C., 2012: Graph-based Urban Land Use Mapping from High Resolution Satellite Images. ISPRS Ann. Photogramm. Remote Sensing Spatial Inf. Sci., **I-4**, S. 119–124.

Automatisierte Einzelbaumerfassung auf Basis von hochaufgelösten TrueOrtho-Mosaiken und photogrammetrisch abgeleiteten Oberflächenmodellen

STEVEN BAYER[1], MOHSEN MIRI[2] & TILMAN BUCHER[3]

Zusammenfassung: Die Forstfernerkundung hat in Deutschland eine lange Tradition. In den letzten 30 Jahren wurden zudem verschiedenste Verfahren der automatischen Extraktion von Einzelbäumen aus Fernerkundungsdaten entwickelt und veröffentlicht. Dabei bestand jedoch eine Fokussierung auf Satellitenbilder und Laserdaten. Zudem haben viele Ansätze ihre Tauglichkeit vor allem im Nadelwald unter Beweis gestellt. Die automatische Einzelbaumextraktion von Laubbäumen ist dagegen bis heute eine Herausforderung. Daher stand im Mittelpunkt des entwickelten Verfahrens nicht nur die Extraktion von Nadelbäumen aus Luftbildern und photogrammetrisch erzeugten Oberflächen, sondern auch eine genaue Einzelbaumextraktion von im geschlossenen Bestand stehenden Laubbäumen. Wie das vorgestellte Verfahren zeigt, lassen sich unter der Verwendung von TrueOrtho-Luftbildern und photogrammetrisch erzeugten Oberflächen Einzelbäume mit einer sehr hohen Genauigkeit ableiten. Dies gilt dabei nicht nur für Nadelbäume sondern auch für Laubbäume.

1 Einleitung

Luftbilder sowie deren Auswertung werden in der Forstwirtschaft schon seit vielen Jahrzehnten für die Unterstützung der Forstinventur eingesetzt. Entsprechend lang ist auch die Forschungstradition in der automatisierten Auswertung von Luftbildern. Mit der Verbreitung von digitalen Luftbildern wurden seither vielfältige Ansätze für die Extraktion von Einzelbäumen entwickelt (STRAUB, 2003). Die automatische Extraktion von im Bestand stehenden Einzelbäumen, insbesondere von Laubbäumen, aus digitalen Luftbildern stellt jedoch bis heute, aufgrund der in den Luftbildern auflösungsbedingt hohen Komplexität, eine große Herausforderung dar. Ziel dieses Beitrages ist die Vorstellung eines leicht nachvollziehbaren Verfahrens der Einzelbaumextraktion aus digitalen Luftbildern unter der Verwendung von bekannten Algorithmen. Das entwickelte Verfahren zeigt dabei nicht nur einen Weg auf, wie Einzelbäume im Nadelwald extrahiert werden, sondern ermöglicht auch eine Einzelbaumextraktion von im geschlossen Bestand stehenden Laubbäumen mit einer hohen Genauigkeit und Vollständigkeit der automatisch extrahierten Baumspitzen und Baumumrisse.

1) Steven Bayer, Institut für Internationale Stadtforschung (InUrban), INA gGmbH an der Freien Universität Berlin, Malteser Straße 74-100, 12249 Berlin; E-Mail: steven.bayer@fu-berlin.de
2) Mohsen Miri, Beuth Hochschule für Technik Berlin, FB III, Labor für Photogrammetrie, Luxemburger Straße 10, 13353 Berlin; E-Mail: mohsen.miri@fu-berlin.de
3) Tilman Bucher, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Fachabteilung Sensorkonzepte und Anwendungen, Rutherfordstraße 2, 12489 Berlin; E-Mail: tilman.bucher@dlr.de

2 Untersuchungsgebiet und Datengrundlage

Ein Großteil der bisher veröffentlichten Verfahren zur automatischen Extraktion von Einzelbäumen stützt sich bei der Analyse auf die Verwendung von Satellitenbildern und Laserdaten. In dem hier vorgestellten Verfahren werden bis auf die Verwendung eines Standard DGM ausschließlich hochauflösende TrueOrtho-Mosaike (TOM) und mittels Semi-Global Matching (SGM) photogrammetrisch erzeugte Digitale Oberflächenmodelle (DOM) verwendet. Als Testdatensätze standen dafür eine DMC-1 Befliegung vom 2. und 3. Mai 2011 mit einer 80/40 Überlappung sowie eine UltraCamXP Befliegung vom 22. September 2010 mit einer Überlappung von 80/30 zur Verfügung. Die x/y-Auflösung betrug bei beiden Befliegungen jeweils 20 cm. Die z-Auflösung des DOM betrug jeweils 10 cm (siehe Abb. 1). Die Eignung der Daten für die forstwirtschaftlichenAnwendungen wurde in einer vorab durchgeführten Studie bereits nachgewiesen (MIRI et. al. 2013).
Da die Entwicklung des Verfahrens der Einzelbaumextraktion in Zusammenarbeit mit dem Projekt „Der Virtuelle Wald“ (RIF E.V., 2012) stattfand, wurden zwei Untersuchungsgebiete in Nordrhein-Westfalen ausgewählt. Bei beiden Regionen handelt es sich um für Mitteldeutschland typische Mischwaldgebiete. Untersuchungsgebiet eins ist das NSG Habichtswald in der Nähe von Osnabrück. Das Gebiet ist geprägt durch einen hohen Anteil an Buchen-Eichenmischwald sowie Kleinflächen von Erlenwald und Fichtenwald (LANUV NRW, 2012). Untersuchungsgebiet zwei liegt ca. 20 km südwestlich von Siegburg und gehört zur Leudscheid, einem bewaldeten Höhenzug an der Grenze von Rheinland-Pfalz und Nordrhein-Westfalen. Die Leudscheid ist geprägt durch einen hohen Anteil an aufgeforsteten Fichtenwäldern (LANIS 2012). Die gewählten Untersuchungsgebiete bieten durch ihre Bestandsvielfalt eine differenzierte Analysengrundlage.

2.1 Semi-Global Matching (SGM)

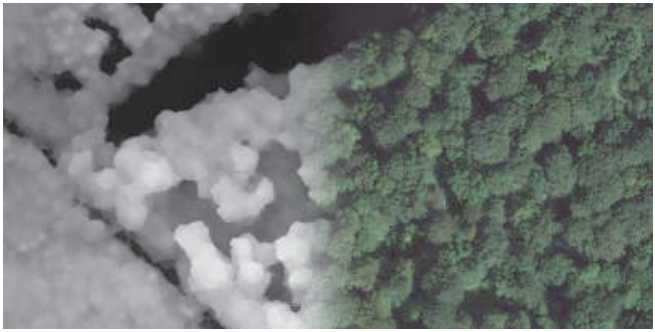


Abb. 1: Mit SGM prozessiertes DOM und TOM

Semi-Global Matching (HIRSCHMÜLLER, 2005) hat sich in der photogrammetrischen Community als führende Methode zur Generierung von Oberflächenmodellen erwiesen (Verleihung des Carl

Pulfrich Preis 2011 an H. Hirschmüller). Aus Luftbildbefliegungen mit idealerweise ≥ 80 % Längsüberlappung lassen sich aus Standard-Luftbilddaten qualitativ hochwertige pixelscharfe und sehr dichte digitale Oberflächenmodelle generieren. Basierend auf diesen werden TrueOrthobild-Mosaike abgeleitet, die bei hohen Überlappungen auch im Wald nur geringe Sichtschatten aufweisen (siehe Abb. 1).

2.2 Digitales Geländemodell (DGM)

Zur Berechnung des normalisierten Oberflächenmodelles (nDOM) wurde ein standard Digitales Geländemodell (DGM 1) mit einer Auflösung von einem Meter von Geobasis NRW verwendet. Auf die direkte Ableitung eines DGM aus dem DOM (MAYER, 2004) wurde aus Gründen der Nachvollziehbarkeit verzichtet. Für die Einzelbaumextraktion ist ein nDOM nicht zwingend notwendig. Dies erleichtert aber die Klassifikation und ist für die Ableitung von Baumhöhen notwendig.

2.3 Kontrolldaten

Für die Berechnung der Genauigkeiten und die Bestimmung der Kontrollbäume wurden im Gebiet des Habichtswaldes mit dem Riegl Laser Scanner LMS Z620 exemplarische Waldgebiete terrestrisch mit lokaler Genauigkeit unter 6 mm vermessen. Die globale Genauigkeit der so erhobenen Kontrollpunkte liegt bei unter 5 mm bezogen auf geodätische Festpunkte. Im Anschluss wurden mit der Software PHIDIAS in Microstation die Kontrollbäume mit der Genauigkeit von ca. 1 cm bestimmt. Dabei wurden die Bäume erhoben, die sowohl in der terrestrischen Punktwolke als auch in dem dazu gehörigen Foto sichtbar waren. Im Zuge der Kontrollbaumbestimmung wurden somit auch sehr kleine Bäume erhoben, um auch Aussagen über den Unterbau treffen zu können. Für die Bestimmung der Vollständigkeit und Korrektheit wurde im Laubwald aber nur auf Bäume mit einem Durchmesser über 20 cm zurückgegriffen. Im Nadelwald wurden nur Bäume mit einem Durchmesser von mindestens 15 cm verwendet.
Im zweiten Waldgebiet wurden im Nadelwald zusätzlich weitere Kontrollbäume visuell aus dem Luftbild bestimmt, um genauere Aussagen über die Vollständigkeit und Korrektheit der automatisch extrahierten Bäume treffen zu können. Insgesamt wurden im Nadelwald ca. 500 Kontrollbäume erhoben.

3 Methoden

Die ersten Methoden der automatischen Extraktion von Einzelbäumen aus Luftbildern wurden ab Mitte der 1980er Jahre entwickelt. Diese Ansätze identifizierten die Einzelbäume mit Hilfe der Detektion von lokalen Helligkeits- oder Intensitätsmaxima im Luftbild (vgl. HAENEL et. al., 1987). Mit zunehmender Bildauflösung gewann neben der Radiometrie auch die Geometrie eine große Bedeutung bei der Erkennung von Einzelbäumen (vgl. FOURNIER et. al., 1995). Bedingt durch die Entwicklung des flugzeuggetragenen Laserscanners und die Ableitung von direkt gemessenen Oberflächenmodellen wurden in der jüngeren Vergangenheit verstärkt Ansätze zur Einzelbaumextraktion aus DOM publiziert (vgl. HYYPPÄ et.al., 2001; KAARTINEN & HYYPPÄ, 2008).

Der in Kooperation mit dem DLR, Abteilung Sensorenkonzepte und Anwendungen in Berlin Adlershof entwickelte Ansatz verwendet neben den radiometrischen Informationen des TOM auch die geometrischen Informationen des DOM für die Extraktion von im Bestand stehenden Einzelbäumen. Dafür wurde ein mehrstufiges Verfahren entwickelt, welches für Nadel- und Laubwaldbestände neben der Bestimmung der Baumspitze auch eine Segmentierung der Baumumringe ermöglicht (siehe Abb. 2). Grundlage für alle eingesetzten Verfahren ist, aufgrund der Morphologie der Bäume, die Erstellung einer Laubwald- und einer Nadelwaldklasse zur differenzierten Betrachtung.
Für die nachfolgend beschriebenen Verfahren wurden die kommerziellen Softwareprodukte ArcGIS 10.1 für die GIS-Analysen und eCognition 8 für die Segmentierung und Klassifizierung verwendet.

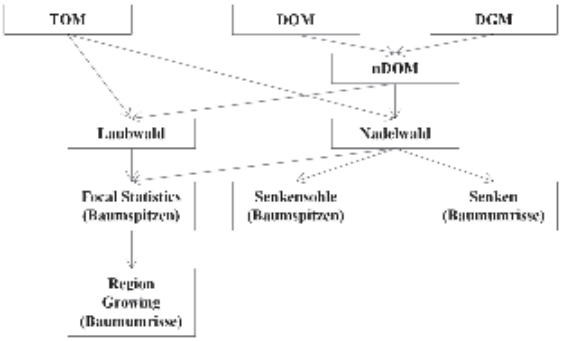


Abb. 2: Verfahren der Einzelbaumextraktion

3.1 Focal Statistics

Mit Hilfe der sogenannten „Focal Statistics“ (ESRI, 2012) wird eine Nachbarschaftsoperation durchgeführt, bei welcher der Wert für eine Ausgabezelle eine Funktion aller Werte der benachbarten Eingabezellen ist. Für die Extraktion der Baumspitzen wird dabei mit Hilfe einer Maximalfunktion in einem vorgegeben Suchraster nach ‚lokalen Maxima‘ im Oberflächenmodell

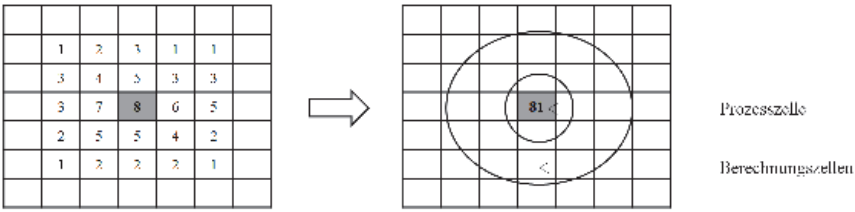


Abb. 3: Focal Statistics Berechnung (Quelle: Abbildung verändert nach ESRI, 2012)

gesucht. Anschließend wird die Summe aller Werte in der Nachbarschaft des lokalen Maximums aufsummiert und ausgegeben. Dabei wird mit einer Kreisringform, welche einen kleinen Kreis innerhalb eines größeren Kreises darstellt, gearbeitet. Bei dieser Form der Auswertung werden nur die Zellen zwischen dem inneren und äußeren Kreis in die Berechnung einbezogen (ESRI, 2012), wodurch die Größe der Bäume über das Verhältnis von innerem und äußerem Ring berücksichtigt werden kann (siehe Abb. 3). Durch die Verschneidung von Oberflächenmodell mit den Ergebnissen der Focal Statistics Analyse können die Baumspitzen sowohl im Nadelwald als auch im Laubwald mit einer hohen Genauigkeit und Vollständigkeit bestimmt werden.

3.2 Region Growing

Zur Bestimmung der Baumumrisse wird das Bildsegmentierungsverfahren Region Growing verwendet (SCHIEWE, 2002). Die mit Hilfe der Focal Statistics Analyse erstellten Baumspitzen werden dabei als initiale Zellen, d.h. als Anfangsregionen für die Segmentierung verwendet. Durch die Berücksichtigung der spektralen und geometrischen Eigenschaften der zu detektierenden Bäume sowie durch die Beachtung von Nachbarschaftsverhältnissen innerhalb der Baumbestände, aber auch durch die Annahme von statistischen Größenverhältnissen für die Laub- und Nadelbäume, lassen sich mit einem Region Growing Verfahren die Baumumrisse für beide Waldtypen sehr genau bestimmen.

3.3 Lokales Maximum und Senken

Parallel zum Verfahren der Detektion von Baumspitzen und Baumumrissen durch den Einsatz einer Focal Statistics Analyse und eines Region Growing Bildsegmentierungsverfahrens wurde ein zweiter Ansatz entwickelt, mit welchem im Rahmen einer GIS-Analyse sowie durch den Einsatz von hydrologischen Tools Baumspitzen und Baumumrisse ebenfalls gut bestimmt werden können (siehe Abb. 4). Dazu ist zu Beginn der Analyse die Invertierung des

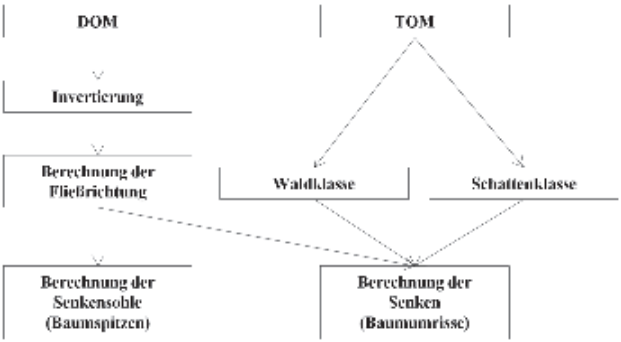


Abb. 4: Berechnung der Lokalen Maxima und Senken aus dem Oberflächenmodell

Oberflächenmodelles notwendig. Anschließend wird für das invertierte DOM die Fließrichtung von jeder Zelle zur nächsten Nachbarzelle mit der steilsten Neigung berechnet. Auf Basis der berechneten Fließrichtung wird das invertierte DOM im Modell ‚geflutet' und die Wassereinzugsgebiete (im folgenden Senken genannt) berechnet. Die Senken werden mit Hilfe

der Berechnung der Wasserscheide bestimmt. Durch die Verschneidung der Senken mit einem vorab klassifizierten Wald- und Schattenlayer können so die Baumumrisse extrahiert werden. Neben der Berechnung der Senken werden auch deren Sohlen bestimmt, welche als lokale Maxima des Oberflächenmodelles die Baumspitzen gut abbilden.

4 Ergebnisse

Mit Hilfe der entwickelten Verfahren können sowohl Nadel- als auch Laubbäume, welche in geschlossenen Beständen stehen mit einer sehr hohen Genauigkeit und Vollständigkeit extrahiert werden. Um die Ergebnisse quantifizieren zu können, wurde nach STRAUB (2003) die Vollständigkeit und Korrektheit der Ergebnisse berechnet.

Die Vollständigkeit (V) der Ergebnisse wird dabei wie folgt berechnet:

$$\text{Vollständigkeit}\,(\%) = \frac{\text{richtig extrahierte Bäume}}{\text{richtig extrahierte Bäume} + \text{nicht extrahierte Bäume}} * 100$$

Die Korrektheit (K) der Ergebnisse wird dabei wie folgt berechnet:

$$\text{Korrektheit}\,(\%) = \frac{\text{richtig extrahierte Bäume}}{\text{richtig extrahierte Bäume} + \text{falsch extrahierte Bäume}} * 100$$

4.1 Laubwald

Die automatische Einzelbaumextraktion von im geschlossenen Bestand stehenden Laubbäumen ist bis heute eine wichtige Aufgabe der Fernerkundung. Aus diesem Grund war es das Hauptziel der Arbeit ein Verfahren zu entwickeln, mit dem auch die Baumspitzen und Baumumringe von Laubbäumen mit einer hohen Vollständigkeit und Genauigkeit automatisch erhoben werden können. Für die Berechnung der Vollständigkeit und Genauigkeit wurden 89 terrestrisch eingemessene Laubbäume mit einem Durchmesser von mindestens 20 cm verwendet. Von diesen 89 Kontrollbäumen konnten für 76 die Baumspitzen richtig extrahiert werden, was einer Vollständigkeit von 87,25 % entspricht. Dabei wurde aufgrund der Morphologie der Baumkrone und des Wachstumsverhaltenes der Laubbäume ein Maximalabstand von fünf Metern zwischen Kontrollbaum und automatisch erhobener Baumspitze für die Bestimmung der Vollständigkeit verwendet. Bei den mit Hilfe der Focal Statistics Analyse automatisch bestimmten Baumspitzen konnten nur fünf als falsch identifiziert werden, was einer Korrektheit von 94,68 % entspricht. Wie die Ergebnisse zeigen, lassen sich mit Hilfe der Focal Statistics Analyse gute Ergebnisse bei der automatischen Extraktion der Baumspitzen von Laubbäumen erzielen. Nicht einzeln detektierbar sind lediglich sehr dicht stehende Bäume gleicher Größe, die überwiegend als Einzelbaum mit nur einer Baumspitze detektiert werden, im Luftbild aber auch visuell nicht zu unterscheiden sind. Vereinzelt wurden auch einige im TOM visuell offensichtlich richtig extrahierte Baumspitzen als nicht extrahierte Baumspitzen ausgegeben. Die Gründe hierfür sind mit dem teilweise sehr schrägen Stammwachstum einiger Laubbäume sowie durch die eventuelle Aufzweigung der Krone eines Baumes zu erklären. In diesen Fällen kann die Baumkrone

mehrere nicht eindeutig bestimmbare Spitzen enthalten. Dadurch kann der Abstand zwischen der auf Brusthöhe gemessenen Kontrollposition des Baumes und der im DOM extrahierten Baumspitzen mehr als fünf Meter betragen. Bei einer rein visuellen Beurteilung von TOM und DOM wären diese Baumspitzen als richtig extrahiert eingeordnet worden, was dementsprechend auch die Vollständigkeit erhöht hätte.

Da die Baumumringe aus den Baumspitzen berechnet wurden, beträgt auch hier die Vollständigkeit 87,25 %. Für die Beurteilung der Korrektheit wurden die automatisch erstellten Baumsegmente visuell in richtig und falsch extrahiert eingeteilt. Die so ermittelte Korrektheit der Baumumrisse liegt bei 93,68 %. Durch die Begrenzung des Segmentwachstums auf ein Größenmaximum und festgelegte Spektralbereiche liegt die Korrektheit der Baumumrisse über der der extrahierten Baumspitzen, weil auch die schräg wachsenden Bäume und die sich teilenden Kronen korrekt extrahiert werden (siehe Abb. 5).

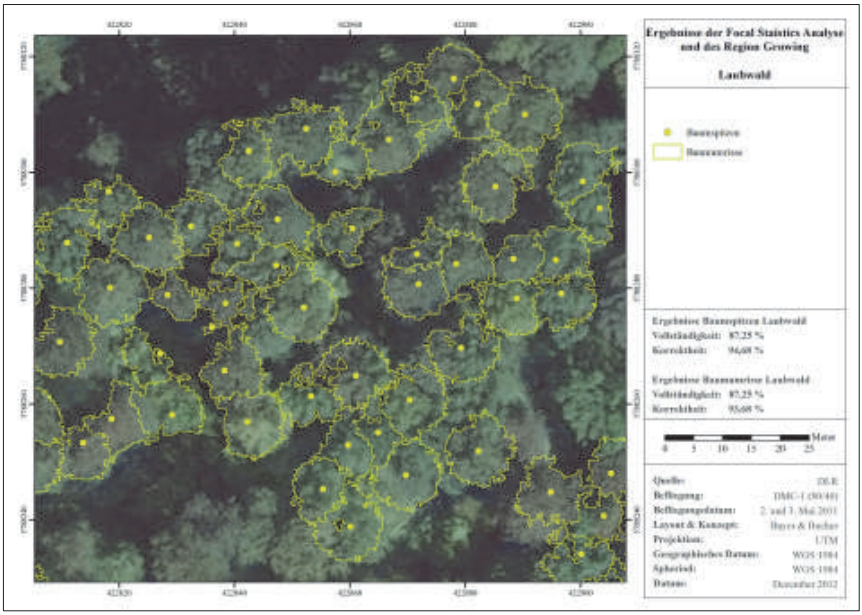


Abb. 5: Ergebnisse der automatischen Einzelbaumextraktion im Laubwald

4.2 Nadelwald

Obwohl für den Nadelwald mit der Focal Statistics Analyse und dem Region Growing Verfahren geringfügig höhere Genauigkeiten erreicht werden, werden im folgenden Kapitel trotzdem die Ergebnisse der Senken- und Senkensohlenanalyse vorgestellt, um auch deren Potential aufzeigen zu können.

Von den Kontrollbäumen im Nadelwald wurden bei einem Maximalabstand von 2,5 m zwischen Kontrollpunkt und extrahiertem Punkt 91,15 % der Baumspitzen vollständig bestimmt. 86,3 %

der ausgegeben Baumspitzen werden bei dieser Analyse korrekt bestimmt. Im Vergleich zur Focal Statistics Analyse zeigt sich bei der Bestimmung der Baumspitze mit der Berechnung der Senkensohle, dass mehr lokale Maxima automatisch extrahiert werden. Auch ist die Lagegenauigkeit der mit der Focal Statistics Analyse berechneten Baumspitzen minimal genauer.

Werden die Baumumrisse mit Hilfe der Senken aus dem Oberflächenmodell bestimmt, liegt die Vollständigkeit bei über 99 %, d.h. es werden fast alle Nadelbäume automatisch extrahiert. Da die Extraktion der Baumumrisse eine für sich eigenständige GIS-Analyse ist und diese nicht mit Hilfe der Baumspitzen berechnet werden, ist eine annähernd hundertprozentige Vollständigkeit erreichbar. Die Korrektheit der Analyse liegt bei 91,74 % (siehe Abb. 6).

Insgesamt ist bei diesen beiden Analyseverfahren eine deutliche Abhängigkeit von der Auflösung des Oberflächenmodelles zu beobachten. Bei der Originalauflösung von 20 cm sind eine leichte Übersegmentierung und die Detektion von kleineren Nebenmaxima zu beobachten. Bei den vorliegenden TOM und DOM wurden bei einer Auflösung von 50 cm die genausten Ergebnisse hinsichtlich Vollständigkeit und Korrektheit erzielt.

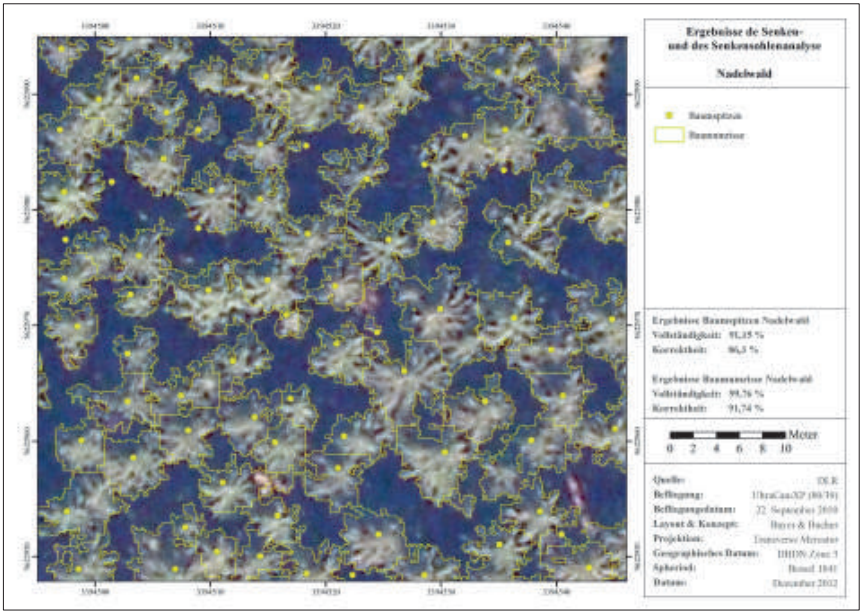


Abb. 6: Ergebnisse der automatischen Einzelbaumextraktion im Nadelwald

5 Konklusion

Wie die Ergebnisse der beiden Verfahren zeigen, können mit Hilfe von photogrammetrisch abgeleiteten Oberflächenmodellen sehr genau die Baumspitzen sowohl im Nadelwald, als auch

im Laubwald abgeleitet werden. Für die Detektion der Baumstandorte kann dabei auf die Verwendung der spektralen Bildinformationen verzichtet werden, da diese keinen Mehrwert für die Extraktion der Baumspitzen darstellen.
Die Bildinformationen lassen sich jedoch für die Extraktion der Baumumrisse und eine mögliche Klassifikation der Baumarten sehr gut nutzen. Mit den beschriebenen Verfahren werden auch bei der automatischen Extraktion der Baumumrisse sehr gute Ergebnisse hinsichtlich Vollständigkeit und Korrektheit erzielt. Zusammenfassend lässt sich damit feststellen, dass eine Einzelbaumextraktion aus photogrammetrisch abgeleiteten Oberflächenmodellen und hochauflösenden TrueOrtho-Mosaiken in Nadel- und Laubwald möglich ist. Mit Hilfe der automatisch extrahierten Daten können so wichtige Informationen, wie beispielsweise Baumhöhe, Kronenfläche oder Baumart, für die Forstinventur abgeleitet werden.
Aufbauend auf den Ergebnissen werden sich die zukünftigen Arbeiten mit der Verbesserung der automatischen Extraktion im Laubwald beschäftigen. Hierbei wird ein Schwerpunkt auf der Einzelbaumextraktion in jüngeren und homogenen Baumbeständen liegen. Zusätzlich ist eine Erweiterung der terrestrisch erhobenen Kontrolldaten geplant, um noch genauere Aussagen über Vollständigkeit und Korrektheit treffen zu können.

6 Danksagung

Die Fernerkundungsdaten wurden für das Projekt „Der Virtuelle Wald“ aufgenommen und prozessiert, welches von der Europäischen Union (European Regional Development Fund - EFRE) und dem Land Nordrhein-Westfalen teilfinanziert wurde. Ein großer Dank gilt an dieser Stelle den Projektbeteiligten des Institutes für Mensch-Maschine-Interaktion (MMI) der RWTH Aachen und des RIF e.V. für die Bereitstellung der Daten und die organisatorische sowie fachliche Unterstützung. Ein Dank geht auch an Geobasis NRW für die kostenfreie Bereitstellung des Digitalen Geländemodelles. Ein besonderer Dank geht auch an ZeDe3D Ltd. für die personelle und technische Unterstützung während der Erhebung der Kontrolldaten.

7 Literaturverzeichnis

ESRI, 2012: ArcGIS Resource Center: Focal Statistics. *Online in Internet: http://help.arcgis.com/de/arcgisdesktop/10.0/help/index.html#//009z000000qs000000.* [Stand: 2012.12-18]

FOURNIER, A. R., EDWARDS, G. & ELDRIGE, N. R., 1995: A Catalogue of Potential Spatial Descriptors for High Spatial Resolution Digital Images of Individual Crowns. Canadian Journal of Remote Sensing, **3** (21), S. 285-298.

HAENEL, S., TRÄNKNER, H. & ECKSTEIN, W.,1987: Automatische Baumkronenentdeckung in Luftbildern - Der Weg durch den Engpaß. Oberpfaffenhofen, S. 53-66.

HIRSCHMÜLLER, H., 2005: Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. – Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 20-26 June 2005, **2**, S.807-814.

HYYPPÄ, J., SCHARDT, M., HAGGRÉN, H., KOCH, B., LOHR, U., SCHERRER, H., PAANANEN, R., LUUKKONEN, H., ZIEGLER, M., HYYPPÄ, H., PYYSALO, U., FRIEDLÄNDER, H., UUTTERA, J., WAGNER, S., INKINEN, M., WIMMER, A., KUKKO, A., AHOKAS, E.,

KARJALAINE, 2001: HIGH-SCAN: The first European-wide Attempt to Derive Single-Tree information from Laserscanner Data. The Photogrammetric Journal of Finnland, 2 (17), pp. 58-68.

Kaartinen, H., Hyyppä, J. 2008: Tree Extraction. EuroSDR / ISPRS Project, Commission II "Tree Extraction". Final Report.

LANDESAMT FÜR NATUR, UMWELT UND VERBRAUCHERSCHUTZ NORDRHEIN-WESTFALEN (LANUV NRW), 2012: Naturschutzgebiet Habichtswald (ST-092). *Online in Internet: http://www.naturschutzinformationennrw.de/nsg/de/fachinfo/gebiete/gesamt/ST_092.* [Stand: 2012.12-18]

LANDSCHAFTSINFORMATIONSSYSTEM DER NATURSCHUTZVERWALDUNG RHEINLAND-PFALZ (LANIS) 2012: Leuscheid. *Online in Internet: http://map1.naturschutz.rlp.de/ landschaften_rlp/landschaftsraum.php?lr_nr=330.00.* [Stand: 2012.12-18]

MAYER, S. 2004: Automatisierte Objekterkennung zur Interpretation hochauflösender Bilddaten in der Erdfernerkundung. Dissertation, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II.

MIRI, M., BAYER, S., BUCHER, T., 2013: Evaluation of high resolution digital surface models for single tree extraction approaches in mixed forests. DGPF Tagungsband 22 /2013 (in Veröffentlichung).

RIF E.V. 2012: Der Virtuelle Wald. *Online in Internet: http://www.virtueller-wald.de/de/ der_virtuelle_wald.* [Stand: 2012.12-18]

SCHIEWE, J., 2002: Segmentation of High-Resolution Remotely Sensed Data-Concepts, Applications and Problems. Proceedings of Symposium on Geospatial Theory, Processing and Applications.

STRAUB, B.-M., 2003: Automatische Extraktion von Bäumen aus Fernerkundungsdaten. Wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover, Nr. 249.

Hochauflösende Datensätze urbaner Oberflächentemperaturen zur Beobachtung der städtischen Wärmeinsel

BENJAMIN BECHTEL[1], KLEMEN ZAKŠEK[2] & JÜRGEN BÖHNER[3]

Zusammenfassung: Trotz zunehmender Bedeutung der städtischen Wärmeinsel ist bislang kein Sensor verfügbar, der ein Monitoring städtischer Oberflächentemperaturen in hoher räumlicher und zeitlicher Auflösung gewährleisten könnte. Einen Lösungsansatz bietet hier das Downscaling von geostationären Messungen. Erste Ergebnisse in urbanen Gebieten erwiesen sich diesbezüglich als sehr vielversprechend und trotz eines großen Skalenfaktors konnten rund 70 % der räumlichen Varianz in 100 m Auflösung erklärt werden. Der mittlere quadratische Fehler lag bei 2.2 – 2.5 K bei einem anteilig recht großes Bias von ~ 1.3 K. In diesem Beitrag wird ein Vergleich möglicher Prädiktoren vorgestellt und das Bias näher untersucht. Es zeigten sich hochsignifikante Unterschiede nach Landnutzungsklassen, wobei diese nicht alleine auf die Höhenstatistik zurückzuführen waren. Eine Schätzung des Bias aus MODIS-Daten war nicht möglich, da diese selbst ein Cold-Bias aufwiesen.

1 Einleitung

Auf Grund von Klimawandel und (Re-)Urbanisierungsprozessen rückt die Überhöhung urbaner Temperaturen gegenüber dem Umland (städtische Wärmeinsel = *urban heat island*, UHI) immer weiter ins Blickfeld. Räumlich und zeitlich hoch aufgelöste Informationen der Temperaturen in urbanen Gebieten wären daher nicht nur zur Vertiefung des Verständnisses steuernder Faktoren von Interesse, sondern im Rahmen eines operationellen Monitorings der Intensität und räumlichen Verteilung der UHI auch von hoher praktischer Relevanz für Stadtplanung und Katastrophenschutz. Während die Lufttemperatur nur *in situ* und damit an wenigen Punkten erfasst werden kann, lässt sich die Erhöhung der Oberflächentemperatur (= *land surface temperature*, LST) in mittlerer Auflösung flächendeckend von satellitengestützten Thermalsensoren erfassen (*Thematic Mapper* auf Landsat 5 in 120 m Auflösung, *Enhanced Thematic Mapper Plus* auf Landsat 7 in 60 m und *Advanced Spaceborne Thermal Emission and Reflection Radiometer* (ASTER) auf Terra in 90 m). Diese Instrumente fliegen auf sonnensynchronen polaren Orbits und haben daher nur zwei Aufnahmezeitpunkte (tags und nachts) und eine lange Wiederkehrdauer von 16 Tagen. Diese zeitliche Auflösung ist für ein UHI-Monitoring völlig inadäquat, zumal die Datenverfügbarkeit zusätzlich durch Wolken eingeschränkt wird. Sensoren auf geostationären Plattformen (z. B. der *Spinning Visible Infra-Red Imager* (SEVIRI) auf Meteosat-8) können dagegen den Tagesgang gut abbilden, haben dafür jedoch eine geringe räumliche Auflösung (~3-5 km), in der sich städtische Strukturen kaum differenzieren lassen und die daher keine stadtklimatischen Untersuchungen ermöglicht. Dazwischen gibt es Instrumente auf polaren Orbits mit breiterem Swath (z.B. *Moderate Resolution Imaging Spectroradiometer* (MODIS) auf Terra und Aqua mit ~ 1 km räumlicher Auflösung und täglicher Wiederkehr).

[1] Benjamin Bechtel, KlimaCampus, Uni Hamburg, benjamin.bechtel@uni-hamburg.de
[2] Klemen Zakšek, Centre of excellence Space-Si, Ljubljana, Slovenia, klemen.zaksek@zmaw.de
[3] Jürgen Böhner, KlimaCampus, Uni Hamburg, boehner@geowiss.uni-hamburg.de

Eine Möglichkeit, die jeweiligen Defizite zu überwinden, besteht in der Kombination dieser Datenströme mit Hilfe von statistischen Downscaling-Verfahren zu einem räumlich sowie zeitlich ausreichend hoch aufgelöstem Datensatz. Dabei wird zwischen geeigneten Prädiktoren und der Zielgröße (hier Oberflächentemperatur) ein Modell in niedriger räumlicher Auflösung kalibriert und unter Annahme eines skalenunabhängigen Zusammenhangs auf eine höhere räumliche Auflösung übertragen. Die Überprüfung erfolgt jeweils mit einer unabhängigen Messung in der höher aufgelösten Domäne. Diese Methode wurde bereits häufig zum Downscaling von Oberflächentemperaturen verwendet, wobei sehr unterschiedliche Auflösungen und Prädiktoren zum Einsatz kamen (KUSTAS et al. 2003, INAMDAR & FRENCH 2009, ZAKŠEK & OŠTIR 2011). In BECHTEL et al. (2012) wurde eine größere Anzahl von Prädiktoren im Hinblick auf ihre Eignung getestet und mit einem Downscaling von SEVIRI zu ASTER erstmals gezeigt, dass auch bei einem Skalenfaktor von > 2000 brauchbare Ergebnissen zu erzielen sind. Daneben stellte sich heraus, dass die vorhergesagten Temperaturen im Mittel zu warm waren, wohinter Unterschiede in der Aufnahmegeometrie vermutet wurden. In diesem Beitrag soll daher neben einer Zusammenfassung vorheriger Ergebnisse untersucht werden, inwiefern das Bias von der Landnutzung abhängig ist und ob er sich aus einem weiteren Datensatz (hier MODIS) schätzen lässt um dadurch die Methode zu verbessern.

2 Daten

Als Untersuchungsgebiet wurde der südliche Teil Hamburgs ausgewählt (53.38 – 53.63° Nord, 9.75 – 10.38° Ost). Es umfasst die Elbe (von Süd-Ost nach Nord-West), ländlichen Gebiete, Wälder im Südwesten und Osten und den größeren Teil der Stadt Hamburg. Hamburg hat eine Bevölkerung von etwa 1.8 Millionen Einwohner und liegt in der norddeutschen Tiefebene nahe von Nord- und Ostsee. Als Zeitpunkt für die Fallstudie wurde der 2.8.2011, ~ 10:30 Uhr UTC (entspricht 12:30 MESZ) gewählt, da hierfür Aufnahmen aller drei Sensoren verfügbar waren (siehe Abb. 1).

2.1 Oberflächentemperaturen

Wie eingangs erwähnt stehen grundsätzlich Sensoren auf polaren Orbits unterschiedlicher Auflösung sowie Sensoren im geostationären Orbit zur Verfügung. Für diese Studie wurden SEVIRI (~ 5 km), MODIS (~ 1000 m) und ASTER (~ 100 m) gewählt. Eine sehr gute Übersicht aller verfügbaren Sensoren geben TOMLINSON et al. (2011).

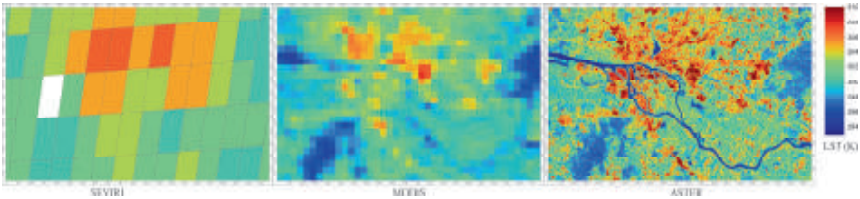


Abb. 1: Verwendente Oberflächentemperaturen in Kelvin in unterschiedlichen Auflösungen. Links: SEVIRI (~ 5 km), Mitte: MODIS (~ 1 km), rechts: ASTER (~ 100 m).

2.1.1 SEVIRI

Die LST in niedriger räumlicher Auflösung zur Kalibrierung stammt vom SEVIRI auf Meteosat-8. Die räumliche Auflösung beträgt für die Fallstudie in der Gegend von Hamburg etwa 3.3 km in Ost-West und 6.7 km in Nord-Süd Richtung, die zeitliche Auflösung ist 15 Minuten. Konkret wurde das operative LSA SAF Produkt verwendet. Zur Berechnung der Oberflächentemperatur kommt ein *generalized split-window* Algorithmus mit den Bändern 10.8 und 12.0 µm [WAN & DOZIER 1996] zur Anwendung. Die Maskierung von Wolkenpixeln basiert auf der *Nowcasting and Very Short Range Forecasting* SAF Software. Die Qualität der abgeleiteten LST hängt neben der Wolkenklassifikation auch stark von der Sensor-Performance, der Genauigkeit der atmosphärischen Korrektur und der Emissivität ab.

2.1.2 MODIS

Als Datensatz in 1000 m Auflösung wurde das MODIS Level 2 LST Produkt in Version 5 verwendet, das ebenfalls auf einem *generalized split-window* Algorithmus basiert. Die Emissivität für beide Kanäle wird aus der Landnutzung geschätzt, der Fehler für LST wird mit < 1 K angegeben (WAN et al. 2004). Das Instrument fliegt auf dem Satelliten Terra, der Teil des NASA Earth Observing System (EOS) ist und etwa 15 Minuten nach Landsat 7 im gleichen Orbit fliegt. Das Swath-Produkt wurde auf Grundlage der mitgelieferten *geolocation fields* mit *piecewise cubic interpolation* georeferenziert. Anschließend wurde aus den Einzelpunkten ein Shapefile erstellt, das anschließend zu UTM Zone 32 projiziert und mit *multilevel B-spline* Interpolation auf ein regelmäßiges 1000 m Raster resampled wurde. Die MODIS-Daten wurden sowohl zur Kalibrierung als auch zur Validierung verwendet.

2.1.3 ASTER

Die hochauflösenden LST Daten zur Validierung stammen vom ASTER Instrument (ebenfalls an Bord des Satelliten Terra). ASTER misst fünf thermische Bänder (zwischen 8 und 11 µm) bei einer horizontalen Auflösung von 90 m. Das Oberflächentemperatur-Produkt wird mit dem *temperature and emissivity separation algorithm* (GILLESPIE et al. 1998) berechnet. ASTER nimmt nur auf Bestellung Daten auf, was zu einer geringeren Verfügbarkeit von entsprechenden Szenen führt. Diese wird im gewählten Untersuchungsgebiet weiterhin durch die Wolkenklimatologie stark eingeschränkt, so das ein Request über den gesamten Sommer 2011 nur zur einer einzigen geeigneten wolkenfreien Szene führte. Zur Georeferenzierung der SWATH-Daten wurde zunächst die mitgelieferte *geometric correction table*, die geographische Koordinaten für bestimmte Pixel enthält, stückweise kubisch interpoliert. Anschließen wurden die Einzelmessungen zu einem Vektordatensatz konvertiert und in UTM Zone 32 projiziert. Die projizierten Daten wurden zunächst in 15 m Auflösung mit *nearest neighbour* interpoliert und anschließend flächengemittelt auf das 100 m Zielraster resampled. Diese Methode garantiert einen minimalen Datenverlust und stellte sich gegenüber der ebenfalls getesteten Multilevel B-Spline-Interpolation als vorteilhaft heraus. Für die Präprozessierung kam das *open source* GIS SAGA (http://saga-gis.org/) zum Einsatz.

2.2 Prädiktoren

In vorherigen Studien stellen sich aggregierte Parameter aus multitemporalen thermischen Infrarotdaten als besonders geeignete Prädiktoren heraus (BECHTEL et al. 2012, BECHTEL et al. 2013). Daher sollen diese hier kurz vorgestellt werden, auch wenn im Folgenden noch kurz auf die Eignung weiterer Prädiktoren eingegangen wird.

2.2.1 Annual cycle parameters

Annual cycle parameters (ACP) beschreiben in Anlehnung an bekannte Tagesgangmodelle (beispielsweise GÖTTSCHE & OLESEN 2009) den Jahresgang der Oberflächentemperatur. Dafür werden mit einer quadratischen Minimierung eine Sinusfunktion als vereinfachtes Modell an die multitemporalen aus Landsat Daten abgeleiteten Oberflächentemperaturen aus einem längeren Zeitraum (insgesamt 26 Szenen zwischen 2008 und 2010) gefittet und daraus die Parameter *mean annual surface temperature* (MAST) und *yearly amplitude of surface temperature* (YAST) sowie optional eine Phasenverschiebung θ abgeleitet (BECHTEL 2011, BECHTEL 2012). Zur Erstellung der Oberflächentemperaturen wurde dabei der NASA *Atmospheric Correction Parameter Calculator* (BARSI et al. 2005) verwendet, der die aufwärts und abwärtsgerichtete Strahldichte sowie die Transmissivität der Atmosphäre zum Aufnahmezeitpunkt mit Hilfe des Strahlungstransfercodes MODTRAN 4 und raumzeitlich interpolierter NCEP Reanalysedaten bestimmt. Die räumlichen Muster von YAST und MAST enthalten somit aggregierte Informationen über die thermischen Eigenschaften der Oberflächenmaterialien unter verschiedenen Einstrahlungsbedingungen.

2.2.2 Thermische Hauptkomponenten

Eine weitere Möglichkeit zur Reduktion der Dimensionalität multitemporaler Thermalmuster ist eine Hauptkomponentenanalyse. Diese rotiert den Datensatz im Merkmalsraum derart, dass der größte Teil der Varianz in der ersten Komponente liegt während die weiteren Komponenten abnehmenden Informationsgehalt haben. Die erste Hauptkomponente kann damit als dominantes (saisonales) Erwärmungsmuster betrachtet werden. Zur Analyse wurden insgesamt 32 Landsat-Szenen verwendet. Auf eine Kalibrierung zu Strahldichte und atmosphärische Korrektur wurde hier verzichtet, da für die Hauptkomponenten nur die dominanten Muster erfasst werden müssen.

2.2.3 Upscaling der Prädiktoren

Zum Downscaling müssen zunächst die Prädiktoren auf das Raster der niedrig aufgelösten Zielgröße gebracht werden. Für das 1000 m Raster wurden sie dafür einfach mit einem flächengemittelten Ansatz resampled. Zur Simulation der SEVIRI-Pixel und ihrer "irregulären" Form auf Grund der Aufnahmegeometrie wurden die Prädiktoren dagegen mit der *point spread function* des Sensors gewichtet, die durch eine Gauß-Funktion approximiert wurde. Prädiktoren im Zentrum eines SEVIRI-Pixels gehen entsprechend stärker ein, da der Sensor gegenüber der Oberflächentemperatur in diesem Bereich ebenfalls sensitiver ist.

2.3 Local climate zones

Die verwendeten Landnutzungsdaten stammen aus einer automatischen Klassifikation von Stadtstrukturen (BECHTEL & DANEKE 2012). Dabei kam die von STEWART & OKE (2009, 2012)

vorgeschlagene Typologie der *local climate zones* (LCZ) zum Einsatz, die Landschaften in lokaler Skala gemäß ihren stadtklimatischen Eigenschaften einteilt. Um nur zusammenhängende homogene Gebiete einer Mindestgröße zu berücksichtigen, wurde die Zugehörigkeit zu jeder Klasse durch eine morphologische Öffnung mit einem kreisförmigen Strukturelement mit 3 Pixeln Radius gefiltert.

3 Downscaling von Oberflächentemperaturen

3.1 Eignung von Prädiktoren

Die Eignung unterschiedlicher Prädiktorensets wurde von BECHTEL et al. (2012) untersucht. Die Ergebnisse sind in Tab. 1 dargestellt. Bereits erwähnt wurde, dass sich aggregierte Parameter aus multitemporalen Thermaldaten besonders eignen. Weiterhin ist festzuhalten, dass die Größe des Prädiktorensets kritisch ist und möglichst klein gewählt werden sollte. Andernfalls kommt es zu einem Overfit in der niedrig und einem großen Fehler in der hoch aufgelösten Domäne.

Tab. 1: Genauigkeit des Downscalings von SEVIRI zu ASTER mit linearer Regression für verschiedene Prädiktorsets. Anzahl der Prädiktoren (Num), erklärte Varianz des Modells (R^2) in der niedrig aufgelösten Domäne; mittlerer absoluter Fehler (MAE), mittlerer quadratischer Fehler (RMSE), R^2 und Bias für die unabhängige Validierung mit ASTER Daten in 100 m Auflösung (nach BECHTEL et al. 2012).

Prädiktorset		Kalibr. (SEVIRI)	Validierung (ASTER)			
Name	**Num**	**R^2**	**MAE (K)**	**RMSE (K)**	**R^2**	**BIAS (K)**
Vordefinierte Sets						
heighstat	6	0.72	4.31	6.09	0.04	1.40
LUCshare	3	0.63	2.23	2.96	0.33	1.33
NDVI	32	0.82	6.40	9.15	0.00	1.46
TIR	32	0.87	3.36	4.46	0.11	1.42
Einzelne Prädiktoren						
bestNDVI[lres]	1	0.36	2.51	3.73	0.03	1.36
bestTIR[lres]	1	0.63	2.08	2.63	0.53	1.32
Aggregierte multitemporeale Thermaldaten						
MAST	**1**	**0.47**	**1.76**	**2.17**	**0.71**	**1.29**
ACP	2	0.59	1.82	2.21	0.68	1.31
tirpca1	**1**	**0.58**	**1.79**	**2.17**	**0.70**	**1.31**
tirpca3	3	0.62	1.89	2.37	0.63	1.37
tirpca5	5	0.63	1.89	2.42	0.61	1.38
Expertenauswahl						
expert1	7	0.70	1.86	2.35	0.63	1.35
expert2	2	0.51	1.85	2.47	0.57	1.34
expert3	2	0.59	1.82	2.31	0.65	1.33
Automatische Merkmalsauswahl						
MRMR2	2	0.64	1.98	2.51	0.54	1.31
MRMR6	6	0.70	2.03	2.64	0.52	1.35
MRMR10	10	0.72	2.12	2.91	0.43	1.35
ForwSel	5	0.80	3.07	4.18	0.19	1.42

Auffällig ist zudem das relativ einheitliche Bias von etwa 1.3 K, der einen großen Teil des gesamten Fehlers ausmacht. Dieser ist auch in Abb. 2 (rechts) ersichtlich (Dominanz hellroter Farbtöne), wo der Fehler des Downscalings mit verschiedenen Prädiktorensets dargestellt ist. Es wurde vermutet, dass dieses Bias aus den unterschiedlichen Beobachtungsgeometrien resultiert (SEVIRI sieht aus dem geostationären Orbit über dem Äquator einen größeren Anteil südorientierter Flächen, die sich schneller erwärmen). Daraus ergeben sich die Fragen, ob durch eine unabhängige Schätzung des Bias die Genauigkeit verbessert werden kann und ob dieser darüber hinaus Informationen über die Wärmespeicherung in vertikalen Objekten (insbesondere Wänden) enthält. Diesen wird im Folgenden durch erste Untersuchungen nachgegangen.

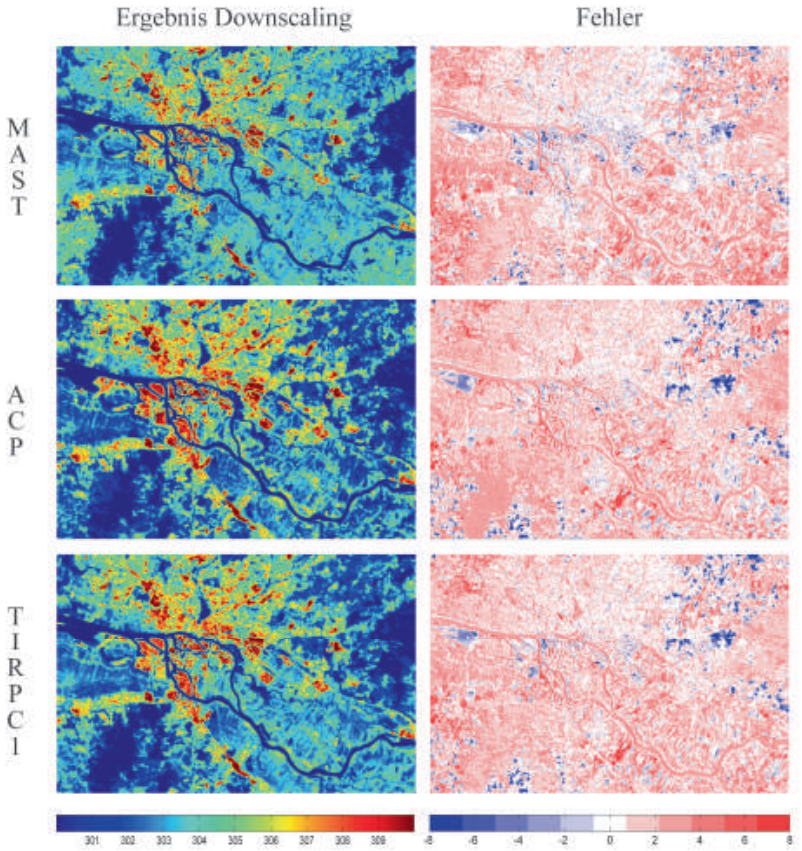


Abb. 2: Downscaling von SEVIRI zu 100 m. Links: Modellierte LST-Muster in Kelvin mit unterschiedlichen Prädiktorensets: *mean annual surface temperature*, *annual cycle parameters*, erste Hauptkomponente aus multitemporalen TIR Daten. Rechts: Fehler gegenüber ASTER-Testdaten in Kelvin.

3.2 Untersuchung des Bias

Zunächst wurde untersucht, inwiefern das Bias von der Landnutzung abhängt. Das Ergebnis ist in Abb. 3 dargestellt und lässt zunächst keine eindeutigen Schlüsse zu. Einerseits sind die Unterschiede in der mittleren Abweichung hochsignifikant und auch zwischen vielen einzelnen Klassen signifikant. Somit erklärt die Landnutzung einen gewissen Teil der Varianz in den Residuen. Andererseits ist die Streuung ebenfalls erheblich und die Vermutung, dass Klassen mit großem Anteil vertikaler Flächen höhere Residuen aufweisen bestätigt sich zunächst nicht. So lassen sich die vergleichsweise hohen Residuen im Wald zwar gut mit der komplexen Topographie des Kronendaches in Verbindung bringen, in den dichten urbanen Klassen *urbdens* und *urbcore* werden die LSTs dagegen nur wenig überschätzt. Dies ist auch aus der räumlichen Verteilung der Residuen in Abb. 2 sichtbar ist (helleres rot im dicht besiedelten Stadtgebiet). Eine mögliche Erklärung wäre, dass bei diesen Klassen SEVIRI einen geringeren Anteil an Straßen sieht, die sich ebenfalls durch sehr hohe Oberflächentemperaturen auszeichnen. Die Korrelation der Residuen mit Parametern der Höhenstatistik aus interferometrischen Radardaten war durchweg gering, was ebenfalls darauf hindeutet, dass der Anteil vertikaler Flächen nicht der alleinig ausschlaggebende Faktor ist.

Im zweiten Schritt wurde untersucht, ob sich das Bias aus MODIS-Daten abschätzen lässt. Dafür wurde das Downscaling nicht nur zwischen SEVIRI und ASTER sondern auch zwischen SEVIRI und MODIS und MODIS und ASTER durchgeführt. Das Ergebnis ist in Tab. 2 dargestellt. Überraschender Weise ergaben sich trotz der wesentlich geringeren Faktoren keine deutlich besseren Ergebnisse. Während die erklärte Varianz in beiden Fällen etwas höher lag, waren die mittleren Fehler für MODIS zu ASTER gleich bis etwas niedriger und für SEVIRI zu MODIS sogar höher. Dies liegt an einem höheren Bias, das beim Downscaling von MODIS zu SEVIRI den größten Teil des Gesamtfehlers ausmacht.

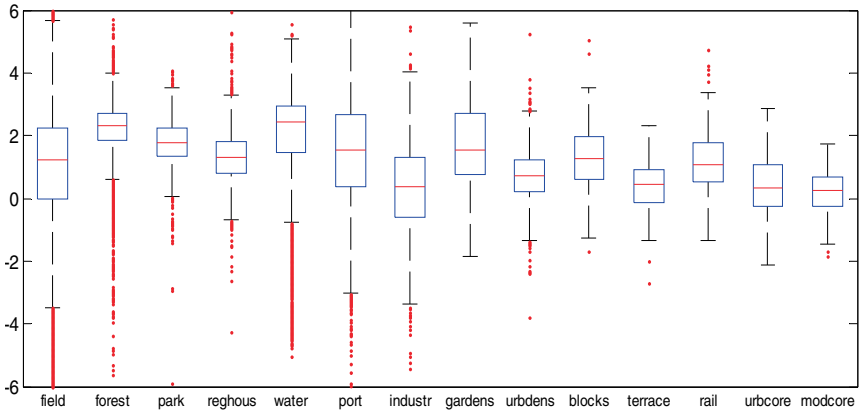
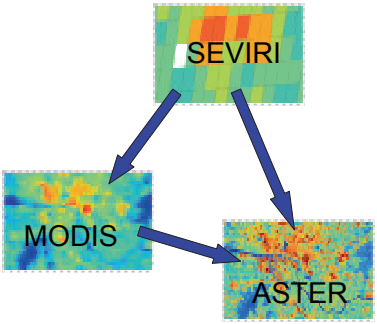


Abb. 3: Bias beim Downscaling von SEVIRI zu ASTER nach Landnutzung in Kelvin.

Tab. 2: Downscaling-Ergebnisse der Fallstudie für unterschiedliche Auflösungen.



	Kalibrierung	*Validierung*			
Prädiktorset	**R^2**	**MAE**	**RMSE**	**R^2**	**BIAS**
SEVIRI zu ASTER					
ACP	0.59	1.82	2.21	0.68	1.31
tirpc1	0.58	1.79	2.17	0.70	1.31
tirpc1&3	0.59	1.74	2.14	0.73	1.34
MAST	0.47	1.76	2.17	0.71	1.29
SEVIRI zu MODIS					
ACP	0.59	2.68	2.87	0.71	2.64
tirpc1	0.58	2.68	2.86	0.73	2.65
tirpc1&3	0.59	2.69	2.87	0.75	2.68
MAST	0.47	2.67	2.88	0.71	2.66
MODIS zu ASTER					
ACP	0.47	1.59	2.11	0.75	-1.35
tirpc1	0.76	1.64	2.22	0.70	-1.35
tirpc1&3	0.73	1.66	2.15	0.73	-1.34
MAST	0.75	1.70	2.22	0.71	-1.36

Dies ergibt sich auch aus Abb. 4, in der der lineare Zusammenhang zwischen verschiedenen Prädiktoren und der Zielgröße LST in den drei Auflösungen dargestellt ist. Es ist klar erkennbar, dass jeweils ein linearer Zusammenhang besteht (während die Prädiktoren YAST, MAST und erste thermische Hauptkomponente eine hohe Korrelation aufweisen, ist diese erwartungsgemäß für die vierte Hauptkomponente sehr viel geringer) und die Steigung zwischen den Auflösungen übertragbar ist. Gleichzeitig weisen diese aber eine unterschiedlichen Mitteltemperatur (bzw. Achsenabschnitt) auf. Diese relative „Unterkühlung“ von MODIS lässt sich nicht aus der Aufnahmegeometrie erklären, da der Sensor auf der gleichen Plattform wie ASTER fliegt und sich diese daher nicht grundlegend unterscheiden sollten. Wahrscheinlicher sind entsprechend Fehler durch die unterschiedlich komplexen Atmosphärenkorrekturen sowie Emissivität. In jedem Fall demonstriert das vorliegende Beispiel, dass sich das Bias nicht ohne weiteres aus MODIS schätzen lässt und neben der Aufnahmegeometrie weitere Aspekte eine wichtige Rolle spielen, die einer vertieften Untersuchung bedürfen.

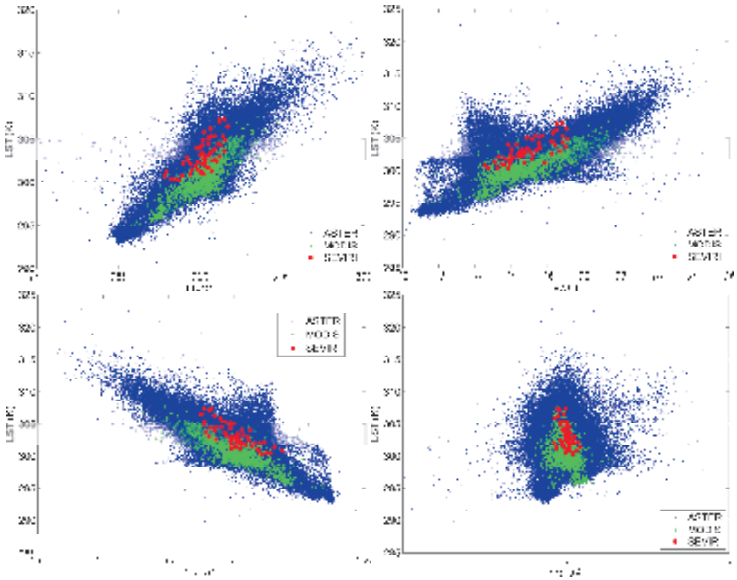


Abb. 4: Linearer Zusammenhang zwischen Prädiktoren und Oberflächentemperatur in unterschiedlichen Auflösungen. Oben MAST und YAST, unten Hauptkomponenten aus multitemporalen TIR Daten (tirpca1, tirpca4). ASTER (100 m) in blau, MODIS (1000 m) in grün, SEVIRI (~5) km in rot.

4 Fazit & Ausblick

Es wurde gezeigt, dass ein Downscaling von geostationären Oberflächentemperaturen in urbanen Gebieten möglich ist. In ersten Fallstudien konnten damit trotz eines Skalenfaktors von über 2000 rund 70 % der räumlichen Varianz in 100 m Auflösung erklärt werden und der mittlere quadratische Fehler lag bei 2.2 – 2.5 K. Angesichts der Genauigkeit der SEVIRI-LST von 2 K ist dieses Ergebnis im Hinblick auf ein operationelles Downscaling zur Beobachtung der Oberflächenwärmeinsel sehr vielversprechend. Weiterhin wurde die Eignung verschiedener Prädiktoren verglichen, wobei sich aggregierte Parameter aus multitemporalen Thermaldaten als besonders geeignet erwiesen. Ergänzende Untersuchungen zeigten, dass sich das Bias zwar signifikant nach Landnutzung unterscheidet, jedoch nicht einfach mit der jeweiligen Häufigkeit vertikaler Flächen zu erklären ist. Schließlich wurde der Frage nachgegangen, ob das Bias aus MODIS Daten geschätzt werden kann. Dabei wiesen die MODIS Daten jedoch selbst ein Cold-Bias auf, das nicht aus der Aufnahmegeometrie erklärt werden kann. Beide Ergebnisse deuten darauf hin, dass diverse Einflussfaktoren das Bias bestimmen. Daher wird eine umfassendere Untersuchung mit einer größeren Anzahl von Aufnahmen und Standorten angestrebt.

5 Literaturverzeichnis

BARSI, J. A.; SCHOTT, J. R.; PALLUCONI, F. D. & HOOK, S. J., 2005: Validation of a web-based atmospheric correction tool for single thermal band instruments. Earth Observing Systems X, Proc. SPIE, **58820E**, S. 1–7.

BECHTEL, B., 2011: Multisensorale Fernerkundungsdaten zur mikroklimatischen Beschreibung und Klassifikation urbaner Strukturen. Photogrammetrie-Fernerkundung-Geoinformation, **2011**(5), S. 325–338.

BECHTEL, B., 2012: Robustness of annual cycle parameters to characterize the urban thermal landscapes. Geoscience and Remote Sensing Letters, IEEE, **9**(5), S. 876 –880.

BECHTEL, B.; BÖHNER, J.; ZAKŠEK, K. & WIESNER, S. Downscaling of diurnal land surface temperature cycles for urban heat island monitoring, in: Urban Remote Sensing Event (JURSE), 2013 Joint. IEEE. Accepted.

BECHTEL, B. & DANEKE, C., 2012: Classification of local climate zones based on multiple earth observation data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, **5**(4), S. 1191 –1202.

BECHTEL, B., ZAKŠEK, K. & HOSHYARIPOUR, G., 2012: Downscaling land surface temperature in an urban area: a case study for hamburg, germany. Remote Sensing, **4**(10), S. 3184–3200.

GILLESPIE, A.; ROKUGAWA, S.; MATSUNAGA, T.; COTHERN, J. S.; HOOK, S. & KAHLE, A. B., 1998: A temperature and emissivity separation algorithm for advanced spaceborne thermal emission and reflection radiometer (ASTER) images. IEEE Transactions on Geoscience and Remote Sensing, **36**, S. 1113–1126.

GÖTTSCHE, F. M. & OLESEN, F. S., 2009: Modelling the effect of optical thickness on diurnal cycles of land surface temperature. Remote Sensing of Environment, 113(11), S. 2306–2316.

INAMDAR, A. K. & FRENCH, A., 2009: Disaggregation of goes land surface temperatures using surface emissivity. Geophysical Research Letters, **36**(2), p. L02408. Retrieved December 13, 2012,

KUSTAS, W. P.; NORMAN, J. M.; ANDERSON, M. C. & FRENCH, A. N., 2003: Estimating subpixel surface temperatures and energy fluxes from the vegetation index–radiometric temperature relationship. Remote Sensing of Environment, **85**(4), S. 429–440.

Tomlinson, C. J.; Chapman, L.; Thornes, J. E. & Baker, C., 2011: Remote sensing land surface temperature for meteorology and climatology: a review. *Meteorological Applications*, *18*(3), pp. 296–306.

STEWART, I. D. & OKE, T. R., 2009: Newly developed "thermal climate zones" for defining and measuring urban heat island magnitude in the canopy layer. Preprints, T.R. Oke Symposium & Eighth Symposium on Urban Environment, January 11–15, Phoenix, AZ.

STEWART, I. D. & OKE, T. R., 2012: "local climate zones" for urban temperature studies. Bulletin of the American Meteorological Society, doi: 10.1175/BAMS-D-11-00019.1.

WAN, Z.; ZHANG, Y.; ZHANG, Q. & LI, Z.-L., 2004: Quality assessment and validation of the modis global land surface temperature. Int. Journal of Remote Sensing, **25**(1), S. 261–274.

WAN, Z. & DOZIER, J., 1996: A generalized split-window algorithm for retrieving land-surface temperature from space. IEEE Transactions on Geoscience and Remote Sensing, **34**(4), S. 892 –905.

ZAKŠEK, K. & OŠTIR, K., 2012: Downscaling land surface temperature for urban heat island diurnal cycle analysis. Remote Sensing of Environment, **117**, S. 114–124.

Automatische Berechnung einer nationalen Gehölzmaske

RUEDI BOESCH[1] & CHRISTIAN GINZLER[2]

Im Rahmen des Schweizerischen Landesforstinventar (LFI) soll eine flächendeckende Gehölzmaske berechnet werden. Als Datengrundlage dienen ADS80 Luftbildstreifen mit 25cm/50cm Bodenauflösung, welche von der swisstopo operationell alle 3 Jahre aufgenommen werden. Aus dem CIR-Komposite der Luftbildstreifen ist mittels Stereo-Korrelation ein digitales Oberflächenmodell der Schweiz erzeugt worden, welches in Kombination mit dem digitalen Bodenmodell der swisstopo die Berechnung eines Kronenhöhenmodells mit 1m Auflösung erlaubt.

Ausgehend von diesem Kronenhöhenmodell wird mit einem robusten geometrischen Modell nach Gehölzkronen gesucht. In der lokalen Umgebung der gefundenen Kandidaten von Gehölzkronen wird mit einer radial gewichteten Suchmaske eine Statistik der spektralen Eigenschaften berechnet. Ausgehend von den gefundenen Zentrumspunkten der Gehölzkronen werden mit einer distanz-basierten Triangulation die Einzelobjekte gruppiert. Die gefundene Gruppierung wird anschliessend mit einem Verfahren basierend auf Alpha Shapes zu einer Gehölzfläche generalisiert, wobei die Generalisierung im Sinne einer vorgegebenen Walddefinition parametrisierbar ist. Dadurch lassen sich komplette Gehölzmasken mit unterschiedlichen Konnektivitäten berechnen.

1 Einleitung

Die Datenerfassung im Schweizerischen Landesforstinventar (LFI) setzt sich aus mehreren Erfassungsschritten zusammen. Primär handelt es sich um eine bodengestützte Inventur, welche mit systematischen Stichproben in einem Netz mit 1.4 km Maschenweite verschiedene Parameter wie Baumanzahl, Stammdurchmesser, Totholz, etc. erfasst (BRASSEL 2001). Vorgängig zur terrestrischen Erhebung wird eine Luftbildinterpretation in einem 500 m Netz durchgeführt, sodass anhand der stereoskopischen Luftbildauswertung entschieden werden kann, ob die Stichprobe im Wald liegt, da nur solche terrestrisch erhoben werden. Neben der Vorselektion wird unter anderem auch die Bestandeshöhe der Stichprobe auf einer Fläche von 50 x 50 m mit einem 5 x 5 Punkteraster geschätzt.

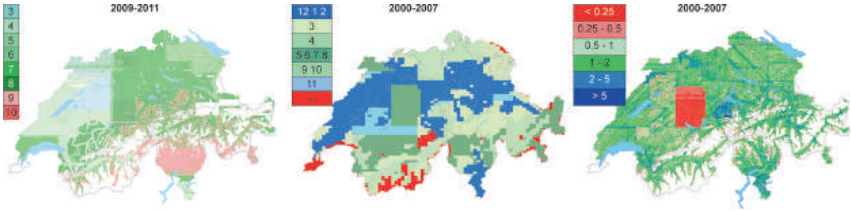
Für die statistische Auswertung auf grossregionaler und nationaler Ebene genügt diese Erhebungsmethodik. Der vermehrte Bedarf an kleinräumigen Auswertungen zeigt, dass die kleine Fläche der Stichprobe zu signifikanten Schätzfehlern führt. Das gesamte Luftbild-Stichprobennetz deckt nur 1% der Gesamtfläche der Schweiz ab.

Daher ist es notwendig aus den vorhandenen nationalen Luftbildern und Höhenmodellen als Ergänzung zu der bestehenden Stichproben-Inventur auch flächenhafte Produkte zu erzeugen. In erster Priorität soll daher eine nationale Gehölzmaske für den Zeitraum 2008-2011 berechnet werden.

1) Ruedi Boesch, Eidg. Forschungsanstalt für Wald, Schnee und Landschaft WSL, 8903 Birmensdorf, Schweiz, ruedi.boesch@wsl.ch
2) Christian Ginzler Eidg. Forschungsanstalt für Wald, Schnee und Landschaft WSL, 8903 Birmensdorf, Schweiz, christian.ginzler@wsl.ch

2 Ausgangsdaten

Die Prozessierung der Gehölzmaske basiert auf den spektralen und geometrischen Eigenschaften von bereits existierenden digitalen Luftbildern und LiDAR-Daten. Die Kombination der Sensorcharakteristik von LiDAR und spektraler Information erlaubt eine robustere Berechnung als nur mit einem einzigen Sensortyp (NELSON 1997, SONG 2003). Da nationale Datensätze über eine Periode von mehreren Jahren aufgenommen werden, sind diese Rohdaten nie genügend homogen. Variierende Beleuchtung, unterschiedliche Vegetationsperioden, wechselnde Aufnahmeverfahren und Vorverarbeitungsschritte führen zwangsläufig zu regional unterschiedlichen Daten. Daher sind zeitlich und räumlich redundante Datensätze notwendig um den partiellen Ersatz mit besseren Daten aus einem möglichst ähnlichen Zeitraum zu ermöglichen. Für die Berechnung einer Gehölzmaske ist die daraus entstehende zusätzliche Unschärfe des Aufnahmezeitpunktes von einigen Jahren nicht ideal, aber tolerierbar.



a) Aufnahmemonate ADS80 b) Aufnahmemonate LiDAR c) Punktdichte [p/m^2] LiDAR

Abb. 1: Räumliche und zeitliche Verteilung von ADS80 und LiDAR

Die hochaufgelösten ADS80/SH52-Bildstreifen des Bundesamt für Landestopografie (swisstopo) sind gesamtschweizerisch verfügbar. Seit 2008 wird in einem 6-Jahresrythmus die Vegetationsperiode Juni-September landesweit beflogen, innerhalb von 3 Jahren wird die ganze Schweiz räumlich optimiert erfasst.

Da der erste 6-Jahreszyklus noch nicht abgeschlossen ist, müssen auch weniger geeignete Bilder im Frühjahr oder Spätherbst als Ersatz verwendet werden. In Abb. 1a ist die aktuelle Streuung des Aufnahmedatums in den Jahren 2009-2011 ersichtlich. Die weissen Flächen im Hochgebirge bezeichnen Gebiete, welche wegen dem fehlenden LiDAR-Bodenmodell über 2000 m.ü.M. nicht verwendet werden können.

Die nationalen LiDAR-Datensätze des digitalen Oberflächen- und Bodenmodell (DOM-LiDAR und swissALTI3D der swisstopo) sind bis zu einer Höhe von über 2000 m.ü.M flächendeckend verfügbar, weisen aber lokal erhebliche variierende Punktdichten und Aufnahmezeitpunkte auf (Abb. 1b + 1c). Insbesondere die lange Projektdauer mit unterschiedlichen Prozessierungen führt zu grossen lokalen Unterschieden und ist daher als primäre Grundlage für die Berechnung einer hochaufgelösten Gehölzmaske zu inhomogen. In der Folge wird bevorzugt ein eigenes stereo-korreliertes Höhenmodell verwendet (DOM-ADS80, GINZLER 2011), da das vorliegende Bildmaterial flächendeckend, eine doppelte Auflösung aufweist und in der Periode 2008-2011 für einige Regionen bereits mehrfach vorhanden ist. Dabei wird aus dem NIR-Kanal der ADS80-Daten mittels Stereo-Korrelation ein digitales Oberflächenmodell erzeugt.

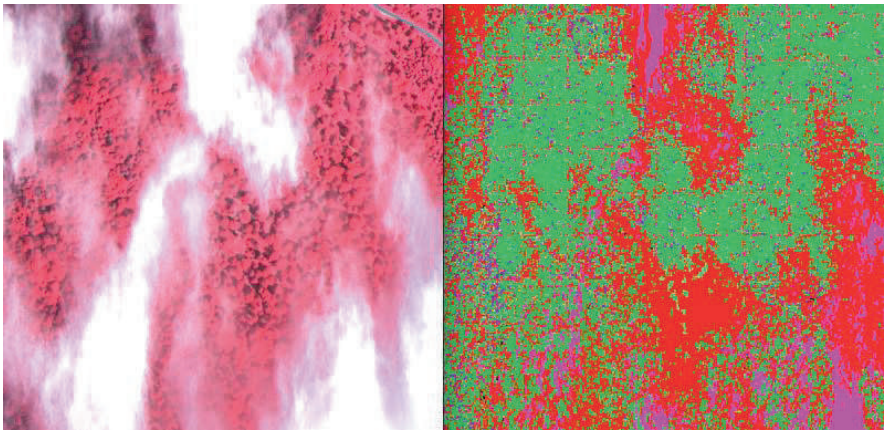


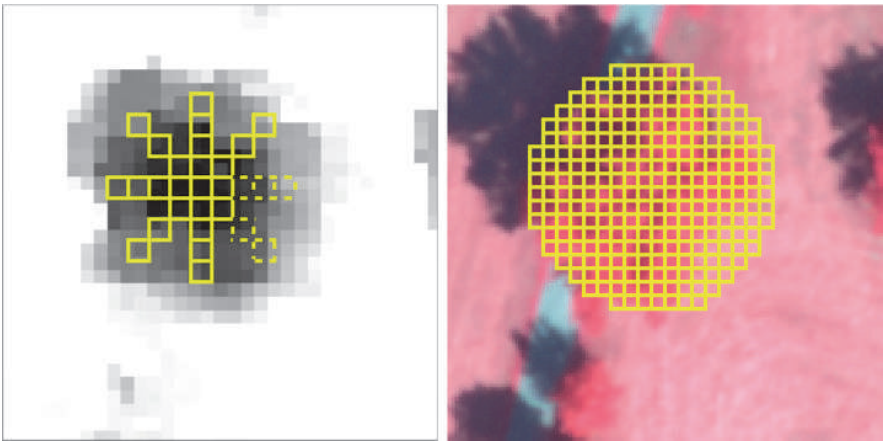
Abb. 2: a) CIR ADS-80 b) FOM aus CIR ADS-80

Die Bildstreifen der swisstopo haben mehrheitlich eine Bodenauflösung von 25 cm, in hochalpinen Gebieten wird wegen der schwierigen Topografie mit 50 cm aufgenommen. Mit der eingesetzten Korrelations-Software SOCET SET NGATE 5.4 und einer „most-Nadir" Strategie bei zeitgleichen Bildstreifen resultiert ein Oberflächenmodell mit 1 m, respektive 2 m Auflösung.
Aus dem DOM-ADS80 und swissALTI3D ist ein Kronenhöhenmodell (CHM) erstellt worden, welches die wichtigste Grundlage für die Berechnung der Gehölzmaske liefert.

3 Methoden

Die Berechnung der Gehölzmaske direkt aus diesem Kronenhöhenmodell liefert nur bedingt brauchbare Ergebnisse (NELSON 1997). Die Auflösung von 1 m, resp. 2 m reicht nicht aus, um Einzelbäume in Beständen zuverlässig zu erfassen und häufig zeugen lokale Artefakte in Oberflächen- oder Bodenmodell von unterschiedlichen Aufnahme- und Prozessierungsfehlern (VAUHKONEN 2010).
Neben dem eigentlichen Kronenhöhenmodell muss daher auch die lokale Qualität für Oberflächen- und Bodenmodell einbezogen werden. Mit Qualität soll im vorliegenden Fall verifiziert werden können, ob die Korrelation des Oberflächenmodells erfolgreich war und ob die Punktdichte des swissALTI3D genügend gross ist.
Im Falle des DOM-ADS80 erzeugt die verwendete Software NGATE eine spezielle rasterbasierte Gütekarte (figure of merits (FOM)) mit gleicher räumlicher Auflösung wie für das berechnete Gebiet. Dabei wird bei einer erfolgreichen Korrelation der Koeffizient geschrieben (Abb. 2b, grün klassiert), bei einem Fehler wird ein FOM-Fehlercode verwendet (Abb. 2b, rot,violett). Für die einzelnen Punkte des DOM und swissALTI3D ist eine Punktdichte-Verteilung mit 10 x10 m Auflösung für die ganze Schweiz berechnet worden (DOM-LiDAR- und DTM-LiDAR-DENS). Anhand dieser beiden Hilfsdatensätze kann die lokale Qualität des CHM numerisch beurteilt werden. Falls beide Hilfsdatensätze am gleichen Ort zu schlecht sind, kann damit auch ein NODATA-Gebiet definiert werden.

In ca. 9 % der korrelierten Fläche verursachen Wolken, Wasserflächen und sehr steiles Gelände grossflächige Korrelationsartefakte, welche auch in den FOM-Daten als zusammenhängende Fehlerflächen zu erkennen sind (rot,violett in Abb. 2b). Mittels einer Formanalyse werden Fehlerflächen mit einer Mindestfläche von 5000 m^2 und einer Zirkularität > 0.05 als signifikant erkannt. Bei Flächen mit signifikant hohem Weiss-Anteil handelt es sich um Bewölkung und in diesem Fall werden die fehlerhaften Gebiete des DOM-ADS80 mit Werten des DOM-LiDAR zu einem synthetischen DOM kombiniert.
Leider sind die FOM-Werte nicht immer zuverlässig, d.h. vereinzelt existieren auch die problematischen Kombinationen von korrektem Luftbild ("true") und "negativem" FOM, resp. fehlerhaftem Luftbild ("false") und korrektem FOM ("positive"), welche nicht automatisch erkannt werden können.
Das oft verwendete pixelweise Schwellwertverfahren zur Extraktion einer Gehölzmaske direkt aus dem Kronenhöhenmodell liefert bei der vorliegenden Auflösung unbefriedigende Resultate. Auch die in der Gelände-Analyse häufig verwendete lokale Krümmung, 2.Ableitung der Höhenwerte in einer 3x3 Umgebung (HODGSON 2005, LIU 2008, ZEVERBERGEN 1987) eignet sich schlecht, weil sehr viele Artefakte, insbesondere in städtischen Gebieten und bei aufgelockerten Beständen entstehen.

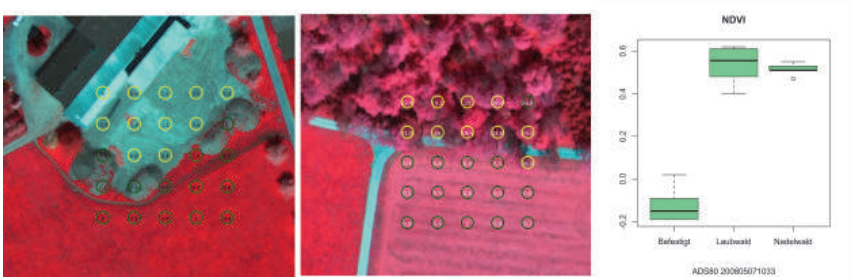


a) CHM mit radialer Suchmaske b) CIR ADS80 mit radialem Histogramm

Abb. 3: Modell-basierte Suchmaske

Aus diesem Grund wird mit einem robusteren modell-basierten Ansatz nach Gehölzkronen gesucht (BOESCH 2010). Dabei wird in einem ersten Schritt ein lokales Höhenmaximum in einer 3 x 3 m Umgebung gesucht. Anschliessend wird in einer erweiterten Umgebung mit einer richtungs-spezifischen Suchmaske (Abb. 3a, gelb) die Neigung der Höhenwerte entlang der 8 Suchrichtungen berechnet. Falls in der Suchrichtung mindestens 2 nachfolgende Neigungswerte zwischen 10 und 80 Grad vorkommen, wird die Suchrichtung als potentielle Kronenform klassiert. Falls die Neigungswert-Statistik in mindestens 6 zusammenhängenden Suchrichtungen erfolgreich ist, wird das gefundene Zentrum mit Umgebung als einzelner

Kronenschluss definiert. Der empirisch gewählte grosse Wertebereich für zulässige Neigungswerte erlaubt den sehr heterogenen Kronenformen bei verschiedenen Baumarten gerecht zu werden. Kleinere Defekte in der Kronenform können mit dem "mindestens 6 aus 8" Kriterium der Neigungswert-Statistik zugelassen werden (Abb. 3a, gelb gestrichelt).
Diese sequentielle Suchstrategie resultiert in einer Punktwolke von Kronenschluss-Zentren. Neben der rein geometrischen Information werden mit einem radial gewichteten Histogramm auch die Spektralwerte des Luftbildes um das gefundene Kronenzentrum erfasst (Abb. 3b). In einem Plausibilitätstest werden spektrale Ausreisser (Farbsättigung und -helligkeit < 20 %) eliminiert.



a) Punktraster "Befestigt" b) Punktraster "Vegetation" c) NDVI-Streuung

Abb. 4: Automatische NDVI-Schwellwertbestimmung

Um Vegetation von Gebäuden möglichst robust zu differenzieren, wird der NDVI des Kronenzentrums verwendet. Unterschiedliche Beleuchtungsverhältnisse und verschiedene Vegetationsperioden der Luftbildstreifen erfordern eine dynamische Bestimmung des optimalen NDVI-Schwellwertes. Mit Hilfe des Punkteraster der LFI-Luftbildinterpretation werden die manuell erhobenen Bodenbedeckungsklassen "Vegetation" und "Befestigte Fläche" mit den Spektralwerten des jeweiligen Luftbildstreifens in einer Umgebung von 5x5m um jeden Rasterpunkt verglichen (Abb. 4a, 4b gelb). Abb. 4c zeigt die gute Differenzierung mit NDVI, beim gewählten Luftbildstreifen 200805071033 wird der Schwellwert daher auf 0.2 festgesetzt.
In einem weiteren Schritt müssen Einzelpunkte mit einem Gruppierungsverfahren zu Polygonen zusammengefasst werden. Dabei werden alle Punkte zuerst mit einem distanzabhängigen Kriterium trianguliert. Mit der maximalen Distanz zwischen einzelnen Punkten kann eine vorgegebene Walddefinition entsprechend modelliert werden. Gemäss der Definition im LFI beträgt die maximale Distanz von Gehölzen bei der Waldbegrenzungslinie 25 m.



Abb. 5: Triangulierte Punktwolken mit Begrenzungslinien (hellgrün)

Die triangulierten Punktwolken werden mit Alpha Shapes (CHANG 2009, EDELSBRUNNER 1995) zu Begrenzungslinien zusammengefasst (Abb. 5). Im Gegensatz zur konvexen Hülle (convex hull) sind mit Alpha Shapes auch partiell konkave Formen möglich, sodass bei Buchten in der Punktwolke die Fläche nicht überschätzt wird. Neben der eigentlichen Begrenzungslinie werden auch Lücken erzeugt, welche konform zur gewählten Abstandsdefinition sind.

4 Prozessierung

Alle Arbeitsschritte erfolgen automatisch und werden mit einer SQL-Tabelle verwaltet. Der einzelne Rechenklient verlangt mittels SQL nach einer freien Kachel zur Berechnung, anschliessend werden die notwendigen Bilddaten direkt über das Filesystem (NFS) gelesen.
Um die parallele Prozessierung zu vereinfachen ist die ganze Schweiz in 500 x500 m grosse Kacheln aufgeteilt worden. Die Ausgangsdatensätze CHM, ADS80 und FOM sind in verschiedenen Vorverarbeitungschritten bereits berechnet und liegen als georeferenzierte Kacheln mit unterschiedlicher Bodenauflösung vor (GINZLER 2011). Eine weitere Vereinfachung zur Prozessierung wird erreicht, indem die gebietsweise reduzierten Auflösungen von 50 cm beim Luftbild auf die höhere Auflösung mittels Resampling angepasst worden sind, sodass alle Luftbild-Kacheln die gleiche Auflösung aufweisen.
Damit keine unerwünschten Randeffekte, z.B. polygonale Randspalten bei geschlossenen Gehölzflächen) entstehen, ist bei der modell-basierten Suche nach Gehölzkronen eine Überlappung von 20 m verwendet worden.
Grundsätzlich ist für die Implementierung der raster- und vektor-bezogenen Methoden (z.B. die modell-basierte Gehölzkronensuche) C++ und die Bibliotheken GDAL/OGR für die georeferenzierten IO-Operationen verwendet worden (GDAL 2012). Für die Formanalyse der FOM-Dateien und andere Bildverarbeitungs-Operationen ist die Bibliothek OpenCV eingesetzt worden (OPENCV 2012). Die distanz-basierte Triangulation und die Alpha Shapes sind mit der Bibliothek CGAL realisiert worden (CGAL 2012). Als Resultat werden für jede Kachel ein Point-Shapefile mit Gehölzkronen und ein Polygon-Shapefile mit der Gehölzmaske erzeugt.
Ausgehend von ca. 15 TB Bilddaten und einem Zeitschnitt von 2008-2011 dauert die vollständige Berechnung mit 30 Prozessoren etwa 8 Tage.

5 Resultate

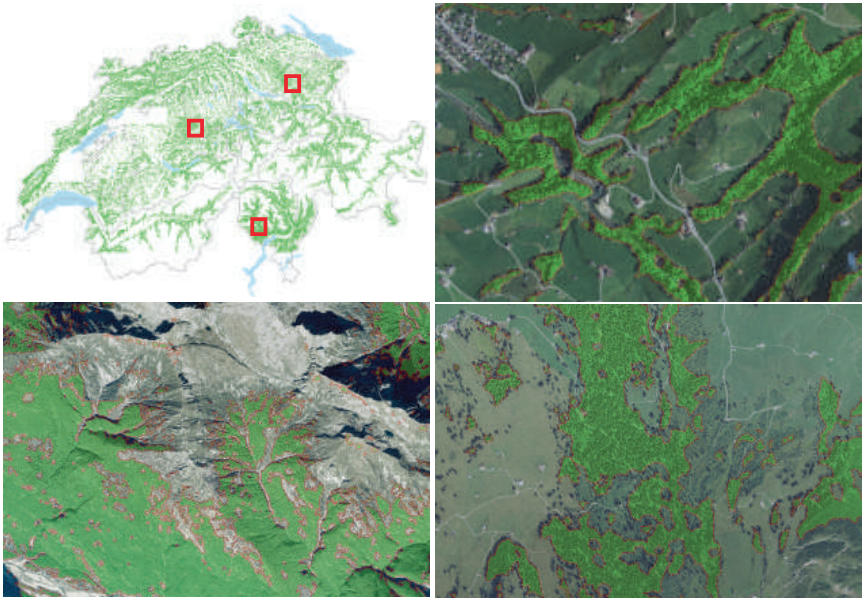


Abb. 6: Gehölzmaske Schweiz mit 3 Ausschnitten (Mittelland, Voralpen, Tessin)

In Abb. 6 sind die fast vollständige Gehölzmaske und 3 repräsentative Ausschnitte dargestellt. Als Hintergrund ist zum qualitativen Vergleich das Orthophotomosaik 50 cm der swisstopo eingeblendet. Zur Verifikation der Gehölzbegrenzungslinie ist der Wald-Entscheid der Luftbildinterpretation des LFI verwendet worden (primär anhand des Deckungsgrades der Punkteraster-Information wird entschieden, ob die Stichprobe im Wald oder ausserhalb liegt).

Tab. 1: Vergleich Luftbildinterpretation mit Gehölzmaske

		LFI-LB		
		Wald	NichtWald	
ADS80	Gehölz	2046	132	2178

Ein Vergleich mit der manuellen Interpretation (Tab. 1) ergibt eine User Accuracy von 94 %, aber bei genauer Betrachtung der 132 NichtWald-Stichproben zeigt sich, dass die Resultate der Luftbildinterpretation und die Gehölzbegrenzungslinie nur beschränkt vergleichbar sind.

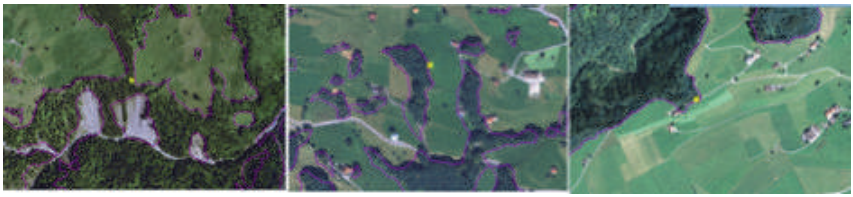


Abb. 7: Stichproben „NichtWald“ im Vergleich mit der Gehölzbegrenzungslinie

Sehr häufig ist das Zentrum der Stichprobe am Rande der Bestockung und wegen dem zu geringen Waldflächenanteil wird die Stichprobe gemäss LFI-Definition als NichtWald definiert (gelbe Stichproben-Zentren in Abb. 7).

6 Literaturverzeichnis

BOESCH, R., C. GINZLER, ET AL., 2010: Robust canopy matching using local shape analysis. ForestSat 2010 Conference, Lugo, Spain.

BRASSEL, P., 2001: Swiss National Forest Inventory: Methods and Models of the Second Assessment. Swiss Federal Research Institute WSL, Birmensdorf.

CHANG, M. C., F. F. LEYMARIE, ET AL., 2009: "Surface reconstruction from point clouds by transforming the medial scaffold." Computer Vision and Image Understanding **113**(11), S. 1130-1146.

CGAL 2012: "Computational Geometry Algorithms Library." from http://www.cgal.org.

EDELSBRUNNER, H., 1995: "Smooth surfaces for multi-scale shape representation." Lecture Notes in Computer Science 1026, S. 391-412.

GDAL 2012: "Geospatial Data Abstraction Library." from http://www.gdal.org.

GINZLER, C., R. BOESCH, ET AL., 2011: High resolution digital surface models for forest monitoring, a photogrammetric approach for large areas. 1st Forestry Workshop: Operational Remote Sensing in Forest Management, Czech Technical University, CTU), Czech Republic, Prague.

HODGSON, M. E., J. JENSEN, ET AL., 2005: "An evaluation of lidar-derived elevation and terrain slope in leaf-off conditions." Photogrammetric Engineering and Remote Sensing 71(7), S. 817-823.

LIU, X. Y., Z. Y. ZHANG, ET AL., 2008: Large area DEM generation using Airborne LiDAR data and quality control. Liverpool, World Acad Union-World Acad Press.

NELSON, R., 1997: "Modeling forest canopy heights: The effects of canopy shape." Remote Sensing of Environment 60(3), S. 327-334.

OPENCV 2012: "Open Source Computer Vision." from http://opencv.willowgarage.com/wiki/.

SONG, C. AND C. E. WOODCOCK, 2003: "Estimating tree crown size from multiresolution remotely sensed imagery." Photogrammetric Engineering and Remote Sensing 69(11), S. 1263-1270.

VAUHKONEN, J., 2010: "Estimating crown base height for Scots pine by means of the 3D geometry of airborne laser scanning data." International Journal of Remote Sensing **31**(5), S. 1213-1226.

ZEVERBERGEN, L. W. AND C. R. THORNE, 1987: "Quantitative Analysis of Land Surface Topography." Earth Surface Processes and Landforms 12, S. 47-56.

Klassifikation von Fernerkundungsdaten zur Früherkennung von Borkenkäferschäden im Bayerischen Wald

MARTIN DRAUSCHKE[1] & PATRICK REIDELSTÜRZ[2]

Zusammenfassung: Im Kontext einer Machbarkeitsstudie zur Früherkennung von Borkenkäferbefall in Fichte wird ein Verfahren zur Klassifikation von Multispektraldaten vorgestellt und diskutiert. Die verwendeten Bilddaten stammen von einer 6-kanaligen Multispektralkamera aus einer Sportflugzeugbefliegung mit einer Flughöhe von 400 m über Grund. In zwei Experimenten werden Merkmale (Vegetationsindices) berechnet, die einerseits auf Erfahrungen in der Präzisionslandwirtschaft und anderseits auf dem Maschinellen Lernen basieren. Zur Klassifikation wurde der effektive Klassifikator Random Forest eingesetzt. Die Erkennung markanter Bäume (Strukturmerkmale) fällt mit einer Erkennungsrate von über 70% zufriedenstellend aus, der gesamte Klassifikationserfolg liegt bei 68% bei Verwendung von zehn Merkmalen bzw. bei 75%, wenn 101 Merkmale ausgenutzt werden.

1 Einleitung und Motivation

Der Borkenkäfer (insbesondere Ips Typographus) ist einer der wirtschaftlich relevantesten Schädlinge in Wirtschaftswäldern und führt allein am Rand des Nationalparks Bayerischer Wald zu einem jährlichen Einschlag von 100 000 Festmeter Fichten-Käferholz.
Borkenkäferbefall wird nach internationalem Sprachgebrauch in drei Befallsstadien unterschieden: (1) *green attack stage* (Erstbefall eines Baumes mit vital erscheinender Baumkrone), (2) *red attack stage* (Rot-Färbung der Baumkrone) und (3) *grey attack stage* (Verlust der Nadelkrone und Tod des Baumes). Ein frischer Befall im *green attack stage* lässt sich äußerlich nicht ohne weiteres von gesunden Fichten unterscheiden.
Diese Arbeit bereitet eine Machbarkeitsstudie vor, in der der Einsatz schmalbandiger Multispektraldaten für die frühzeitige Erkennung befallener Fichten im *green attack* Stadium und deren Unterscheidung von unbefallenen Bäumen untersucht wird. Dazu werden Vegetationsindices analog zu erfolgreichen Ansätzen in der Präzisionslandwirtschaft wie z.B. (ZARCO-TEJADA et al., 2004) und ein effektiver, robuster Klassifikator für mehrere Klassen eingesetzt. Aufgrund fehlender terrestrischer Vergleichsdaten wird in dieser Studie eine UAV gestützte aeriale Waldzustandsinterpretation als Referenz für eine multispektrale Klassifikation herangezogen.

2 Versuchsaufbau

Als Datengrundlage dienen Bilder einer Cessna-Befliegung mit 400 m über Grund über einem Testgebiet im Nationalpark Bayerischer Wald am 5. August 2011. Im Außenbehälter wurde eine

[1] Martin Drauschke, Universität der Bundeswehr München, Institut für Angewandte Informatik, Werner-Heisenberg-Weg 39, 85577 Neubiberg; Email: martin.drauschke@unibw.de
[2] Patrick Reidelstürz, Hochschule Deggendorf, Technologie Campus Freyung, Grafenauer Straße 22, 94078 Freyung; Email: patrick.reidelstuerz@hdu-deggendorf.de

Multispektralkamera „TetraCam Mini MCA 6“ mitgeführt, die mit sechs schmalbandigen Filtern konfiguriert war (Tab. 1). Bei der Filterauswahl wurden diejenigen berücksichtigt, die erfolgreich in der Präzisionslandwirtschaft zur Bestimmung von Vegetationsindices eingesetzt werden. Unter dieser Priorität wurde bewusst auf einen Blau-Kanal verzichtet und damit auf die Möglichkeit, normale Farbbilder ableiten zu können.

Tab. 1: Verwendete Filter für die Erfassung schmalbandiger Intensitätsbilder.

Band	b_1	b_2	b_3	b_4	b_5	b_6
Wellenlänge	550 nm	670 nm	710 nm	780 nm	900 nm	950 nm

Auf der Untersuchungsfläche (Abb. 1) befinden sich zum Aufnahmezeitpunkt neben jungen und älteren unauffälligen Fichten und Buchen auch vom Borkenkäfer befallene Bäume in allen drei Stadien sowie Sturmholz.

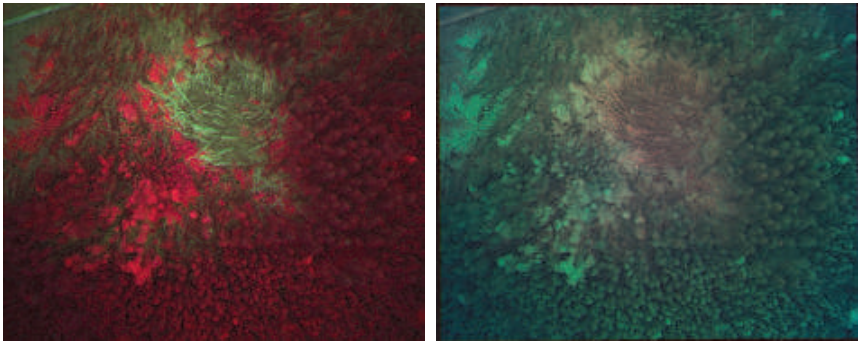


Abb. 1: Untersuchungsfläche im Nationalpark Bayerischer Wald. Links: CIR-Bild mit den Bändern (b_4,b_2,b_1). Rechts: Bildkomposition aus den verbleibenden drei Bändern (b_3,b_5,b_6). Das Untersuchungsgebiet zeigt in der Bildmitte eine abgebrochene Sturmwurfaufarbeitung mit geschältem Holz und oben links einen Sturmwurf. Im linken Bereich befinden sich Bodenbewuchs und Buchen, am unteren Bildrand Fichtenstangenholz. Rechts oberhalb der Aufarbeitungsfläche befinden sich Fichten im red attack stage, dazwischen zum Auswertezeitpunkt nicht näher lokalisierte green-attack-Stadien sowie gesunde Fichten.

Die Kontrollbefliegung mit einem unbemannten Luftfahrzeug (3,2m Motorseglerkonfiguration) erfolgte am 18. November 2011, d.h. ca. drei Monate nach dem Multispektralflug. Die Bilder dieser Befliegung wurden photogrammetrisch orientiert und nach dem Fichten-Interpretationsschlüssel zur aerialen Waldzustandsinventur bei dreidimensionaler Betrachtung interpretiert (Abb. 2, nach HILDEBRAND, 1992). Diese Waldzustandsinventur auf Basis dreidimensionaler Strukturmerkmale ermöglichte eine manuelle Annotation (Abb. 3 links) von elf Zustandsklassen in 21 Bildern der Multispektralbefliegung, bei der diverse Bäume aus verschiedenen Blickwinkeln und unter verschiedenen Beleuchtungsverhältnissen erfasst wurden. Dabei wurden vier Klassen für Fichten definiert: gesunde (1) bzw. junge (2) Fichten, markante (3) Fichten aus der Waldzustandsansprache mit Hilfe von Strukturmerkmalen sowie Fichten im *red*- und *grey-attack*-Stadium (4). Hinzu kommen noch eine Klasse für die Buchen sowie

Klassen für Erdboden, Weg, Bodengewächse, Wasser, Holz (bearbeitet) und Sturmholz. Mit den letzten sechs genannten Klassen wird der Hintergrund modelliert.

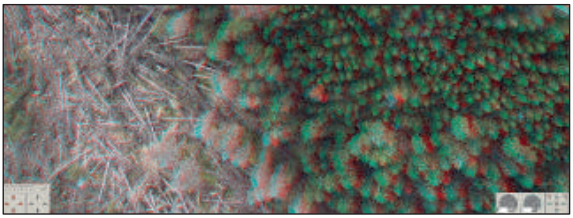


Abb. 2: Stereomodell (anaglyph) zur großmaßstäbigen Kronenzustandsinterpretation aus der UAV-Befliegung

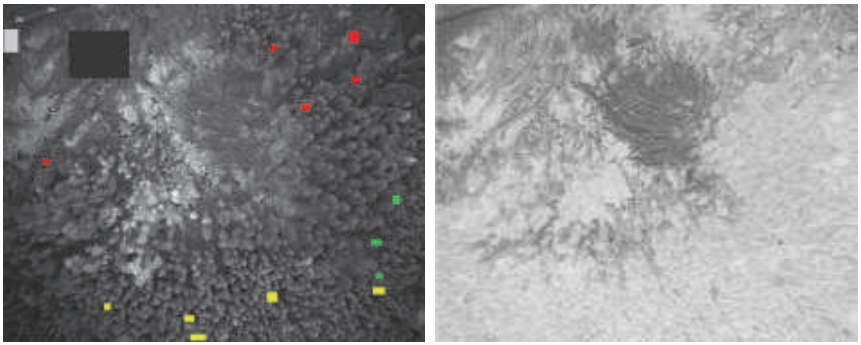


Abb. 3: Links: Manuelle Annotation einzelner Bereiche im Untersuchungsgebiet. Als Hintergrund wird das NIR-Band b_4 als Grauwertbild gezeigt. Rot markiert wurden markante Fichten, grün: gesunde Fichten, gelb: junge Fichten. Bereiche der Restklassen sind in verschiedenen Grautönen inkl. Schwarz visualisiert. Rechts: NDVI-Grauwertbild derselben Szene.

3 Methodik

Es werden zwei verschiedene Experimente durchgeführt. Zum einen werden zehn Vegetations- bzw. Chlorophyll-Indices manuell ausgewählt, die basierend auf Analysen von Multispektraldaten in der Präzisionslandwirtschaft auch für die Klassifikation von Borkenkäferschäden sinnvoll sein könnten. Zum anderen wird ein 101-dimensionaler Merkmalsraum verwendet, der die Intensitätswerte der sechs beobachteten Spektralbereiche sowie zahlreiche Kombinationen verschiedener Bänder analog zu den Vegetationsindices enthält. Die Vergrößerung des Merkmalsraums wird beim Maschinellen Lernen oft angewandt, um den Klassifikationserfolg zu verbessern (GUYON & ELISSEEFF, 2003).
MARX (2010) verwendet für seine Klassifikation von Borkenkäferschäden in Satellitenbildern einen einzelnen Entscheidungsbaum, der mit Hilfe der Software C5 der Firma RuleQuest konstruiert wird. Auf Grund der gelernten sieben Entscheidungen wird den Datensätzen eine Klasse zugeordnet. Die Klassifikationsmethode Random Forest verwendet dagegen viele solcher

Entscheidungsbäume und umgeht somit das Problem der Überangepasstheit des Klassifikators an die Trainingsdaten. Daher werden in beiden Experimenten die Merkmalsvektoren mit Random Forest nach (SHOTTON et al., 2008) klassifiziert. Um auch die Robustheit des Klassifikators beurteilen zu können, wurden aus den annotierten Bildbereichen zehn Mal voneinander verschiedene Trainings- und Testdaten ausgewählt.

3.1 Verwendete Merkmale: Vegetations- und Chlorophyll-Indices

Im ersten Versuch werden zehn Vegetations- bzw. Chlorophyll-Indices verwendet, von denen die meisten von ZARCO-TEJADA et al. (2004) einander gegenübergestellt und diskutiert werden. Alle diese Indices werden erfolgreich in der Präzisionslandwirtschaft eingesetzt, weil man mit ihnen gut vitale Vegetation von anderen Objekten unterscheiden kann.
Die ersten beiden ausgewählten Indices sind der Normalized Difference Vegetation Index NDVI (Abb. 2, rechts), d.h. die Verhältniszahl b_4 - b_2 zu b_4 + b_2, sowie der Green NDVI mit grünem anstatt roten Lichts (b_4 - b_1 zu b_4 + b_1). Es folgen die beiden Verhältniszahlen aus Red Edge zu Nahem Infrarot (b_3 zu b_4) sowie Grün zu Nahem Infrarot (b_1 zu b_4). Alle vier Merkmale werden schon seit vielen Jahren für die Erkennung von vitaler Vegetation in Multispektraldaten verwendet. Als fünftes Merkmal wird der Soil-Adjusted Vegetation Index SAVI, definiert nach (ZARCO-TEJADA et al., 2004) durch die Gleichung

$$m_5 = (1+L) * {(b_4 - b_2)}/{(b_4 + b_2 + L)} \quad \text{mit } L = 0{,}4 \tag{1}$$

verwendet, der durch den Parameter L den Bodenanteil auf der beobachteten Fläche berücksichtigt. Die nächsten drei Merkmale, die Vegetationsindices TCARI und OSAVI sind nach (ZARCO-TEJADA et al., 2004) durch die Gleichung definiert:

$$m_6 = 3 * \left(b_3 - b_2 - 0.2(b_3 - b_1)\frac{b_3}{b_2}\right) \quad \text{und} \quad m_7 = 1{,}16 * \frac{b_4 - b_2}{b_4 + b_2 + 0{,}16}. \tag{2}$$

Hinzu kommt noch ihr Verhältnis. Bei den letzten beiden Vegetationsindices handelt es sich um den Triangular Vegetation Index TVI bzw. eine seiner Modifikationen, der MTVI, definiert nach ZARCO-TEJADA et al. (2004) durch die Gleichung

$$m_9 = 60 * (b_3 - b_1) - 100 * (b_2 - b_1), \qquad m_{10} = 1{,}44 * (b_4 - b_1) - 3 * (b_2 - b_1). \tag{3}$$

So ergibt sich ein 10-dimensionaler Merkmalsvektor für die pixelweise Klassifikation.

3.2 Verwendete Merkmale: Ansatz des Maschinellen Lernens

Bei einer genauen Betrachtung der oben erwähnten Merkmale kann man gut die Ähnlichkeit von NDVI, SAVI und OSAVI erkennen. Der Unterschied liegt in der unterschiedlichen Belegung des Parameters L. Beim NDVI gilt $L=0$, beim SAVI in diesen Experimenten $L=0{,}4$, und beim OSAVI gilt $L=0{,}16$. Für eine Interpretation der Szene bzgl. der Vitalität von Vegetation ist der Parameter L sicherlich hilfreich, um Aufnahmen unterschiedlicher Zeitpunkte oder Orte miteinander vergleichen zu können. Für die Klassifikation kann die Wahl von L allerdings unerheblich sein, da es nur zu einer Stauchung bzw. Streckung des Wertebereichs des Merkmals führt. Wenn es einen Klassifikator gibt, der die zu untersuchenden Klassen mittels NDVI gut voneinander trennen kann, dann lässt sich aus diesem auch ein erfolgreicher Klassifikator ableiten, der SAVI oder OSAVI verwendet. Somit kann aus Sicht des Maschinellen Lernens hinterfragt werden, ob die Auswahl der Merkmale in vorherigen Abschnitt 3.1 sinnvoll ist. Eine ähnliche Diskussion wird auch von TOKARCZYK et al. (2012) angeregt.

Stattdessen können beliebige Kombinationen von Bändern analog zu den Berechnungen der Vegetationsindex-Ratios oder der NDVI-Formel zur Konstruktion eines hochdimensionalen Merkmalsraums verwendet werden. Eine Vergrößerung der Dimension des Merkmalsraums mittels solcher kombinierter Merkmale wird von GUYON & ELISSEEFF (2003) empfohlen, um den Klassifikationserfolg weiter zu verbessern.
In Konsequenz wird in unserem zweiten Experiment ein Merkmalsvektor mit 101 Komponenten konstruiert, dessen erste sechs Komponenten die Intensitätswerte der verschiedenen Bänder sind, d.h. $m'_k = b_k$, mit $k = 1, \ldots, 6$. Dann wurden *6* über *2*, d.h. 15 viele Kombinationen von Verhältniszahlen bestimmt:

$$m'_k = \frac{b_i}{b_j}, \text{ mit } k = 7, \ldots, 21, i = 1, \ldots, 6, j = i+1, \ldots, 6. \quad (4)$$

Andere Kombinationen sind nur Reziproke dieser 15 Merkmale und verbessern daher nicht die Klassifikation und können deshalb unberücksichtigt bleiben. Es folgen 15 Merkmale analog zum NDVI, d.h.

$$m'_k = \frac{b_i - b_j}{b_i + b_j}, \text{ mit } k = 22, \ldots, 36, i = 1, \ldots, 6, j = i+1, \ldots, 6. \quad (5)$$

Hinzu kommen noch dreimal 6 über 3, d.h. insgesamt 60 Merkmale mit einer Kombination von drei Merkmalen, wobei diese auf drei verschiedene Weisen angeordnet werden:

$$m'_k = \frac{b_i - b_j}{b_h + b_j}, \qquad m'_k = \frac{b_j - b_i}{b_h + b_i}, \qquad m'_k = \frac{b_j - b_h}{b_j + b_h}. \quad (6)$$

Die letzten fünf Merkmale werden aus dem ersten Experiment übernommen, u. z. TCARI, OSAVI, ihr Verhältnis, TVI und MTVI. Der OSAVI wurde nur deshalb berücksichtigt, weil die Verhältniszahl von TCARI und OSAVI berücksichtigt bleiben sollte.

3.3 Klassifikation mit Random Forests

Die pixelweise Klassifikation an Hand der beobachteten Merkmale wurde mit Random Forests aus *N* vielen binären Entscheidungsbäumen nach (SHOTTON et al., 2008) durchgeführt. Ein Entscheidungsbaum besteht aus zwei verschiedenen Knotenarten: den inneren Knoten und den Blattknoten. In den inneren Knoten werden die dort vorliegenden Trainingsdaten in zwei Teile geteilt und beide Teile in den jeweiligen Kinderknoten übertragen. In den Blattknoten werden die Wahrscheinlichkeiten der Klassenzugehörigkeiten aller in diesem Knoten gelandeten Trainingsdaten berechnet. Der rekursive Aufbau eines Binärbaums endet entweder, wenn alle Trainingsdaten zu einer Klasse gehören oder wenn zu wenige Trainingsdaten im Knoten vorhanden sind, um diese zu teilen.
Die *N* Entscheidungsbäume heißen Random Forest, weil in drei Schritten der Zufall ausgenutzt wird, um die Komplexität des Lernverfahrens klein zu halten. Zunächst wird jeder einzelne Entscheidungsbaum nur auf einem zufällig ausgewählten Teil der Trainingsdaten gelernt. Um keine Klasse zu bevorzugen, werden die Trainingsdaten so ausgewählt, dass eine Gleichverteilung der Klassen im Wurzelknoten eines jeden Entscheidungsbaums vorliegt. Dann ist auch die Entropie im Wurzelknoten maximal. Außerdem spielt der Zufall bei jeder Aufteilung der Trainingsdaten in den inneren Knoten eine Rolle: Mehrere Merkmale und Schwellwerte werden zufällig im entsprechenden Wertebereich ausgewählt. Nach SHOTTON et al. (2008) ist das Kriterium der maximalen Entropiereduzierung in den Kinderknoten eine sehr gute Entscheidung für die Auswahl des besten Merkmals und des besten Schwellwerts.
Wenn ein Random Forest auf neuen Daten angewandt wird, wird für jeden Datensatz in jedem der *N* Entscheidungsbäume der Blattknoten bestimmt, zudem der Datensatz auf Grundlage der hierarchisch angeordneten Entscheidungen gehört. Die Wahrscheinlichkeiten für alle

Klassenzugehörigkeiten werden über alle N Blattknoten aufsummiert, die Klasse mit der höchsten Summe wird als Klassifikationsergebnis ausgegeben.
Die Experimente dieser Arbeit werden mit jeweils N = 200 Entscheidungsbäumen in einem Random Forest durchgeführt. Wird N deutlich kleiner gewählt, sinkt der Klassifikationserfolg, bei einer größeren Anzahl von Entscheidungsbäumen, steigt nur die Laufzeit des Lernens, aber nicht mehr der Klassifikationserfolg. In allen Entscheidungsbäumen werden zehn Merkmale und zehn Schwellwerte zufällig ausgewählt, um die beste Entscheidung innerhalb eines inneren Knotens zu treffen.

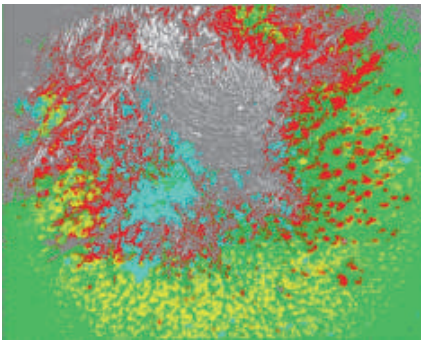


Abb. 4: Visualisierung der Klassifikation mit Random Forests mit gesunden Fichten (grün), jungen Fichten (gelb), markanten Fichten aus der aerialen Waldzustandsinterpretation nach Strukturmerkmalen (rot), Fichten im *red-* bzw. *grey-attack*-Stadium (weiß), Buchen und Unterwuchs (türkis) und alle Restklassen (grau).

4 Ergebnisse

Für beide Experimente wird eine Kreuzvalidierung mit zehn Versuchen durchgeführt. Dazu werden dieselben 5000 Pixel je Versuch und Klasse aus den manuell annotierten Daten für das Training zufällig ausgewählt. Hinzu kommen weitere 5000 von den Trainingsdaten verschiedene und ebenfalls zufällig ausgewählte Pixel je Versuch und Klasse für die quantitative Beurteilung des Lernerfolgs. In jedem Versuch setzt sich der Klassifikator Random Forest aus 200 Entscheidungsbäumen zusammen. Nach Anwenden der gelernten Entscheidungsbäume des Random Forests werden die sechs Klassen *Erdboden*, *Weg*, *Bodengewächse*, *Wasser*, *Holz (bearbeitet)* und *Sturmholz* zu einer neuen Klasse *Rest* zusammengefasst. Das vereinfacht insbesondere die Darstellung von Konfusionsmatrizen (Tab. 2 und 3).
Die quantitative Analyse der Klassifikation ergibt einen Erkennungserfolg von 68% bei den zehn manuell ausgewählten Merkmalen sowie von 75% bei den 101 definierten Merkmalen. Jede Zeile dieser beiden Tabellen gibt den Prozentsatz an, mit dem ein Datensatz der Klasse (siehe Zeile) zu einer Klasse (siehe Spalte) klassifiziert wird. Die fett gedruckten Angaben stellen dabei den Erfolg der korrekt klassifizierten Datensätze an.
Die Klassifikation von Fichten im *red-* und *grey-attack*-Stadium (F-red/grey), Buchen und dem Rest liefert in beiden Versuchen die zuverlässigsten Ergebnisse. Bei der Klassifikation von gesunden bzw. jungen Fichten sowie markanten Fichten gibt es die größten Fehler. Die

Klassifikation mit zehn Merkmalen liefert in allen Fällen schlechtere Ergebnisse, teilweise um bis zu 8%.

Tab. 2: Konfusionsmatrix mit sechs Klassen für Klassifikation mit Random Forest bzgl. 10 Merkmale.

	F-gesund	F-markant	F-red/grey	F-jung	Buche	Rest
F-gesund	**0,388**	0,197	0,000	0,305	0,101	0,008
F-markant	0,110	**0,728**	0,010	0,006	0,008	0,138
F-red/grey	0,000	0,000	**0,746**	0,000	0,000	0,254
F-jung	0,173	0,021	0,000	**0,662**	0,143	0,001
Buche	0,063	0,042	0,000	0,118	**0,752**	0,025
Rest	0,007	0,073	0,087	0,001	0,037	**0,795**

Tab. 3: Konfusionsmatrix mit sechs Klassen für Klassifikation mit Random Forest bzgl. 101 Merkmale.

	F-gesund	F-markant	F-red/grey	F-jung	Buche	Rest
F-gesund	**0,465**	0,150	0,000	0,287	0,090	0,008
F-markant	0,087	**0,755**	0,004	0,012	0,017	0,125
F-red/grey	0,000	0,004	**0,824**	0,000	0,000	0,172
F-jung	0,152	0,015	0,000	**0,742**	0,090	0,001
Buche	0,038	0,041	0,000	0,062	**0,835**	0,023
Rest	0,002	0,042	0,067	0,000	0,021	**0,868**

Zu berücksichtigen bleibt, dass die Klasse der markanten Fichten nach dem Schlüssel zur aerialen Waldzustandserfassung aus dreidimensionalen Strukturmerkmalen nach den Stufen 2 und 3 abgeleitet wurden, wonach die Baumkrone nicht mehr optimal ausgeprägt, aber auch noch nicht offensichtlich stark geschädigt oder tot ist (HILDEBRAND, 1992). Das Interessante ist dabei einerseits, dass die Bäume in der schlechter aufgelösten Multispektralbefliegung drei Monate zuvor nach dem Interpretationsschlüssel unauffällig waren und andererseits, dass die aus dem UAV Flug abgeleitete Klasse nicht über thematische Spektralinformationen definiert ist. Die Interpretation analysiert dabei nicht die Ursache einer schlecht ausgeprägten Baumkrone, sondern lediglich, dass sich der Baum nicht im optimalen äußeren Erscheinungszustand befindet. Das muss nicht zwingend durch eine *green attack stage* des Borkenkäferbefalls ausgelöst sein, sondern kann genauso beispielsweise durch einen Schaden im Randbereich des Sturmwurfes durch Anschieben und Abreißen des Feinwurzelsystems des Baumes verursacht sein.
Die durch dreidimensionale Strukturmerkmale definierte Klasse *markante Fichte* (F-markant) ist in der schlechter aufgelösten Multispektralbefliegung drei Monate zuvor nach diesen Strukturmerkmalen also unauffällig und die Klassifikation erfolgt unabhängig von diesen Strukturmerkmalen nach thematischer Spektralinformation.
Es hat daher den Anschein, dass man Auffälligkeiten einer Baumkrone bei den dreidimensionalen Strukturmerkmalen auch mit einer Analyse inhaltlicher Information (Spektren) deutlich machen kann, ohne dass die dreidimensionalen Strukturmerkmale in den Multispektralbildern deutlich sind. Dabei ist nicht geklärt, ob die Merkmale in den Multispektralbildern wegen der geringeren Auflösung oder wegen des zeitlichen Vorlaufes von drei Monaten nicht sichtbar werden.

5 Ausblick

Die in diesem Aufsatz beschriebene Studie, wurde im Jahr 2012 mit einer durch die Bayerische Landesanstalt für Wald- und Forstwirtschaft finanzierten Machbarkeitsstudie zur Früherkennung von Borkenkäferbefall in Fichten mit Hilfe von Multispektraldaten fortgesetzt. Dadurch war es möglich, im Jahr 2012 harte Referenzdaten über den Verlauf des Borkenkäferbefalls zu dokumentieren. In insgesamt sechs Multispektralbefliegen können für einzelne Bäume nicht nur tatsächlicher Borkenkäferbefall im *green attack stage*, sondern auch die aktuellen Käferentwicklungsstadien zugewiesen werden. Das war durch ein intensives Dauermonitoring auf den Flächen sowie über die Rückrechnung von angebrachten Thermologgerdaten mit Abgleich der verfügbaren Daten aus dem Pheromonfallen-Monitoring über die Korrelation zur Tagestemperatursumme möglich.
Die Auswertungen sind zum aktuellen Zeitpunkt noch nicht abgeschlossen. Es bleibt abzuwarten, ob mit der vorgeschlagenen Methodik weitere Hinweise für die Machbarkeit einer Borkenkäferfrüherkennung mit schmalbandigen Multispektralinformationen abgeleitet werden können.

Danksagung

Die Autoren danken Christian Matera für die Annotation der Daten und für seine Unterstützung bei den vorbereitenden Arbeiten. Ein weiterer Dank für die Unterstützung geht an die Nationalparkverwaltung Bayerischer Wald und das Institut für Robotik und Mechatronik des DLR Oberpfaffenhofen. Nicht unerwähnt bleiben soll die uneingeschränkte Unterstützung und Begleitung der Studie durch Dr. Heinrich Rall, ehemaliger Leiter der Abteilung Forschung und Dokumentation im Nationalpark Bayerischer Wald.

6 Literaturverzeichnis

GUYON, I. & ELISSEEFF, A., 2003: An Introduction to Variable and Feature Selection, Journal of Machine Learning Research, **3** (Mar), S. 1157-1182.

HILDEBRANDT, G. (Hrsg.), 1992: Anwendungen der Fernerkundung zur Beurteilung des Gesundheitszustandes der Wälder. Walphot, Belgien.

MARX, A., 2010: Erkennung von Borkenkäferbefall in Fichtenreinbeständen mit multi-temporalen RapidEye-Satellitenbildern und Datamining-Techniken. Photogrammetrie – Fernerkundung – Geoinformation, **2010** (4), S. 243-252.

SHOTTON, J.; JOHNSON, M. & CIPOLLA, R., 2008: Semantic Texton Forests for Image Categorization and Segmentation, IEEE Conference on Computer Vision and Pattern Recognition.

TOKARCZYK, P.; MONTOYA, J. & SCHINDLER, K., 2012: An Evaluation of Feature Learning Methods for High Resolution Image Classification, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **I-3**, S. 389-394.

ZARCO-TEJADA, P.J.; BERJÓN, A. & MILLER, J.R., 2004: Stress Detection in Crops with Hyperspectral Remote Sensing and Physical Simulation Models, Airborne Imaging Spectroscopy Workshop.

Improvement of Motion Estimation of the Taku Glacier using Spaceborne SAR Images

LI FANG [1], OLIVER MAKSYMIUK [1], MICHAEL SCHMITT [1] & UWE STILLA [1]

In dieser Arbeit wird die Oberflächengeschwindigkeit des Taku Gletschers (Juneau, Alaska) mit einem Korrelationsansatz aus hoch aufgelösten TerraSAR-X Daten mehrerer Überflüge geschätzt. Die Genauigkeit liegt dabei im Bereich von 2 cm/Tag mit einer hohen Zuverlässigkeit. Letztere wird durch einen Vergleich zwischen dem Maximum der Korrelation und seiner gemittelten Nachbarschaft bestimmt. Um die Zuverlässigkeit und die Abdeckung weiter zu erhöhen wird in dieser Arbeit eine Markov Zufallsfeld (MRF) verwendet, um die Daten eines aufsteigenden und abfallenden Orbits zusammenzuführen. Das Verfahren hilft an Stellen, die durch Schatten oder andere Artefakte keine zuverlässigen Messungen erlauben, eine Schätzung der Geschwindigkeit durchzuführen. Des Weiteren wird die Genauigkeit im Gesamten verbessert, da redundante Informationen konsequent genutzt und Widersprüche durch das Modell aufgelöst werden.

1 Introduction

The outlet glaciers of ice sheets, the main means of transporting ice from the interior to the oceans, play a very important role in the research of ice sheet variations, due to their flow velocity as one control parameter determining the mass balance. Especially, for small icefields, the mass-loss recently has become an excellent cryosphere contributor to sea level changes.
Even if in situ observations (e.g., GPS) of glacier motion can be of high accuracy level, they are costly and limited at the spatial coverage compared with remote-sensing data. The satellite remote-sensing data acquired from several sensors can help in those cases. Satellite optical imagery data have been examined to be useful sources as they provide visible features for the tracking procedure (SCAMBOS et al., 1992). However, the method relies on illumination by the sun, resulted in their dependence on weather conditions which cannot provide timely information of the target area. In contrast, SAR, an active microwave remote-sensing instrument, can obtain timely data with a spatial resolution compatible with the topographic variation by emitting its own energy, providing the ability to measure centimeter scale changes on the Earth's surface (HØGDA et al., 2010).
Much work has been successfully done to map glacier velocities by radar imagery using the feature tracking method, a cross-correlation based technique, for fast-moving ice (i.e. outlet glaciers), when interferometry and coherence tracking methods are ineffective because of the loss of phase coherence and large displacements (STROZZI et al., 2002; DE LANGE et al., 2007).

1) Photogrammetrie & Fernerkundung, Technische Universität München (TUM),
Arcisstraße 21, 80333 München, http://www.pf.bv.tum.de

Some research work has reported that the accuracy of ice velocity estimation using the intensity tracking technique can be on the order of 2cm/d when applied on high resolution repeat pass TerraSAR-X data with 11 days interval (FLORICIOIU et al., 2008; FLORICIOIU et al., 2009; EINEDER et al., 2011).
However, often, the cross-correlation technique does not find matches of features in "data holes" in the SAR scenes affected by radar shadowing and specular backscattering of smooth ice surface, and therefore fails to derive the surface velocities of glaciers in mountainous areas. For this case, this paper proposes a fusion of velocity maps at same region and acquisition time derived from the two different imaging geometries from ascending and descending satellite tracks, aiming to reduce the effects of data holes on ice motion estimation.

2 SAR Data Preprocessing

The Juneau Icefield is a low-latitude glacier system of small scale located in southeast Alaska, with the Taku Glacier as its principal outlet glacier. Current researches in Taku Glacier suggest that ice fluxes have been nearly stable for a long period (MCGEE et al., 2007; PELTO et al., 2008). In this paper, the scenes of Taku Glacier imaged by high resolution TerraSAR-X are applied for experiments. All the experimental works in this paper are based on TerraSAR-X repeat cycle data acquired in stripmap mode from two different imaging geometries (ascending and descending) between end of June and early August 2009 (see Table 1).

Table 1. TerraSAR-X data delineation used for ice motion estimation (SM-stripmap mode).

Imaging Mode	Acquisition Time	Incidence Angle Min	Incidence Angle Max	Pass Direction	Orbit
SM	2009-06-30T02:31:42	33,77	36,93	ascending	137
SM	2009-06-30T15:27:06	26,95	30,57	descending	145
SM	2009-07-11T02:31:42	35,81	38,79	ascending	137
SM	2009-07-22T02:31:43	35,80	38,79	ascending	137
SM	2009-08-02T02:31:44	33,77	36,93	ascending	137
SM	2009-08-02T15:27:08	26,94	30,57	descending	145

The TerraSAR-X Single Look Slant Range Complex (SSC) data were operationally processed into the geo-coded data using a precise available DEM (courtesy of the Chair of Geodesy, TU München). The enhanced data have a ground resolution of 3m with pixel spacing of 1.5 m in both azimuth and range direction. The geo-coded SAR scenes used in this paper are plotted on the ground truth area shown in Fig.1.

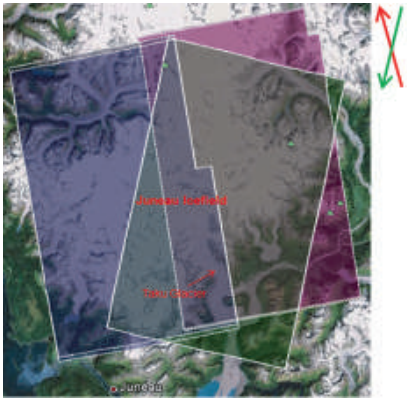


Fig. 1. Ground coverage of geo-coded SAR data plotted on an optical image of the scene. Red arrow: azimuth direction of ascending orbit; Green arrow: azimuth direction of descending orbit. In both tracks, TerraSAR-X is right looking. (Underlying optical image ©2012 Google Earth)

3 Offset Tracking

3.1 Cross-Correlation Technique

The feature tracking technique, which has been described in detail in (STROZZI et al., 2002; DE LANGE et al., 2007), was applied on the high resolution geo-coded ground range TerraSAR-X images to generate two dimensional velocity maps. The offset maps are generated by seeking the peak in the normalized cross correlation field obtained by patches from two sequential SAR intensity imageries. The successful estimation of offsets relies on the identical glacier surface features (e.g., crevasses or drainage patterns) at the scale of patches used. According to the coarse information about the velocity rate of Taku Glacier, we assumed that the maximal displacement between two successive images, with 11days cycle, is 22 meters (about 15 pixels) on the ground. Then, the sizes of sample widow and search window employed in this work are in the order of 216x216 and 296x296 pixels which are enough for the assumed displacements. In order to achieve sub-pixel accuracy, around the correlation peak, a two dimensional regression fit is employed around the correlation peak in order to obtain the newly modeled correlation peak, which is used to replace the original one, by a window (5x5 pixels). A correlation signal-to-noise ratio (SNR) expressing the comparison of the modeled correlation peak relative to the remained average value of the original correlation field is employed as confidence measurement.

3.2 Error Analysis

The absolute error of the glacier motion estimation can be calculated using the cross-correlation method for individual patches from the zero-velocity zones (i.e. rock areas) in SAR scenes. In this paper, the absolute error is computed by dividing the averaged offsets of several individual patches from stable zones located in different parts of the scene by the observation cycle.

3.3 Tracking Results

Since we perform the technique described above on the geo-coded SAR intensity images, the region of interest (ROI) – the Taku Glacier – is extracted from every scene using the geometric annotation information available in the metadata. The 2D velocity vector maps, at the same region and acquisition time, are generated using images from ascending and descending satellite tracks separately.

The calculated absolute error in this paper, using the method above, is 0.032m/d, sufficient for the ice motion estimation, which is lightly smaller than 0.036m/d in (FLORICIOIU et al., 2009). The raw result of the estimated motion map with 11.37 as the averaged SNR value, and velocities with high correlation confidence level from ascending and descending tracks are shown in Fig.2, separately.

Note that for this illustration, the outliers were culled out first by applying various rejection criteria, for instance an SNR value of 6 as lower threshold, neighborhood motion values, or flux direction.

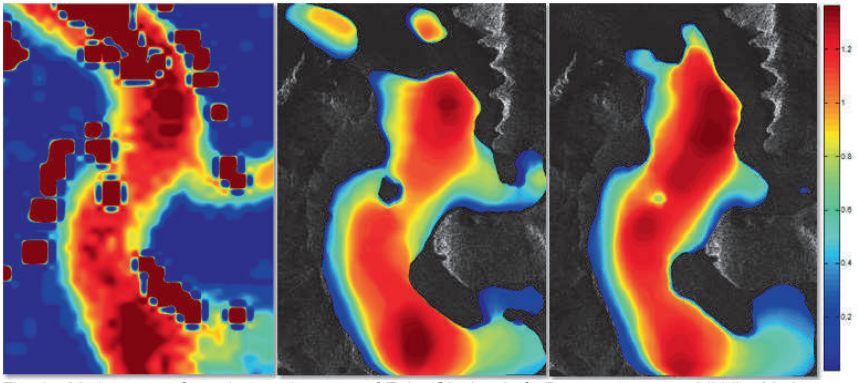


Fig. 2. Motion maps from the surging area of Taku Glacier. Left: Raw motion map. Middle: Motion map from ascending data after culling outliers. Right: Motion map from descending data after culling outliers.

It was found that some estimated data holes caused by radar shadowing that are found in the motion map derived from the ascending SAR data are successfully filled with information from the descending data with reliable SNR values. That is also true when reversely applied. Additionally, there are some holes located at the same area in motion maps derived from both satellite tracks. This should be due to the lack of features in some ice areas caused by specular backscattering due to the smooth ice surface that acts like a mirror. Thus, fusing the motion maps derived from the two different imaging geometries can help to reduce the 'data holes' and improve estimation accuracy by redundant measurements.

4 Fusion Procedures

The fusion of estimated velocities from different satellite tracks has two main concerns. The first is to enlarge the coverage area of estimation results for the scene and the second is to improve the estimation accuracy.

4.1 Weighted Averaging

The simplest method to fuse estimated motion maps from different imaging geometries is the weighted average method. In Fig.4, on the left side, the fused ground 2D velocity maps using raw motion maps from both satellite geometries by applying the weighted average method using the related SNR value as weight is shown. Obviously, this fusion result is not reliable because of the sharp edge of motion values in the glacier body. Then, we try to find another fusion method to generate more reasonable results.

4.2 Markov Random Field

The other problem concerns the estimated velocities where the measurements or estimations from a single track are contradictory to each other. If both estimations have almost the same SNR the most simple solution to fuse the tracks, namely weighted averaging, would not work. In this case it is advantageous to use local context to resolve the contradictions. We propose Probabilistic Graphical Models (PGM), more specific a Markov Random Field (MRF), as a framework for the fusion. The flexibility of PGMs to model the domain allows to benefit from the contradictions and redundant measurements. Additionally it is possible to impose some further model knowledge. One evident constraint is the smoothness of the surface velocity. It is very unlikely that a discontinuity appears in just a few measurements. The domain model used in this work is shown in Fig. 3.

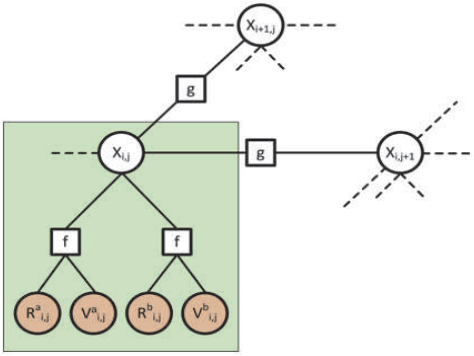


Fig. 3. Structure of the PGM used for the fusion of multiple tracks. The observations (red) are connected to the variables representing the estimated state of the surface velocity. These variables themselves are connected in a grid to impose a smoothness constraint.

The variable $X_{i,j}$ represent the discretized surface velocity of the glacier at discrete position (i, j) and are the true state which is subject to estimation. Variables $R^x_{i,j}$ and $V^x_{i,j}$ are the observations namely the SNR and the estimated velocity from track x. The model for posterior distribution, after (LI, 2009), is given as

$$P(X \mid R,V) = \frac{1}{Z}[P(R,V \mid X)P(X)] \tag{1}$$

Our state variables set X is connected to the observations by factors and represents the "data term" of the network. These factors connect three different entities with the following relationship

$$P(R,V \mid X) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp^{-\frac{(X-V)^2}{2\sigma_1^2}} \tag{2}$$

where factor σ_1 expresses the strength of σ_0 relative to the SNR.

The interaction between the variables in a PGM can be interpreted as probabilities or as energies. Despite the interpretation the optimization tries to find an extremum for the joint assignment of all variables in the network. We use a Gaussian likelihood to model the relationship between the variable and its observations. Therefore we use the causal model that the true state causes the appearance of the pixels in the different tracks. Furthermore, the model implies a conditional independence of the pixels of the tracks given the true state X. As can be seen in (2) we consider the SNR as a reliability measure for the single measurements. The higher the SNR, the more reliable the corresponding measurement and the smaller the variance will be. As a consequence the factor is much more restrictive in the deviation of the true state from the observation if the SNR is high. It should be noted that the model can be used with any number of tracks.

The other kind of factors is used to define the "smoothness term" and allows imposing some regularization to the state variables. They are particular useful in the aforementioned case of contradictory observations. In this case the regularization resolves the contradiction by using local context to determine the best state. The smoothness factors are defined by

$$P(X) = \prod_i \prod_{j \in N_i} p(X_i, X_j) \tag{3}$$

$$P(X) = \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp^{-\frac{(X_i - X_j)^2}{2\sigma_2^2}} \tag{4}$$

where X_j are the 8 neighbor pixels of X_i.

It should be noted that PGMs are not the only approach to implement this kind of model. There are even analytical methods to obtain a solution for this problem. One advantage is the flexibility of the PGMs to model almost any kind of dependency. To improve the fusion we allow discontinuities in the model. This results in a smoother velocity distribution and leads to more accurate boundaries of the glacier. To adapt the model we have to use bounded functions, deduced from equations (2) and (4), which now are defined by

$$P(R,V \mid X) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp^{-\frac{[\min(|X-V|,\beta)]^2}{2\sigma_1^2}} \quad (5)$$

$$P(X) = \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp^{-\frac{[\min(|X_i - X_j|,\gamma)]^2}{2\sigma_2^2}} \quad (6)$$

where factor β is the threshold for the maximum difference between state X and the observation value, and, also, γ is the threshold for the maximum difference between state X_i and neighbor pixels X_j. σ_0, β, σ_2 and γ are input parameters in this graphical model, and, considering the physic of glacier flow, we set these parameters as 30, 1.5, 5, 0.2 separately. The overall solution is obtained by running inference (loopy belief propagation) on the resulting graph.
The fusion result using raw motion maps from ascending and descending tracks, by applying the above graphical model, is plotted on the right side of Fig. 4, which is smoother in velocity distribution and reasonable concerning the physics of glacier flow.

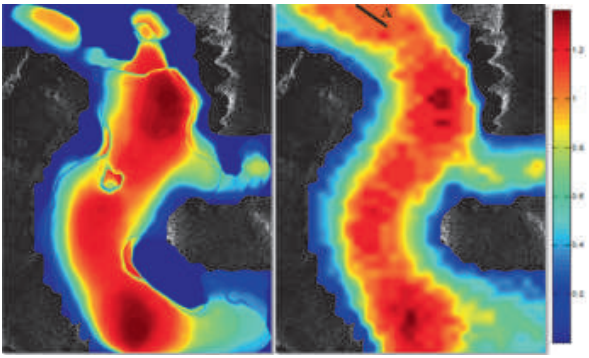


Fig. 4. Motion maps after fusion of estimation results from ascending and descending satellite tracks. Left: Fusion result using weighted average. Right: Fusion result using PGM.

4.3 Glacier Motion Result

The velocity pattern on Taku Glacier agrees well with the ground GPS observation exploited by the Juneau Icefield Research Program (JIRP). As the ice fluxes have been nearly stable for a long period, we can compare the estimated result with recent ground GPS observations. In the right graph of Fig. 4., along the center line A (black) located in upper part of the examined glacier body, the average velocity value is 1.09m/d and with the maximum value at the lowest part of line A as 1.107m/d which is slightly lower than that of 1.14m/d by ground survey in (McGee et al., 2001), possibly due to the absolute error 0.032m/d calculated in this paper.

5 Conclusion and Recommendations

The study on the outlet area of the Taku Glacier demonstrates the remarkable capabilities of TerraSAR-X stripmap mode data for deriving ice motion maps by a using cross correlation based technique. However, due to the "data holes" in the SAR scenes affected by radar shadowing and backscattering properties, especially in mountainous areas, the feature tracking method fails to derive the surface velocities of glaciers in those areas. In this case, fusion of motion maps deduced from both satellite imaging geometries can be a good solution. We have shown that the use of the weighted average method for fusion is not reliable. On the contrary, the markov random field has shown its unique capabilities on fusion of motion maps from different data sources. Additionally, fusion of motion maps can also enhance the overall accuracy by exploitation of redundant measurements. Therefore, for velocity estimation of glacier using spaceborne SAR images, the fusion of motion maps deduced from complementary orbits can greatly improve the result.

Acknowledgment

The authors would like to thank Prof. Thomas Wunderlich , Chair of Geodesy, Technische Universität München (TUM) and Dr. Oliver Lang of Infoterra GmbH / Astrium GeoInformation Services for providing the data used in this study.

6 References

DE LANGE, R.; LUCKMAN, A. & MURRAY, T., 2007: Improvement of satellite radar feature tracking for ice velocity derivation by spatial frequency filtering. IEEE Transactions on Geoscience and Remote Sensing, vol. 45, no. 7, pp. 2309-2318.

EINEDER, M.; ABDEL-JABER, W.; FLORICIOIU, D.; ROTT, H. & YAGUE-MARTINEZ, N., 2011: Glacier flow and topography measurements with TerraSAR-X and TanDEM-X. IEEE International Geoscience and Remote Sensing Symposium (IGARSS).

FLORICIOIU, D.; EINEDER, M.; ROTT, H. & NAGLER, T., 2008: Velocities of major outlet glaciers of the Patagonia Icefield observed by TerraSAR-X. IEEE International Geoscience and Remote Sensing Symposium (IGARSS).

FLORICIOIU, D.; EINEDER, M.; ROTT, H.; YAGÜE-MARTINEZ, N. & NAGLER, T., 2009: Surface velocity and variations of outlet glaciers of the Patagonia icefields by means of TerraSAR-X. IEEE International Geoscience and Remote Sensing Symposium (IGARSS).

HØGDA, K. A.; STORVOLD, R. & LAUKNES, T. R., 2010: SAR imaging of glaciers. In: P. Pellikka, and W. G. Rees (Eds). Remote Sensing of Glaciers. Taylor & Francis Group, London: 153.

Li, S. Z., 2009: Markov Random Field Modeling in Image Analysis. 3rd edn, Springer, London, pp. 15-16.

MCGEE, S. & STEMPFHUBER, W. ET AL., 2007: Surface Velocity Changes on the Taku Glacier System - 1993 to 2007. Juneau Icefield Research Program, Field Session.

MCGEE, S. ET AL., 2001: The GPS data, in GPS Science on the Juneau Icefield, Alaska. Viewed on 02 December 2012, <http://www.crevassezone.org/survey-results/surface-velocity/Longitudinal-A/Longitudinal-A.htm>.

PELTO, M. S.; MCGEE, S. R.; ADEMA, G. W.; BEEDLE, M. J.; MILLER, M. M.; SPRENKE, K. F. & LANG, M., 2008: The equilibrium flow and mass balance of the Taku Glacier, Alaska 1950-2006. The Cryosphere Discussions, vol. 2, issue. 3, pp. 275-298.

SCAMBOS, T. A.; DUTKIEWICZ, M. J.; WILSON, J. C. & BINDSCHADLER, R. A., 1992: Application of image cross-correlation to the measurement of glacier velocity using satellite image data. Remote Sens. Environ., vol. 42, no. 3, pp. 177–186.

STROZZI, T.; LUCKMAN, A.; MURRAY, T.; WEGMÜLLER, U. & WERNER, C. L., 2002: Glacier motion estimation using SAR offset-tracking procedures. IEEE Transactions on Geoscience and Remote Sensing, vol. 40, no. 11, pp. 2384-2391.

Grossflächige hochaufgelöste Schneehöhenkarten aus digitalen Stereoluftbildern

CHRISTIAN GINZLER[1], MAURO MARTY[2] & YVES BÜHLER[3]

Kenntnisse über Schneemächtigkeiten und deren räumliche Verteilung sind wichtige Informationen für eine Reihe von Anwendungen in der Schnee- und Lawinenforschung. Heute werden Schneeverteilungen üblicherweise mittels Interpolationsverfahren von punktuellen manuellen Messungen und automatischen Messstationen ermittelt. Die hohe räumliche Variabilität im alpinen Gelände ist mit dieser Methode allerdings nicht erfassbar. Terrestrisches Laserscanning (TLS) ermöglicht eine hohe räumliche Auflösung, aber die erfassbare Fläche ist klein und die Kosten dafür hoch. Mittels Airborne Laserscanning (ALS) können zwar grössere Gebiete erfasst werden, die Kosten bleiben allerdings hoch.
Aktuelle optische Sensoren, wie der ADS80, erlauben durch Stereokorrelation die Berechnung von digitalen Oberflächenmodellen (DOM). Die hohe radiometrische Auflösung von 11bit und die Kombination der spektralen Bänder von Blau bis zu nahem Infrarot bieten sich für ein Bild-Matching auch unter schwierigen Bedingungen an. Für die Korrelation werden die multispektralen , backward-looking' und ,nadir-looking' Sensordaten verwendet. Stehen digitale Geländemodelle (DTM) aus Sommerbefliegungen zur Verfügung, so lassen sich daraus Schneehöhen berechnen.
Im Untersuchungsgebiet Davos (Graubünden, Schweiz) wurden im Winter 2009/10 und 2011/12 Flugstreifen mit Bodenauflösungen von ~25 cm akquiriert. Wir vergleichen die Ergebnisse mit differentiellen GPS Messungen und Messungen von terrestrischem Laserscanning um die erreichbare Genauigkeit der Schneehöhenkarten zu quantifizieren und mögliche Einschränkungen zu identifizieren.

1 Einleitung

Schnee spielt eine wichtige Rolle im Alpinen Raum, nicht nur für den Tourismus (e.g. ELSASSER & BÜRKI 2002; RIXEN et al. 2011) sondern unter anderem auch für die Energiegewinnung und Wasserversorgung (e.g. MARTY 2008, FARINOTTI et al. 2012), die lokale Fauna und Flora (e.g. WIPF et al. 2009) und die Sicherheit im Berggebiet. Allerdings ist seit langem bekannt, dass die Verteilung des Schnees im Gebirge räumlich sehr stark variiert (e.g. ELDER et al. 1991; Schweizer et al. 2008, LEHNING et al. 2008; EGLI 2008, Grünewald et al. 2010). Allerdings kann diese räumliche Variabilität der Schneedecke bis heute noch nicht befriedigend über grosse Flächen erfasst werden.
Heute wir die Schneehöhe meist anhand von Punktmessungen an automatischen Wetterstationen oder von Beobachtern an definierten Orten gemessen (BRÜNDL et al. 2004; EGLI 2008).

1) Christian Ginzler, Eidg. Forschungsanstalt für Land, Schnee und Landschaft WSL, CH-8903 Birmensdorf; christian.ginzler@wsl.ch
2) Mauro Marty, Eidg. Forschungsanstalt für Land, Schnee und Landschaft WSL, CH-8903 Birmensdorf; mauro.marty@wsl.ch
3) Yves Bühler, WSL-Institut für Schnee- und Lawinenforschung, CH-7260 Davos Dorf; yves.buehler@slf.ch

Danach werden die Messwerte über grosse Gebiete interpoliert und mit satellitengestützten Schneebedeckungskarten kombiniert (e.g. FOPPA et al. 2007). Dabei wird aber die kleinräumige Variabilität der Schneehöhen nicht erfasst. Zudem ist kaum bekannt, inwiefern die Messungen an den Stationen für die Einzugsgebiete repräsentativ sind. Terrestrial Laserscanning (TLS) kann benutzt werden um Schneehöhen räumlich kontinuierlich zu messen (PROKOP 2008; GRUENEWALD et al. 2010). Obwohl die Messgenauigkeit sehr gut ist (meist besser als 20 cm), können mit dieser Methode grosse Einzugsgebiete wie das Dischma Tal nicht vollständig abgedeckt werden. Die Datenerfassung ist zeitaufwändig und nur bei guten Bedingungen (Wetter, Lawinensituation) und in der Nähe von leicht zugänglichen Gebieten möglich. Zudem entstehen zahlreiche Abdeckungs-Lücken, zum Beispiel hinter Hügeln, wo der Scanner keine freie Sichtverbindung hat. Laserscanning vom Flugzeug oder Helikopter aus (Airborne Laser Scanning ALS) kann grössere Gebiete auch in heiklen Situationen abdecken, allerdings sind die Kosten bereits für kleinere Flächen sehr hoch (BÜHLER et al., 2012). In der vorliegenden Untersuchung nutzen wir den opto-elektronischen Scanner ADS80 der Firma Leica Geosystems um Oberflächenmodelle für den Sommer und Winter im Raum Davos zu berechnen. Daraus leiten wir räumlich kontinuierliche Schneehöhen über weite Gebiete (ca. 120 km^2) ab. Diese Methode ist für grosse Flächen (mehrere km^2) deutlich kostengünstiger als Laserscanning. Um die Zuverlässigkeit der photogrammetrischen Oberflächenmodelle abzuschätzen vergleichen wir diese Schneehöhen mit Handmessungen, welche zeitgleich mit dem Überflug im Feld erhoben wurden.

2 Untersuchungsgebiet

Die beiden Untersuchungsgebiete befindet sich im Südosten der Schweiz, im Kanton Graubünden in der Nähe der Stadt Davos. Der Talboden befindet sich auf rund 1500 m ü. M., die höchsten Gipfel reichen bis über 3000 m ü. M.. Die Untersuchungen werden in den zwei Testgebieten Wannengrat (ca. 28 km^2) und Dischma (ca. 90 km^2) durchgeführt (Abb. 1).
Beide Untersuchungsgebiete beinhalten Siedlungsgebiete im Raum Davos und den Seitentälern Sertig, Dischma und Flüela, Nadelwald, offene Matten, Geröllhalden, kleine Gletscher und felsige Blockgrate, die zu den Gipfeln führen.
Das „Kern"-Testgebiet, wo verschiedene Untersuchungen des WSL- Institut für Wald, Schnee und Lawinenforschung SLF lokalisiert sind, liegt südwestlich des Skigebietes Parsenn im „Steintälli" und dem „Vorder Latschüel". Wir fassen dieses Gebiet nachfolgend unter dem Begriff „Wannengrat" zusammen (Abb. 1). Die Daten für das Test-Gebiet Dischma sind zurzeit noch in Bearbeitung.

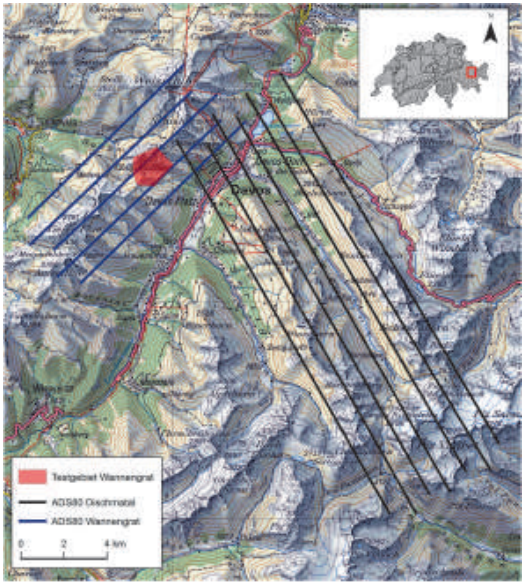


Abb. 1: Fluglinien der ADS80 Zeilensensordaten für die Untersuchungsgebiete Wannengrat und Dischma und das Testgebiet am Wannengrat (rot).

3 Daten und Methoden

3.1 Luftbilder

Die Luftbilddatensätze sind nahe am maximalen Stand der Schneehöhe im Frühjahr 2010 (16.04.2012), Frühjahr 2012 (20.03.2012) und als Referenz im Sommer 2010 (26.08.2012) für die Untersuchungsgebiete Wannengrat und Dischmatal aufgenommen worden. Die Aufnahmen erfolgten mit dem opto-elektronischen Zeilenscanner ADS80 durch die Herstellerfirma Leica Geosystems.

Das Untersuchungsgebiet Wannengrat ist mit fünf Bildstreifen abgedeckt. Es umfasst eine von Südwest nach Nordost verlaufende Gebirgskette zwischen Arosa und Davos (ca. 28 km^2). Für das Dischmatal wurden sieben Bildstreifen aufgenommen. Sie decken ein Gebiet zwischen Davos und dem Engadin ab, welches im Osten und Westen durch das Flüela- bzw. Sertigtal begrenzt wird (Abb. 1).

Die Luftbildstreifen beider Untersuchungsgebiete haben eine Querüberlappung von ca. 70 %. Die räumliche Auflösung der Bilder beträgt 25 cm (GSD). Beim Bildflug erfolgt eine simultane Aufnahme von fünf Spektralbändern (rot, grün, blau , nahes Infrarot und pan) mit einer radiometrischen Auflösung von 11 bit sowie in drei Aufnahmewinkeln (backward -16°, nadir, forward +27°). Die Orientierung mittels integriertem GNSS/IMU während der Bildaufnahme

und der anschliessenden Aerotriangulation ergab eine Lagegenauigkeit der Luftbildstreifen (x,y) von ungefähr einer GSD (25 cm).

Ein Vorteil des ADS80 Zeilensensors im Vergleich zu Matrixkameras ist die geringere Verzerrung in Flugrichtung (Linienprojektion) und die gleiche räumliche Auflösung in allen Spektralbändern, welche durch separate CCD-Sensorlinien für jedes Spektralband ermöglicht wird. Dieser Sensor wurde bereits erfolgreich für Anwendungen wie der Berechnung von Digitalen Oberflächenmodellen im Hochgebirge (BÜHLER et al., 2012), Lawinenkartierungen (BÜHLER et al., 2009) und der Klassifikation von Baumarten (WASER et al., 2011) eingesetzt. Weitere Informationen über den ADS80 Zeilensensor können SANDAU (2010) entnommen werden.

3.2 Feldmessungen

Zeitgleich mit der Befliegung wurden im Feld mit einer Lawinensonde an 15 verschiedenen Orten Plots von 5 mal 5 Messungen im Abstand von 2 Metern vorgenommen und die Eckpunkte mittels differentiellem GNSS festgehalten (Abb. 2). Somit wurden insgesamt 375 einzelne Schneehöhenmessungen von Hand erhoben. Weil die Schneehöhe aber bereits innerhalb weniger Zentimeter stark variieren kann, zum Beispiel wenn am Boden ein Felsbrocken liegt, vergleichen wir nicht die einzelnen Punkte sondern die Mittelwerte und Standardabweichungen der ganzen Plots mit den entsprechenden Gebiete aus den ADS Oberflächenmodellen. Weitere Schneehöhen-Referenzdaten, erhoben mit differentiellem GNSS, TLS und Ground Penetrating Radar GPR wurden ebenfalls zeitgleich mit dem Überflug erhoben. Diese Daten werden in den nächsten Monaten ausgewertet.

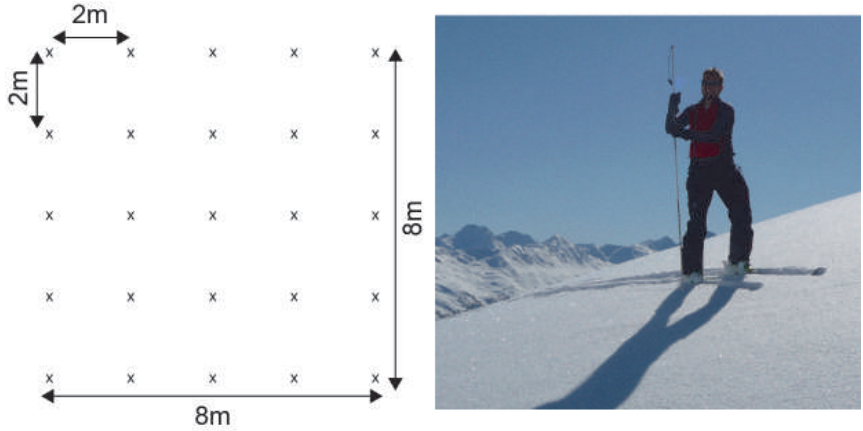


Abb. 2: Messschema für die manuellen Schneehöhenmessungen mit der Lawinensonde.

3.3 Bildkorrelation

Die Berechnung der Oberflächenmodelle erfolgt mit dem Softwarepaket SOCETSET 5.6 der Firma BAE SYSTEMS. Dabei finden die beiden Module NGATE („Next Generation Automatic Terrain Extraction“) und AATE („Adaptive Automatic Terrain Extraction“) Verwendung.

NGATE basiert auf einem hybriden Bildkorrelationsalgorithmus, welcher flächen- und kantenbasierende Methoden zur Bestimmung eines Ähnlichkeitsmasses für korrespondierende Punkte, kombiniert. Dieses wird mittels eines zweidimensionalen Kreuzkorrelationskoeffizienten berechnet. Ein „Multi-Image Matching“ Algorithmus und eine „Back-Matching“ Funktion ermöglicht eine erhöhte Redundanz von Informationen zu jedem korrelierten Punkt (DEVENECIA et al., 2007). Die Korrelation wird auf Basis der einzelnen Pixel eines Bildstreifens berechnet. Das Verhalten von NGATE kann mittels unterschiedlicher Strategien gesteuert werden.
AATE ist eine Weiterentwicklung des älteren Algorithmus ATE, welcher in (ZHANG und MILLER, 1997) beschrieben wird. Die Berechnung des Ähnlichkeitsmasses anhand der zweidimensionalen Kreuzkorrelation wird dabei durch flächenbasierte Methoden durchgeführt. Der „Multi-Image Matching“ Algorithmus und die „Back-Matching“ Funktion sind auch in dieser Softwareversion enthalten. Im Unterschied zu NGATE wird die Steuerung von AATE durch eine sogenannte „Interference Engine“ gesteuert, welche anhand von Geländeeigenschaften, Flughöhe, Signalstärke und der X- und Y- Parallaxe eine automatische Parametersteuerung vornimmt. Die Bildkorrelation wird in einem regelmässigen, Abstand durchgeführt („Post Spacing“).

In der vorliegenden Untersuchung werden die DOMs mit einer räumlichen Auflösung von 1 m (4 x GSD, XU et al., 2008) gerechnet. Dabei ist jedem Punkt ein Qualitätsmass („Figure of Merit“ FOM) zugeordnet, welcher Informationen über die Herkunft des Punktes enthält. In der vorliegende Untersuchung wird dieser Wert zur Unterscheidung von korrelierten (FOM ◘ 40) bzw. interpolierten (FOM < 40) Geländepunkten verwendet.

Die komplexe Topographie im Hochgebirge und die zum Teil sehr kontrastarmen Bildteile mit Neuschnee erschweren die Bildkorrelation. Diese Gegebenheit führt dazu, dass die Anwendung einer einzigen Strategie nie für alle unterschiedlichen Gelände- und Illuminationsbedingungen zu befriedigenden Ergebnissen führt. Daher wurden von jedem Bildstreifen drei unterschiedliche DOMs gerechnet (2 x NGATE, 1 x AATE) und miteinander kombiniert. Mit NGATE fanden die beiden Strategien „low_contrast“ und „urban“ Verwendung.

Die Kombination der verschiedenen DOMs erfolgte nach der Eliminierung der nicht korrelierten Punkte aus allen DOMs („Figure of Merit“ < 40). Die verbleibenden Punkte wurden nach der Priorität 1. „low_contrast“, 2. „urban“, 3. „AATE“ zusammengeführt, d.h. durch „low_contrast“ nicht korrelierte Geländepunkte wurden, falls vorhanden, durch „urban“ oder in einem zweiten Schritt durch „AATE“ Informationen ergänzt.

Die Schneehöhe wurde aus der Differenz des kombinierten DOM und einem DOM aus den Stereoluftstreifen der Sommerbefliegung berechnet.

4 Resultate

Zur Abschätzung der Qualität eines photogrammetrischen Oberflächenmodelles ist es von zentraler Bedeutung ob Geländepunkte aus der eigentlichen Bildkorrelation oder durch eine Interpolaton von angrenzenden Punkte gewonnen werden konnte. In Abbildung 2 sind die korrelierten bzw. interpolierten Geländepunkte aus dem kombinierten DOM für das Testgebiet Wannengrat dargestellt. Insgesamt wird dieser Ausschnitt durch 2‘377‘014 Geländepunkte Modelliert, wobei nur 0.38% (8940) der Punkte interpoliert werden mussten.

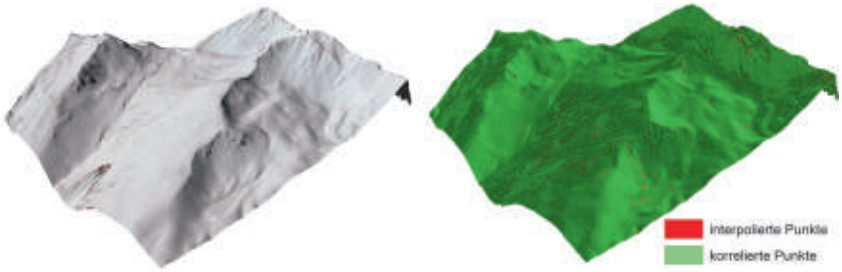


Abb. 2: Korrelierte und interpolierte Geländepunkte des fotogrammetrischen Oberflächenmodell vom Testgebiet Wannengrat, Winter 2012.

Die qualitative Begutachtung der Schneehöhenkarten (Abb. 3) belegt die Plausibilität der Resultate. Aus GRÜNEWALD et al. (2010) bekannte Schneefallen an Luv-Seiten von Geländerücken und Wächten sind sehr schön erkennbar. Wind und Schneeverfrachtungen durch Lawinen führen zu einer gut erkennbaren Verfüllung von Rinnen und insbesondere Bachtobeln.

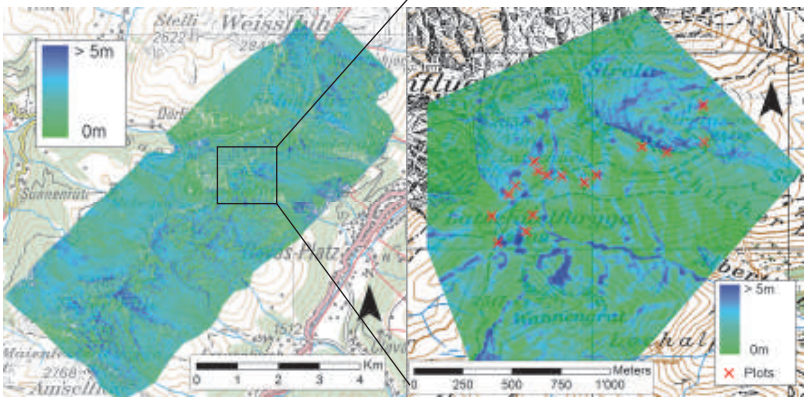


Abb. 3: Schneehöhenkarte von einem Bildstreifen und Ausschnitt im Testgebiet Wannengrat mit eingezeichneten Plots.

Um auch quantitative Aussagen über die Qualität der photogrammetrischen Schneehöhenkarten machen zu können, vergleichen wir die Resultate mit den von Hand gemessenen Plots (Tab. 1.). Allerdings konnte die Schneehöhe nicht an allen Plots manuell korrekt erhoben werden, da die Schneehöhe teilweise die Länge der Lawinensonde von 3.10 m überstiegen hat. Der Korrelationskoeffizient R zwischen den mittleren Schneehöhen der ADS- und Referenzdaten beträgt 0.95.

Tab. 1: Gemessene Schneehöhenwerte an den einzelnen Plots sowie die Differenzen des Mittelwertes und der Standardabweichung verglichen mit den ADS Schneehöhen. Plots mit Messungen über 3.10 m sind grau eingefärbt.

Name	Min	Max	Mean	STD	Min ADS	Max ADS	Mean ADS	STD ADS	Δ Mean	Δ STD
Plot1	1.80	> 3.10	2.81	0.42	1.96	4.4	3.16	0.62	-0.36	-0.19
Plot2	0.85	2.50	1.43	0.53	0.83	2.48	1.68	0.35	-0.25	0.17
Plot3	1.20	1.75	1.43	0.16	0.62	2.16	1.24	0.35	0.19	-0.19
Plot4	0.35	0.90	0.50	0.15	0.00	0.67	0.15	0.18	0.34	-0.03
Plot5	0.55	1.75	1.01	0.34	0.00	1.44	0.51	0.39	0.48	-0.05
Plot6	0.75	1.75	1.19	0.29	0.19	1.98	1.18	0.40	0.01	-0.11
Plot7	1.35	2.90	2.32	0.47	1.98	2.88	2.53	0.21	-0.21	0.26
Plot8	1.85	2.80	2.33	0.25	1.92	2.99	2.43	0.27	-0.09	-0.02
Plot9	1.40	2.20	1.71	0.23	1.35	2.25	1.72	0.23	-0.01	0.00
Plot10	0.55	2.35	1.34	0.56	0.47	2.31	1.15	0.45	0.19	0.11
Plot11	0.65	> 3.10	2.28	0.67	0.05	3.05	1.78	0.72	0.49	-0.04
Plot12	0.15	0.35	0.22	0.06	0.00	0.44	0.22	0.10	0.00	-0.04
Plot13	2.30	> 3.10	2.59	0.33	2.35	4.47	3.59	0.52	-1.00	-0.20
Plot14	0.70	2.00	1.37	0.41	0.59	1.50	1.27	0.34	0.09	0.07
Plot15	0.35	1.60	0.97	0.33	0.03	2.11	1.12	0.39	-0.15	-0.06

5 Diskussion

In dieser Untersuchung wurde erstmals erfolgreich mittels digitaler photogrammetrischer Oberflächenmodelle eine grossflächige und räumlich kontinuierliche Kartierung der Schneehöhen in einem hochalpinen Einzugsgebiet durchgeführt. Entgegen der weit verbreiteten Meinung, dass über schneebedeckten Flächen die Korrelation von Bildpunkten problematisch ist, konnten wir sehr gute Resultate erzielen. Dies obwohl die Befliegung unmittelbar nach dem letzten Schneefall stattfand und die Oberfläche sehr homogen war. Der Vergleich mit Handmessungen, welche zeitgleich mit dem Überflug im Feld erhoben worden sind, zeigt eine hohe Übereinstimmung ($R = 0.95$). Daraus leiten wir ab, dass die vorgeschlagene Technologie für die räumlich hochaufgelöste und kontinuierliche Kartierung der Schneehöhe in Alpinen Einzugsgebieten erfolgreich eingesetzt werden kann und dies deutlich kostengünstiger als mit ALS. Als Einschränkung ist zu nennen, dass bestockte Flächen nicht ausgewertet werden können. Allerdings können aus den ADS Luftbildern Bäume und apere Flächen klassiert, und vor der Schneehöhenberechnung ausgeschieden werden. Wir werden weitere Referenzdatensätze (dGNSS, TLS & GPR) verwenden um die Qualität der ADS-Schneehöhenmessungen in weiteren Gebieten (Dischma, Steintälli) zu analysieren. Zudem werden wir Befliegungen verschiedener Winter, 2010/11 (unterdurchschnittliche Schneemengen) und 2011/12 (überdurchschnittliche Schneemengen), vergleichen können. Die generierten Schneehöhenkarten sollen in Zukunft weitere wichtige Untersuchungen in der Schneehydrologie, Schneeklimatologie, Ökologie, Lawinenforschung und weiteren Alpinen Forschungsgebieten ermöglichen.

6 Literaturverzeichnis

BRÜNDL, M.; ETTER, H.-J.; STEINIGER, M.; KLINGLER, C.; RHYNER, J. & AMMANN, W. 2004: IFKIS - a basis for managing avalanche risk in settlements and on roads in Switzerland Natural Hazards and Earth System Sciences, 4, 257 – 262.

BÜHLER, Y.; MARTY, M. & GINZLER, C. 2012: High Resolution DEM Generation in High-Alpine Terrain Using Airborne Remote Sensing Techniques. Transactions in GIS, 2012, 16 (5), 635 – 647

DEVENECIA, K.; WALKER, S.; ZHANG, B. 2007: New approaches to generating and processing high resolution elevation data with imagery. Proceedings of Photogrammetric Week, 2007, Heidelberg, Germany: 297-308.

EGLI, L. 2008: Spatial variability of new snow amounts derived from a dense network of Alpine automatic stations. Annals of Glaciology, 49, 51-55.

ELDER, K.; DOZIER, J. & MICHAELSEN, J. 1991: Snow accumulation and distribution in an alpine watershed. Water Resources Research, 27, 1541-1552.

ELSASSER, H. & BÜRKI, R., 2002: Climate change as a threat to tourism in the Alps. Climate Research. 20, 253-257.

FARINOTTI, D.; USSELMANN, S.; HUSS, M.; BAUDER, A. & FUNK, M. 2012: Runoff evolution in the Swiss Alps: Projections for selected high-alpine catchments based on ENSEMBLES scenarios. Hydrological Processes, 26, 1909-1924.

FOPPA, N.; STOFFEL, A. & MEISTER, R. 2007: Synergy of in situ and space borne observation for snow depth mapping in the Swiss Alps. International Journal of Applied Earth Observation and Geoinformation, 9, 294-310.

GRUENEWALD, T.; SCHIRMER, M.; MOTT, R. & LEHNING, M. 2010: Spatial and temporal variability of snow depth and ablation rates in a small mountain catchment. Cryosphere, 4, 215-225.

LEHNING, M.; LÖWE, H.; RYSER, M. & RADERSCHALL, N. 2008: Inhomogeneous precipitation distribution and snow transport in steep terrain. Water Resources Research, 44.

MARTY, C. 2008: Regime shift of snow days in Switzerland. Geophysical Research Letters, 35.

PROKOP, A. 2008: Assessing the applicability of terrestrial laser scanning for spatial snow depth measurements. Cold Regions Science and Technology, 54, 155-163.

RIXEN, C.; TEICH, M.; LARDELLI, C.; GALLATI, D.; POHL, M.; PÜTZ, M. & BEBI, P., 2011: Winter tourism and climate change in the Alps: An assessment of resource consumption, snow reliability, and future snowmaking potential. Mountain Research and Development, 31, 229-236.

WIPF, S.; STOECKLI, V. & BEBI, P. 2009: Winter climate change in alpine tundra: Plant responses to changes in snow depth and snowmelt timing. Climatic Change, 94, 105-12.

SCHWEIZER, J.; KRONHOLM, K.; JAMIESON, J. B. & BIRKELAND, K. W. 2008: Review of spatial variability of snowpack properties and its importance for avalanche formation. Cold Regions Science and Technology, 51, 253-272.

XU, F.; WOODHOUSE, N.; XU, Z.; MARR, D.; YANG, X. & WANG, Y. 2008: Blunder elimination techniques in adaptive automatic terrain extraction. International Archives of the Photogrammery, Remote Sensing and Spatial Information Sciences, Bejing, China, Vol. XXXVII, Commission III: 1139-1144.

ZHANH, B. & MILLER, S. 1997: Adaptive Automatic Terrain Extraction. SPIE Proceedings, Volume 3072

Comparison of Image Quality in Optical Radar Fused Image Sets

MARCELLO MARIA GIOVENCO[1] & GIANLUCA GIAQUINTO[2]

Summary: The information of an image can be improved by fusing images coming from the same or from different satellites. Within the satellite image analysis optical and radar images have distinct characteristics: optical images give more complete semantic information whereas radar ones, being independent from the illumination conditions are more suitable in case of clouds or night vision. This is the reason why optical and radar images are widely used in multi-sensor image fusion. In this paper we will focus on prerequisites for fusion. The improvement shall be measured in terms of information content. The information content (IC) can be derived by the data providing that the sensors noise, the spatial and spectral resolutions be adequate to the task under consideration. In other words the information in images is strongly dependent on the system performance or the image quality (IQ). Since the additional use of radar data often gives an insignificant improvement in the final result, we expect an appreciable enhancement of IC in the fused image using IQ comparable sets of optical and radar images. In this contribution the IQ is evaluated through National Image Interpretability Rating Scale (NIIRS), a human-eye-based imagery rating scale, from 0 to 9 in order to assess the quality of pristine and fused images in terms of interpretability and possibly make some comparisons. The General Image Quality Equation (GIQE) is the worldwide recognized standard equation for computing NIIRS and for this goal the calculation of its parameters as Ground Sample Distance (GSD), Relative Edge Response (RER), Signal-to-Noise Ratio (SNR), etc. has been performed. The chosen fusion algorithm was the Principal Component Analysis (PCA) applied on different sets of optical and radar images.

Zusammenfassung: Bildinformation kann durch die Fusion von Daten mehrere Satellitensensoren verbessert werden. In der Bildanalyse haben optische und Radardaten unterschiedliche Anwendungen, weshalb man sie häufig für multisensorale Bildfusion nutzt. In dem Beitrag sollen die Voraussetzungen für eine sinnvolle Fusion von Daten unterschiedlicher Sensoren untersucht werden. Dazu wird der Begriff Informationsinhalt verwendet. Der Informationsinhalt (IC) der Daten für spezielle Aufgabenstellungen kann aus den Bilddaten durch Analyse des Rauschens sowie der räumlichen und spektralen Auflösung abgeleitet werden, wobei festgehalten werden muss, dass Bildqualität stark von der Aufgabenstellung abhängt. Da die Verwendung von Radar Daten oft keine bedeutende Verbesserung des Ergebnises gibt, erwarten wir auch keine beträchtliche Erhörung des IC, wenn die IQ von Radar und optischen Bildern nicht vergleichbar ist. IQ lässt sich mittels NIIRS auswerten. Es entspricht einer Ratingskala von 0 bis 9 und basiert auf der Genauigkeit der Objektdetektion von geschulten Bildauswertern. Darauf basierend lässt sich die Qualität von originalen und fusionierten Bildern schätzen und vergleichen. GIQE ist eine heuristische Standardgleichung zur Berechnung des NIIRS. Dafür wurde auch die Berechnung seiner Parameter (GSD, RER, SNR) durchgeführt. Die gewählte Fusionsalgorithmus war der PCA, auf verschiedenen Sets von optischen und Radar-Bilder angewendet.

[1] Marcello Maria Giovenco, Optische Informationssysteme, Rutherfordstraße 2, 12489 Berlin; E-Mail: marcello.giovenco@dlr.de

[2] Gianluca Giaquinto, Optische Informationssysteme, Rutherfordstraße 2, 12489 Berlin; E-Mail: gianluca.giaquinto@dlr.de

1 Introduction and motivations

Sensors are widely used for Earth observation and provide an efficient mean for detecting changes and analysing its surface. Two types of sensors are commonly used for this purpose: optical and radar sensors whose different characteristics alone are not always sufficient to appropriately provide a perception of the real world. Fusion instead can enhance the amount of information giving a better understanding not only of the scene observed but also of the overall image quality as we will show in this paper.

Image quality is strictly related to specific concepts like spatial resolution , Signal to Noise Ratio (SNR) or Point Spread Function (PSF) (LEACHTENAUER et al. 1997). Spatial resolution is determined by several factors like Ground Sample Distance (GSD), with which is often confused, pixel size and sensor noise and it's generally defined as the ability of the sensor to separate adjacent objects in the observed scene. The image parameters as PSF and SNR are not always easy to compute in real imagery; hence artificial targets are sometimes used instead to evaluate the Edge Response as shown in (REULKE et al. 2011).

The quality of images is usually assessed by visual analysis subjective to the interpreter to evaluate Image quality the *National Imagery Interpretability Rating Scale* (NIIRS) parameter was chosen. This is a subjective scale used for rating the quality of imagery acquired from various types of imaging systems. In particular the scale is a 10-level scale, 0-to-9, defined by interpretation tasks or criteria. The NIIRS is evaluated through the General Imagery Quality Equation (GIQE).

2 Background

2.1 Interpretability Parameters

Many parameters are suitable to describe the quality of an image. The General Image Quality Equation (GIQE) is a heuristic approach to determine the quality of an image in terms of interpretability by returning the NIIRS level. This equation includes the most representative parameters of the image quality, such as ground sampling distance (GSD), modulation transfer function (MTF) or Relative Edge Response (RER), signal to noise ratio (SNR). The single parameters themselves indeed can only indicate interpretability partially. GSD may be considered as the most important but it does not give an overall idea of the quality of images. On the other hand MTF and SNR can specify some other aspects of image quality. The NIIRS involves all these measures and moreover, also when a study of a satellite is proposed NIIRS is one of the required parameters (KIM et al. 2008). For this reason, NIIRS has been proposed as a measure of image quality in terms of interpretability.
Certain sensor parameters, e.g. spectral resolution and number of spectral channels, are not involved in this calculation.
The Edge Spread Function (ESF) or its normalized version – the Relative Edge Response (RER) can be evaluated from real images instead of the Point Spread Function (PSF) or the Modulation Transfer Function (MTF). Indeed, ESF, PSF and MTF are related, the latter being the module of the Optical Transfer Function (OTF) which is in turn the 2D-Fourier Transform of the Point Spread Function. The procedure can be found in (LEACHTENAUER et al. 1997).

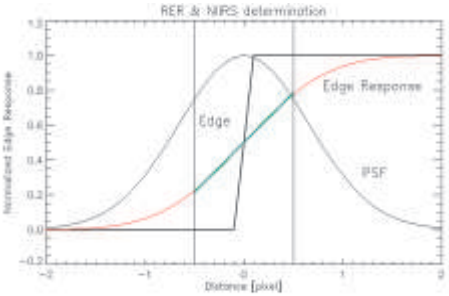


Figure 1 - Normalized Relative Edge Response

2.2 Fusion

There is no particular application underlying this work, which started as a mere scientific research as a first step to find a link between information content and image quality. Usually the choice of the fusion algorithm is limited by the specific application (POHL, VAN GENDEREN, 1998): according to the customer's needs the most suitable algorithm is chosen. Not be limited by this restriction gave the opportunity to choose among several fusion algorithms. We chose for this study the Principal Component Analysis (PCA).

Various fusion techniques have been proposed in the recent and not recent years. As you can find in (CHAVEZ, 1987), there are two main classes of fusion techniques:

- Colour related techniques
- Statistical/numerical methods

The first comprises the colour composition of three image channels in the RGB colour space as well as more sophisticated colour transformations, e.g., IHS and HSV. Statistical approaches are developed on the basis of channel statistics including correlation and filters. Techniques like PCA and regression belong to this group.

The numerical methods follow arithmetic operations such as image differencing and ratios but also adding of a channel to other image bands. A sophisticated numerical approach uses wavelets in a multi-resolution environment.

As mentioned the PCA has been chosen for this study. It is one of the most common and important algorithm of data manipulation and it is largely employed in computer graphics, neuroscience and many other fields. The main reason however is that PCA can perform fusion between any number of input bands, whereas some techniques allow only to select a limited number of bands (e.g. RGB, IHS*)*.Another aspect is that is a *non-parametric* algorithm, that is, there is no parameter to set or modify according to the input. This complete independence from the input makes possible that any data set can be put in and a smaller data set will be returned. Indeed, main aspect of the PCA is his ability in data compression. It happens to manage data of big dimensionality, often redundant and it is therefore desirable to reduce the dimensionality. This is what PCA is meant for. In a very cloudy and confused environment of data of high dimension, the PCA finds the patterns in the data and, highlighting the differences and the similarities between them. The result is a data set with lower dimensionality, where the

information is highlighted ant the noise segregated. In general there is no much loss of information. The relations between data are found through the linear algebra. The mathematical details are in (SHLENS, 2009).

It is worth to mention that before the fusion basic pre-processing such as radiometric and geometric correction would be perform. Some steps involve correction of the remote sensing images from system errors, radiometric processing for instance through speckle reduction in SAR images which can be done by filtering speckle noise before geocoding which implies improved object identification for GCP measurements (DALLEMAND et al. 1992). In some applications the filtering and resampling can be performed during the geocoding process in one step to reduce the number of times the data is resampled. Optical imagery is influenced by the atmosphere during data acquisition and therefore needs correction and/or other radiometric enhancements. Following the radiometric processing the data are geometrically corrected (geocoding, co-registration); this is a major element of pixel level image fusion since these fusion techniques are very sensitive to bad registration. In some cases, especially if image data of very different spatial resolution is involved, the resampling of the low resolution data to the pixel size of the high resolution image might cause a blocky appearance of the data. Therefore a smoothing filter can be applied before actually fusing the images (CHAVEZ, 1987).



Figure 2 - Blue channel optical image from IKONOS sensor of the selected area

3 Data Set

Two sets of data have been used for this analysis. Both sense an area of the river Elbe passing through the centre of the city of Dresden in Germany.

The data we used are:

- one radar image acquired by the Terra SAR-X satellite in Strip-Map mode on 02-07-2008 whose resolution is 2.75 meters
- four channel (blue, green, red and NIR) optical images acquired by the IKONOS satellite on 08-07-2007 whose resolution is 1 meter

The images are quite complex and complete. Jetties, trees, water and many other classes are in the scene, as visible in the Figure 2.

4 Methodology

Images were co-registered to one another through the use of ten Ground Control Points (GCPs) selected in the Image-to-Image mode. The GCPs were chosen in both open spaced and urban areas such as fields, buildings' corners, streets and so on, using then a Nearest Neighbour Resampling method to determine the digital pixel values of the corrected output image.

The images were processed according two different ways, first working on geometric resolution only and then working on radiometric resolution only. These resolutions are present in the General Image Quality Equation. As for the first way, we processed the image in a way to work with worse spatial resolution, degrading the optical images by resampling and each time fusing them with the radar image at coarse resolution. Smoothing of the optical images has been necessary in order to remove the high-frequency components that could generate artefacts and aliased components. At this scope a low-pass Gaussian filter was used. Subsequently the images have been taken to 2, 5 and 10 meters resolution, minifying by averaging with a pixel aggregate technique and magnifying with a bilinear interpolation. In the second way we fused the high resolution optical image each time with better radiometric resolutions. A better radiometric resolution has been obtained by "multi-looking" the radar image. This two steps then followed by the fusion. For every resulting image the image quality parameters were evaluated for comparisons.

5 Results

For the two processing flows described in this paper the following graphs showing the trend of NIIRS and principal components of PCA were drawn. The first graph presented in Figure 3 is very significant in order to understand the work that has been made. It shows, in fact, the NIIRS comparison prior to fusion for the optical and radar images according to the degrading optical resolution in the four channels: to achieve the same NIIRS number (same interpretability) for the two sets of images the optical resolution needs to be degraded from 1 m until 4.5 m giving us a break-even point at a NIIRS value of 2.2. Same results are shown in the Table 1 where the NIIRS values for the four optical channels and the resampling steps are reported. Decreasing the resolution means becoming more difficult for the human eye to "interpret" the image, that would then mean a decrease of the NIIRS value.

Regarding fusion we evaluated two processing flows. In the first one the radar resolution is kept still at 2.75 m (pixel size is set to 1 m) and the results of fusion with the optical at 1 m and 10 m GSD are presented. In the second flow the optical resolution is kept still while the radar has been improved through multi look process.

Relative to the first flow **Fehler! Verweisquelle konnte nicht gefunden werden.** shows the NIIRS value for the fused images at untouched resolutions equal to 4.5 which is slighlty bigger than the average value of the optical bands prior to fusion equal to 4.4. This means that a small improvement in the fusion process has been achieved.

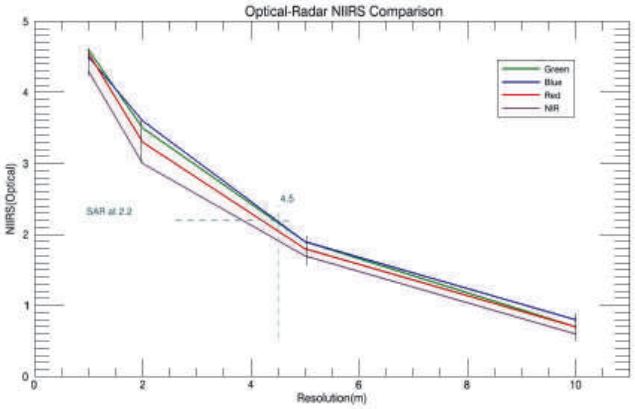


Figure 3 - The same interpretability for the two sets of images is achieved by degrading the optical resolution from 1m until 4.5 m giving us a break-even point at a NIIRS value of 2.2

GSD (m)	**1**	**2**	**5**	**10**
Green	4.6	3.5	1.9	0.7
Blue	4.5	3.6	1.9	0.8
Red	4.5	3.3	1.8	0.7
NIR	4.3	3.0	1.7	0.6

Table 1 – NIIRS values at different Ground Sample Distance for the four optical bands.

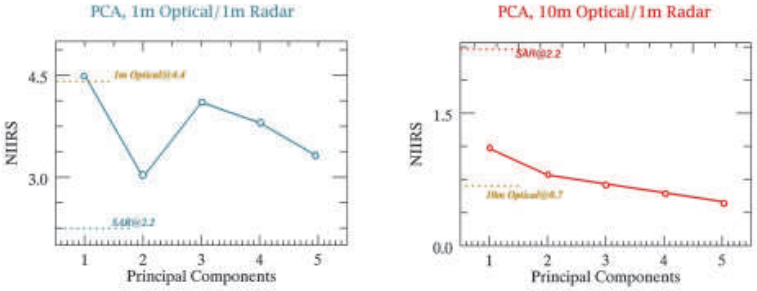


Figure 4 – NIIRS comparison: radar at 2.75m and optical at different GSDs.

Both NIIRS and RER decrease according to the resolution decreasing trend. The principal components are defined so that the first one contains the most of the information (e.g. variance) while the others contain less; this is not valid when the complexity of the scene arises greatly because of the small objects in it which have a small variance that would end up not in the first component but in the other ones which leads to irregularities in the observed trend. That is the case of the second PC in the first graph in Figure 4. It is important to notice how the same trend does not occur in the second graph because, having made the geometric resolution of optical data worse, the contribution of the small objects in the scene is much smaller than in the previous case since they have been flattened. The same analysis was made for a set of images of the new Berlin Brandenburg International Airport. The optical satellite is RapidEye, a constellation of five satellites collecting 5 meters resolution images in 5 bands, blue, red, green, red-edge and near-infrared. The images were sensed on 24th September 2010. The radar satellite is Terra SAR-X, which provides 1 meter resolution images. Date of capture was 18th September 2010. In this set of images the contribution of optical and radar images is more comparable (even if not similar: NIIRS for radar data is 1.1 while NIIRS for optical is data 2.3), hence the trend of the principal components is more linear than the previous one as shown in Figure 5.

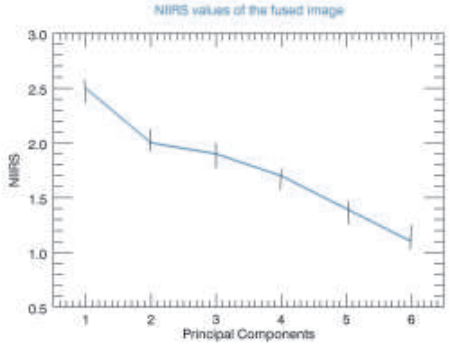


Figure 5 – NIIRS trend for PCs in BBI

In the second processing flow we increased the number of "Looks" of the radar image in order to improve the radiometric resolution: as the number of multi-looks gets higher the NIIRS value grows. This means that the influence of the radar image in the fusion is greater although still little compared to the optical one. This is visible Figure 6.

Last graph shows the NIIRS values of the first PCs (which contain most of the information) of the multi-looked radar images compared to the maximum optical NIIRS value. Optical NIIRS value is situated between the ones of the 4-looks and 16-looks radar images as shown in Figure 7. After that, image fusion takes a slight advantage in respect to the optical image itself making it worth to perform it.

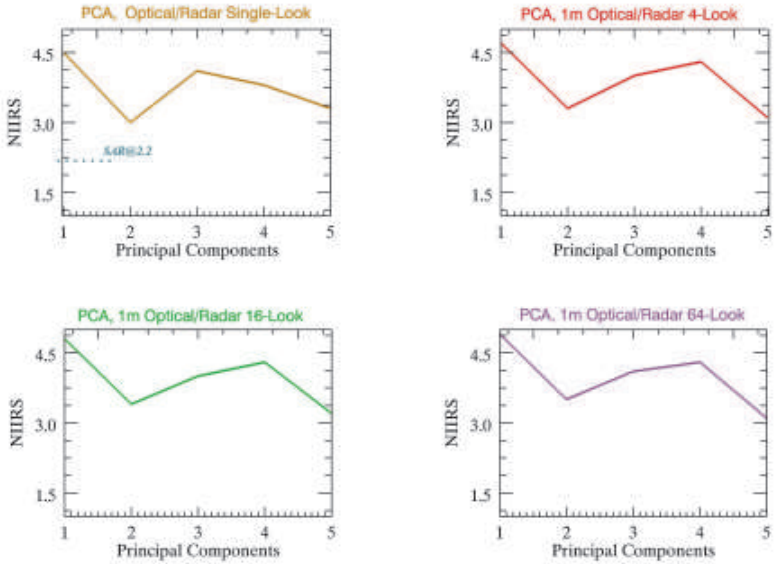


Figure 6 – NIIRS comparison for the PCA bands in the second processing flow

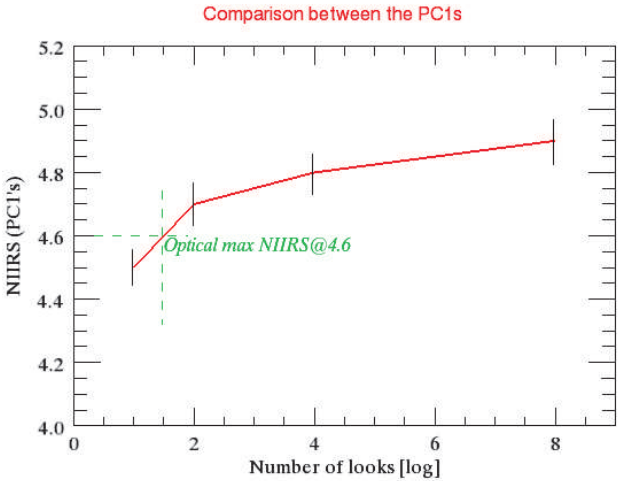


Figure 7 – NIIRS values of the first PCs compared to the maximum optical NIIRS value.

6 Conclusions

The optical and radar fusion is part of multi-sensor image fusion. The goal is to combine data from different sources in order to get as much information as possible in one single output data or integrate different information from the pristine images. In several situations, i.e. night vision, cloudy weather conditions, etc… the conveyance of different contributions wen some information might be missing can be of vital importance, especially in post-natural phenomena situations, fire detection and so on. In this study the image quality has been assumed to be related to the information content. Thus the idea was to assess the image quality of the fused images and compare it with the original images in terms of quality and interpretability. The National Image Interpretability Rating Scale (NIIRS) is a well-known metrics to determine the quality of satellite images in terms of interpretability and many important sensors parameters are adjusted according the required NIIRS level. The RER and the SNR were calculated to assess the NIIRS level. Two different paths of image processing before the fusion have been proposed: in one procedure we merged radar images with decreased optical resolutions and the results suggest that the optical images of 1 meters resolution have to be made worse to approximately 4 meters resolution to be compared with the radar RER response. By calculating the NIIRS, it is immediate to notice that the RER has a strong weight in the NIIRS equation together with the GSD. Therefore the results showed that the RER confers the trend of the NIIRS curves, whereas the GSD the range of values. The value of 4 meters was assumed to be an interpretability-threshold of fusion. The fusion was performed with the Principal Component Analysis, compact the information in the first principal components by segregating the noise in the last ones. When optical resolution in the fused image is better than the threshold resolution, the trend of NIIRS after PCA fusion is unclear and objects cannot be well compacted into major PCs. In fact the radar image is still too poor in information content in terms to be compared with the same scale of contribution of information in the fusion. It ends up in the presence of a bad segregation of the noise in the last principal components that in our case showed up in the second principal component. Conversely, when the resolution exceeds the threshold value even with a lower range determined by a worse GSD, the trend is clear and regular decreasing with lower principal components. The results were validated with another set of images capturing the Brandenburg International Airport in Berlin (Germany). The optical image at 5 meters resolution and the radar at 1 meter, gave a more linear (decreasing) trend at the NIIRS curve. In the second procedure, the radar image was improved with multi-looks, while the optical bands were kept at high resolution. With the multi-looks, which average the speckle noise – a better radiometric resolution was obtained. Results exhibit that this helps the interpretability of areas with a strong bright-dark transition. Indeed the edge response is smoothed, but more regular and with lower variance of the pixels' value in the edge profile. As a matter of fact, also the RER improves. The improvement of the radiometric resolution with the multi-look the NIIRS value had a small improvement of a 0.2-ΔNIIRS. Thanks to comparisons presented, an interpretability-threshold was found. It represents the value from which optical and radar images can be compared in terms of interpretability and in order to perform a reliable fusion. New investigation using other fusion techniques can be assessed. Noise-Adjustment Principal Component for instance, arranges the data according to SNR and not according to variance, performing better data reduction and fusion when the input images are noisy and complex. Moreover, measurements of "goodness" of fusion can be investigated. Quality metrics might be based on information theory. Some objective metrics might be very helpful when no reference image is available and for real time application, edge detection and

change detection, which can be also be thought as an alternative to NIIRS to avoid to go through performance middle-steps parameters (MTF, SNR). As for the RER from natural edges, it is feasible but more targets are required.

References

CHAVEZ, P. S., 1987: DIGITAL MERGING OF LANDSAT TM AND DIGITIZED NHAP DATA FOR 1 : 24 000 SCALE IMAGE MAPPING. PHOTOGRAMMETRIC Engineering and Remote Sensing,., **52**, S.1637-1646.

DALLEMAND, J. F., LICHTENNEGER, J., KAUFMANN, V., PAUDYAL, D. R. & REICHERT, A., 1992: COMBINED ANALYSIS OF ERS-1 SAR AND VISIBLE/INFRARED RS DATA FOR LAND COVER/LAND USE MAPPING IN TROPICAL ZONE: A CASE STUDY IN GUINEA. SPACE AT THE SERVICE OF OUR ENVIRONMENT, PROCEEDINGS FIRST ERS-1 SYMPOSIUM. PARIS: EUROPEAN SPACE AGENCY, SP-359, S.555-561

LEACHTENAUER, JON C., MALILA, JON C., IRVINE, J., COLBURN, L. & SALVAGGIO N., 1997: GENERAL IMAGE-QUALITY EQUATION: GIQE. APPLIED OPTICS. OPTICAL SOCIETY OF AMERICA, **36** (32), S. 8322-8328.

KIM, T., KIM, H. & KIM, H., 2008: IMAGE-BASED ESTIMATION AND VALIDATION OF NIIRS FOR HIGH-RESOLUTION SALTELLITE IMAGE. THE INTERNATIONAL ARCHIVES OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES, BEIJING. PART B1, XXXVII.

POHL, C. & VAN GENDEREN, J. L., 1998: MULTISENSOR IMAGE FUSION IN REMOTE SENSING: CONCEPTS, METHODS AND APPLICATIONS, INT. J. REMOTE SENSING., **19**(5), S.823-829, S.839-845.

REULKE, R., BRUNN, A. & WEICHELT, H., 2011: SPATIAL RESOULTION ASSESSMENT FROM REAL IMAGE DATA. INSTITUTE FOR ROBOTICS AND MECHATRONICS, DLR, GERMANY. HTTP://CALVAL.CR.USGS.GOV/JACIE_FILES/JACIE11/PRESENTATIONS/WEDAM/1140_REULKE_JACIE_11.110.PDF.

SHLENS, J., 2009: A TUTORIAL ON PRINCIPAL COMPONENT ANALYSIS. VERSION 3.01.

UAS-basierte automatisierte Vogelzählung am Beispiel einer Sturmmöwenkolonie der Vogelschutzinsel Langenwerder

GÖRRES GRENZDÖRFFER[1] & SVETLANA BOGDANOV[1]

Zusammenfassung: Der Beitrag stellt Erfahrungen und Ergebnisse mit UAS-Befliegungen zur automatischen Vogelzählung einer Sturmmöwenkolonie der Jahre 2011 und 2012 auf der Vogelinsel Langenwerder vor. Zu lösende Fragestellungen waren unter anderem: Üben Multicopter eine Scheuchwirkung auf die Vögel aus? (Antwort: Nein); wie hoch muss man fliegen und wie zuverlässig können Sturmmöwen erkannt und gezählt werden? (Antwort: 50 – 60 m). Lassen sich Vögel automatisch klassifizieren? (Antwort: Ja, die Erkennungsgenauigkeit der Möwen durch eine Kombination aus multispektraler Klassifikation und anschließenden GIS-Operationen liegt bei 95 - 97%. Zudem wurden stehende Möwen von brütenden Vögeln unterschieden. Wenn auch mit einer deutlich niedrigeren Zuverlässigkeit von nur 74%).

1 Einleitung

Die herkömmliche manuelle Zählung von Vögeln ist ein aufwändiger Prozess, der mehrere Personen benötigt, die von Nest zu Nest gehen und die Gelege bzw. Vögel zählen. Dabei kann nicht ausgeschlossen werden, dass Vögel mehrfach gezählt werden und auf alle Fälle stellt die Zählung eine Störung im Brutprozess dar. Manuelle Zählungen, basierend auf Schätzung und Hochrechnung unterliegen grundsätzlich größeren Ungenauigkeiten. Diese sind abhängig von der zu untersuchenden Vogelart, der Anzahl der Individuen, dem Beobachter sowie dem Beobachtungspunkt. Dabei sind bei Ansammlungen von etwa 2.000 Individuen Über- bzw. Unterschätzungen um 25% keine Seltenheit (z.B. BIBBY et al. 1995, KULEMEYER et al. 2011).
Fernerkundliche oder andere Alternativen gab es bisher nicht, da die notwendige Bodenauflösung von 2 cm oder besser von Flugzeugen kaum zu erreichen ist, bzw. für ein kleines Gebiet auch kaum wirtschaftlich ist. Seit kurzem jedoch bieten UAS (Unmanned Airborne Systems) eine überzeugende Alternative. Erste Untersuchungen, z.B. SARDÀ-PALOMERA et al. (2012), WATTS et al. 2010 haben gezeigt, dass es grundsätzlich möglich ist, mit UAS Vögel zu beobachten. Allerdings haben sich beide Studien auf die UAS-Bildaufnahme beschränkt und die Auswertung, d.h. das Zählen von Vögeln, manuell vorgenommen.

2 Untersuchungsgebiet

Das Untersuchungsgebiet ist die ca. 25 ha große Vogelschutzinsel Langenwerder. Sie liegt nördlich der Insel Poel und ist seit 1903 unter Schutz gestellt. Die Insel ist vorwiegend mit Gras und Schilf bewachsen und nahezu baumfrei sowie für die Öffentlichkeit gesperrt. Um die brütenden Vögel zu schützen wird zur Hauptbrutzeit im Süden der Insel ein Elektrozaun aufgestellt, der z.B. Füchse oder Wildschweine abhalten soll sich auf der Insel reichlich zu bedienen.

1) Görres Grenzdörffer, Svetlana Bogdanov, Universität Rostock, Professur für Geodäsie und Geoinformatik, J.-v.-Liebig Weg 6, 18059 Rostock,
email: {goerres.grenzdoerffer},{svetlana.bogdanov}@uni-rostock.de

Hauptvogelart der Insel ist die Sturmmöwe (*Larus Canus*). Die Insel beherbergt die größte Sturmmöwenkolonie der Ostseeküste, allerdings mit abnehmender Tendenz. Während es 2002 noch 3.000 Brutpaare gab, ist die Zahl in letzten Jahren auf ca. 1.500 Brutpaare zurückgegangen. Neben weiteren Brutvogelarten stellt die Insel einen bedeutenden Rastplatz für Zugvögel, wie z.B. Ringelgänse und Lemikolen dar. Eine ausgewachsene Sturmmöwe hat eine Körperlänge von 40–45 cm und eine Flügelspannweite von 100–130 cm. Von oben betrachtet hat die Sturmmöwe einen typisch weißen Kopf, ein hellgraues Federkleid und grauweiße Schwanzfedern.

2.1 UAS und Vögel?

Vor dem ersten Flug im Mai 2011 stand die große Frage, wie reagieren die Vögel auf ein UAS? Aus der bemannten Fliegerei ist bekannt, dass Flugzeuge, die z.B. zur Zählung von Seevögeln eingesetzt werden und in einer Höhe von 80 m agieren, eine große Scheuchwirkung haben, z.B. KULEMEYER et al. 2010. Andere Autoren berichten, dass eine Mindestflughöhe von 300 - 400 m notwendig sei, um störungsfreie Aufnahmen zu bekommen.
Bei den elektrisch betriebenen Multicoptern (Falcon 8, MD4-1000) zeigten die Vögel glücklicher weise überhaupt kein Fluchtverhalten gegenüber dem UAS. Ab einer Entfernung von ca. 15 m wurde das UAS schlicht und ergreifend ignoriert. Das galt nicht nur für die Sturmmöwen, sondern auch für andere auf der Insel vorkommende Vogelarten, vgl. Abb. 1.



Abb. 1: Vögel lassen sich nicht durch UAS (Falcon 8) stören

2.2 UAS-Bildflugplanung zur Vogelzählung

Nachdem sich die Vögel nicht um das UAS kümmern, stand als nächste Frage, welche Flughöhe bzw. GSD ist notwendig, um einerseits die Vögel noch zu erkennen und andererseits eine maximale Bodenabdeckung pro Bild zu gewährleisten, um die begrenzte Flugzeit von ca. 15 Minuten beim Falcon 8 auszureizen. Eine Annahme war, dass ein Vogel etwa 20 - 30 Pixel groß sein sollte, um ihn sicher identifizieren zu können. Aufgrund der Größe der Vögel von 40 - 45 cm wurde

empirisch überprüft, ob diese Annahme trägt. Die empirisch ermittelte optimale Flughöhe lag bei 50 m und einem dazugehörigen GSD von ca. 1.6 cm.
Bei der Bildflugplanung wurde eine 80 % Längsüberlappung und eine 60 % Querüberlappung gewählt. Die große Querüberlappung soll einerseits gewährleisten, dass Windböen, die unkontrollierbare Rollbewegungen verursachen können, nicht zu Lücken im Bildverband führen. Andererseits hilft die große Querüberlappung auch bei der Georeferenzierung und ermöglicht ein True Orthophoto. Üblicherweise stoppt das UAS bei einer Aufnahme mehrere Sekunden, um auszupendeln und möglichst senkrechte Bilder auszulösen. Nachdem sich herausgestellt hat, dass die so ausgelösten Bilder, insbesondere bei Wind, auch nicht so viel senkrechter sind, wurde diese Funktion ausgestellt und in einem definierten Zeitintervall - entsprechend der Flughöhe und Geschwindigkeit - kontinuierlich ausgelöst. In den beiden Jahren kamen zwei unterschiedliche UAS zum Einsatz, deren technische Parameter in der nachfolgenden Tabelle 1 zusammengefasst sind:

Tab. 1: Eingesetzte UAS für die Befliegungen der Vogelinsel Langenwerder

	Falcon 8* (2011)	**MD4-1000 (2012)**
Anzahl Rotoren	8	4
Max. Payload	400 g	1.200 g
Abfluggewicht	ca. 2. kg	ca. 5 kg
Flugdauer (mit normaler Payload)	ca. 15 min	ca. 30 min
Kamerastabilisierung	Ja (2-Achsen)	Ja (2-Achsen)
Kamera	Sony NEX5	Olympus PEN E2
Auflösung Kamera (Pixel)	4592 x 3056	3000 x 4000

* Ascending Technologies (http://www.asctec.de/uav-anwendungen/luftbild/asctec-falcon-oktokopter/drohne/)

** Microdrones GmbH (http://www.microdrones.com/products/md4-1000/md4-1000-key-information.php)

3 Bildflüge 2011

In 2011 wurden am 30.05.2011 zwei Flüge über Teilen der Insel durchgeführt. Beim ersten Flug gab es technische Probleme die dazu führten, dass die Bildflugplanung durch das UAS nur unzureichend umgesetzt wurde und unsystematisch Bilder ausgelöst wurden. Der zweite Flug verlief reibungslos und 50 Bilder standen für die Weiterverarbeitung zur Verfügung. Eine klassische Aerotriangulation war mangels Passpunkten nicht möglich, da die Insel zur Passpunktmessung leider nicht betreten werden konnte. Eine passpunktlose Aerotriangulation schied ebenfalls aus, da die GPS-Näherungswerte des UAS mit ± 2 - 3 m zu ungenau sind.
Letztendlich wurde ein anderer Ansatz gewählt, der auf einer kommerziellen Panoramasoftware aufbaut. Nach verschiedenen Experimenten mit Open Source Produkten und frei nutzbarer Software fiel die Wahl schlussendlich auf die Software Autopano Giga 2.5. Diese Software lässt mehrere Standorte zu und ermöglicht zudem einzelne Fehlzuordnungen manuell zu korrigieren. Bei der Mosaikierung hat sich herausgestellt, dass die Radialverzeichnung, die bei dem Weitwinkelobjektiv der Sony Nex5 durchaus beträchtlich ist, durch die Software nur unzureichend

modelliert werden kann. Demzufolge wurde die Radialverzeichnung mit der Software ImageIron korrigiert. Im Ergebnis hat das Mosaik 20.520 x 19.297 Pixel und deckt eine Fläche von ca. 4.4 ha ab. Abschließend wurde der mosaikierte Bildverband über 4 Punkte grob georeferenziert.

4 Automatische Vogelerkennung und -zählung 2011

Bei der Vogelerkennung, die durch P. SCHNEIDER (2011) im Rahmen einer Bachelorarbeit erstellt wurde, ist eine einfache Strategie verfolgt worden, die in erster Linie auf den spektralen Eigenschaften der Möwen basiert. Schließlich sind die Vögel mit ihren weißen Köpfen und hellgrauen Körper fast die einzigen weißen Elemente auf den Bildern. Auf einen komplexeren objektorientierten Ansatz wurde bewusst verzichtet, da die Bilddaten recht groß sind und lange Bearbeitungszeiten zur Folge haben können. Erst in einer zweiten Stufe spielen Größe und Form eine Rolle. Im Einzelnen läuft die Auswertung wie folgt ab:

1. Überwachte Multispektralklassifikation:

Da die Möwen einen weißen Kopf und ein hellgraues Federkleid aufweisen, sind sie spektral in den Farbbildern eindeutig zu erkennen. Mit dem Programm ArcGIS 10 wurde eine überwachte Multispektralklassifikation mit insgesamt 7 Klassen erstellt. Anschließend wurde das als Raster vorliegende Ergebnis der Klasse "Vögel" (52.488 Objekte) in Vektorform überführt.

2. Größenbasierte Selektion der Vögel

Im Bereich des Strands kam es zu einer spektralen Überschneidung zwischen den Vögeln und dem hellen Sand. Da auf dem Strand keine Vögel brüten, wurde der Strandbereich digitalisiert und die darin klassifizierten Objekte eliminiert (11.157 Objekte bleiben übrig). Anschließend wurde für die Polygone die Flächengröße berechnet und als neues Attribut hinzugefügt. Über dieses Attribut erfolgt die Selektion nach Größe. In diesem Schritt werden alle Objekte (z.B. kleine helle Steine etc.), deren Flächengröße kleiner als 10 cm² ist, entfernt. Zu große Objekte, z.B. Höckerschwäne wurden durch eine Größenschwelle von > 2.000 cm² ebenfalls entfernt. Aufgrund des Farbwechsels, heller Kopf, graues Gefieder und helle Schwanzfeder der Vögel, ist ein einziger Vogel oftmals als zwei eng benachbarte Vogelobjekte klassifiziert worden. Um die beiden Teile zusammenzuführen, wurden benachbarte Polygon vereint, deren Entfernung zueinander kleiner als 24 cm ist. Abschließend wird abermals nach Größe selektiert, wobei die jetzt angesetzte Schwelle nun 20 cm² beträgt.



Abb. 2a Vogelobjekt vor Aggregierung



Abb. 2b: Vogelobjekt nach Aggregierung

Als Endergebnis konnten 1.568 Vögel gezählt werden. Die Genauigkeit (Über- und Unterschätzung) wurde mit Hilfe visueller Gegenzählungen eines definierten Abschnittes mit 500 Vogelobjekten ermittelt und anschließend hochgerechnet.
Als Ergebnis der Auszählung der 500 Vogelobjekte konnten tatsächlich 477 Sturmmöwen verifiziert werden (95.4 %). Dabei wurden 65 Objekte fehlerhaft als Vogel identifiziert und gleichzeitig sind 42 Vögel nicht als solche klassifiziert worden. Beim Auszählen per Hand konnten verschiedene charakteristische Hauptfehler der Erkennung von Objekten ausfindig gemacht werden: falsch erkannte Objekte waren hauptsächlich Holzstücke, Steine und stellenweise helle Stellen in Grünflächen. Außerdem wurden diejenigen Vögel nicht erkannt, die zu „dunkel“ waren, bzw. in unscharfen Bildbereichen auftraten, im Schatten von Büschen oder zu nahe beieinander waren und daher als ein Objekt klassifiziert wurden. Ursache der Bildunschärfen war einerseits das Weitwinkelobjektiv der Sony Nex5 und anderseits die verwendete Stitchingsoftware, die an den Bildübergängen einen unscharfen Saum erzeugt hat.

5 Automatische Vogelzählung 2012

Auf den Erfahrungen von 2011 aufbauend wurde in 2012 das Untersuchungsgebiet durch die Verwendung der MD4-1000, die über eine Flugzeit von ca. 25 - 30 Minuten pro Batterie verfügt, erweitert und die Flughöhe auf 55 m geringfügig erhöht. Eine Befliegung der relevanten Inselteile konnte am 25.5.2012 mit insgesamt 629 Bildern erfolgreich durchgeführt werden. Die äußeren Bedingungen waren perfekt, nahezu windstill und wolkenloser Himmel. Die Befliegung der Insel erfolgte durch vier Teilbefliegungen, um zwischenzeitlich den Akku zu wechseln und durch eine geschickte Wahl der Startpunkte, die notwendige Flugzeit zu minimieren. Um sicherzustellen, dass die Kamera während des Fluges nicht wegen Strommangels die Bilder aufzunehmen "vergisst", wurde zwischen den Flügen der Akku gewechselt.
Begleitend zur Befliegung wurden an einem Rand der Insel einige signalisierte Passpunkte mit einem RTK-GPS gemessen. Da die Vögel im Brutgeschäft waren, konnte die Insel zur Passpunktmessung wiederum nicht betreten werden. Somit dienten die randlichen Passpunkte nur dazu, um den Block lagemäßig einzupassen.
Die photogrammetrische Datenauswertung erfolgte mit der Software Pix4D, die als Webservice genutzt wurde. Im Ergebnis wurden ein digitales Orthophoto, ein DGM und eine Punktwolke geliefert. Entsprechend dem Report lag die Lage- und Höhenabweichung (1 σ) an den fünf Passpunkten bei ca. 5 mm.

5.1 Automatische Vogelerkennung und -zählung 2012

Das Verfahren der automatischen Vogelzählung 2012 ist mit dem Ansatz von 2011 grundsätzlich vergleichbar. Nur die Größenfilter zur Eliminierung von kleinen Fehlzuordnungen und großen Objekten sind geringfügig geändert worden. Die Zählung ergab für das Jahr 2012 insgesamt 1.945 Vögel, vgl. Abb. 3. Die Klassifikationsgüte ist anhand von 605 Objekten validiert worden. Es wurden 560 Objekte als Vögel verifiziert. 45 Objekte entpuppten sich bei näherem Hinsehen nicht als Vogel, sondern als etwas anderes. Gleichzeitig sind 31 Vögel nicht erkannt worden Das bedeutet im Saldo wurden 97.6 % der Vögel erkannt. Die Steigerung der Klassifikationsgenauigkeit in 2012 lässt sich eindeutig auf die bessere Bildqualität zurückzuführen.

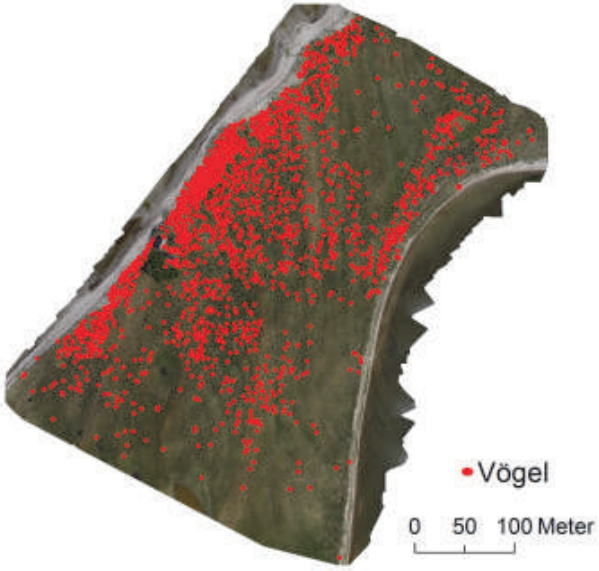


Abb. 3: Identifizierte Vogelobjekte (rote Punkte) aus UAS-Befliegung vom 25.5.2012 auf der Vogelschutzinsel Langenwerder

5.2 Identifikation von Brutvögeln

Bei der Vogelzählung ist die Zählung aller Individuen eine wichtige Größe. Weit wichtiger ist jedoch die Zählung der Brutpaare bzw. Nester oder auch der Anzahl der gelegten Eier. Da die Größe der Gelege aus der Luft nicht erkennbar ist - schließlich sitzen die Vögel drauf und brüten, ist es von großem Interesse zumindest die Anzahl der Brutpaare zu ermitteln. Der vormittägliche Befliegungszeitpunkt ist grundsätzlich ein optimaler Zeitpunkt, da bei den meisten Brutpaaren ein Partner zu dieser Tageszeit auf Futtersuche ist und die Anzahl der gezählten Vögel näherungsweise mit den Brutpaaren übereinstimmt. Dennoch sind auf den Bildern stehende von brütenden Vögeln visuell zu unterscheiden. Bei näherem Betrachten fällt auf, dass diese einen kleinen Schatten werfen, wie die Abb. 4a und 4c schön verdeutlichen.

Die Richtung und Länge des Schattens ergibt sich aus dem Sonnenstand, der sich aus der geographischen Position (54° 01' N; 11° 29' E) und der Uhrzeit (ca. 10:00 Uhr MESZ) berechnen lässt sowie der Größe der aufrecht stehenden Vögel. Somit fällt der Schatten aus einer Elevation von etwa 40° in Richtung 290°. Bei einer Vogelgröße von etwa 30 bis max. 40 cm beträgt die Länge des Schattens 34 - 46 cm.

Der dunkle Schattenbereich lässt sich einfach klassifizieren bzw. spektral vom Rest des Bildes trennen. Um die Vögel wird ein sektoraler Buffer mit einem Durchmesser von 40 cm berechnet. Die Schattenmaske wurde mit den Vogelpufferzonen verschnitten (4.137 Objekte). Dabei zeigt

sich, dass viele Fehlzuordnungen zu eliminieren sind. Ursachen sind zum eine viele kleine Schattenelemente und Vögel brüten gerne in höherem Gras, welches Schatten erzeugt. Ein Größenfilter entfernt viele kleine Schattenelemente, die auch direkt neben sitzenden Vögeln auftreten können. Ein weiterer Filter grenzt diejenigen Schattenobjekte ab, die nicht direkt, d.h. weniger als 15 cm vom Vogelzentroid entfernt, neben den Vögeln anfangen. Ein letzter Filterschritt selektiert diejenigen Vögel, die nicht in höherem Gras (> 15 cm) stehen. Die Grashöhe wird aus der Punktwolke ermittelt (vgl. Abschnitt 6.1).
Schlussendlich konnten 54 Vögel identifiziert werden, die nicht brüten, sondern anderen Aktivitäten nachgehen. Eine visuelle Kontrolle ergab, dass die Strategie leider nicht wirklich erfolgreich war. Vielmehr lag die Trefferquote bei nur etwa 74%. Ursache ist unter anderem, dass in der direkten Umgebung eines sitzenden Vogels der bultenförmige Bewuchs einen Schatten erzeugt. Gleichzeitig wurden mehrere stehende Vögel nicht richtig erfasst, da sie z.B. mit ihrer Längsachse in Ost-West Richtung standen, und so nur einen sehr kleinen Schatten geworfen haben (Abb. 4a-c). Als Fazit lässt sich ziehen, dass vielleicht 50 - 100 Vögel nicht im Brutgeschäft sind, so dass sich die Anzahl der Brutpaare auf der Vogelinsel 2012 demzufolge auf etwa 1.850 - 1.900 beziffern lässt.



Abb. 4a: korrekt klassifizierter stehender Vogel mit Schattenwurf



Abb. 4b: fehlerhaft klassifizierter stehender Vogel (Bewuchs in der Umgebung



Abb. 4c: Nicht klassifizierte stehende Vögel (Ungünstige Ausrichtung, d.h. zu kleiner Schatten)

6 Räumliche Verteilung von Vögeln

Die Brutplätze der bodenbrütenden Vögel sind nicht gleichmäßig über die Insel verteilt - was auch nicht wirklich zu erwarten war. Vielmehr zeigt sich, dass die Masse der Vögel auf der westlichen Seite, bzw. auf dem Uferwall brüten (vgl. Abb. 3). Ursache ist wahrscheinlich, dass dort der vorherrschende Westwind optimal unter die Flügel greifen kann und somit der Energieaufwand beim Starten minimal ist. Im Südosten der Insel sind wenige Vögel zu beobachten, was laut Aussagen der Ornithologen vor Ort auf den höheren Bewuchs zurückzuführen ist. In der Inselmitte und im Inselnorden sind ebenfalls relativ wenige Vögel zu beobachten. Mögliche Ursachen können z.B. sein, dass dort die Brutbedingungen nicht optimal sind.
Gemäß der Literatur, z.B. von BLOTZHEIM & BAUER (2000) sucht das Männchen den Nistplatz aus, wobei oftmals der gleiche Platz wie in den Vorjahren gewählt wird. Der Nistplatz sollte möglichst sicher vor Bodenfeinden, Witterung und Überflutung sein. Meist steht das Nest an

wenig bewachsenen Standorten am Boden. Zudem sind Bodennester gerne auch an unzugänglichen Stellen wie Inselchen, Bülten, Landzungen, Dämmen oder im Schwingrasen.
Anhand der aus dem UAS gewonnenen fernerkundlichen Daten (Orthophoto, DGM und der 3D-Punktwolke) lassen sich die wichtigsten Standorteigenschaften ermitteln.

6.1 Bestimmung der Vegetationshöhe aus der 3D-Punktwolke

Die Höhe der Vegetation lässt sich nicht direkt aus dem digitalen Oberflächenmodell bestimmen. Viel interessanter ist die 3D-Punktwolke. Im Profil dargestellt (Abb. 5) zeigt sich, dass die Bestandesoberfläche nicht durch einen Punkt abgebildet wird, sondern einen Saum darstellt. Die Höhe des Saums variiert, je nach Höhe des Aufwuchses aus Gräsern, Seggen oder kleinen Büschen.

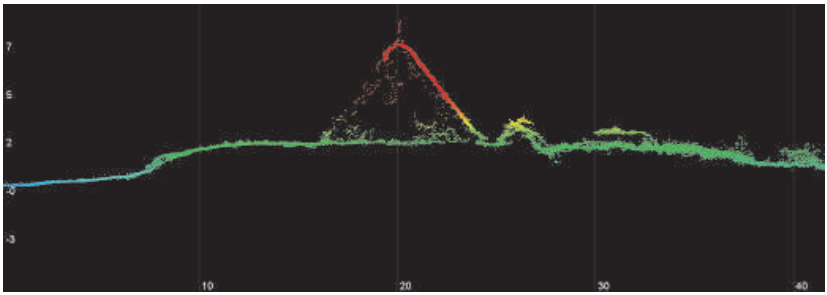


Abb. 5: 2 m breiter und 42 m langer Querschnitt durch 3D-Punktwolke (v.li.n.re.) Strand - Vogelwärterhaus - Nebengelass - Buschwerk)

Um die Höhe des Saumes bzw. der Vegetation zu bestimmen wurde die 3D-Punktwolke mit der Software LAStools / lascanopy analysiert. Die durchschnittliche Punktdichte der Datei beträgt 215 Punkte / m² bzw. einen durchschnittlichen Punktabstand von 6.8 cm. Damit ist die Punktdichte, verglichen mit klassischen Laserscannern, natürlich sehr hoch und trotz der grundsätzlichen Erfassung der Oberfläche durch die Photogrammetrie auch in der Lage innerhalb des unregelmäßigen Aufwuchses die Geländeoberfläche zu erfassen. Aufgrund dessen wurden in einer Rasterweite von 0,5 m das Minimum, Mittelwert, Maximum sowie die Standardabweichung der Punkte pro Rasterzelle berechnet. Da die Extremwerte (Minimum und Maximum) fehlerbehaftet sein können, entspricht die zweifache Standardabweichung am ehesten der tatsächlichen Wuchshöhe (Abb. 6). Wie an obiger Stelle erwähnt ziehen die Vögel Standorte mit geringem Bewuchs vor. Abb. 7 zeigt die Verteilung der Wuchshöhe an den jeweiligen Vogelstandorten.

6.2 Geomorphologie und Vogelverteilung

Die insgesamt sehr flache Insel Langenwerder mit einer durchschnittlichen Höhe von nur 1,09 m über dem Meeresspiegel ist ständig "in Bewegung", da laufend durch Meeresströmungen Sand an unterschiedlichen Stellen an- und abgetragen wird. Die während der Inselbildung vorherrschende südliche Strömungsrichtung lässt sich aus der Richtung der Strandwälle, die von NNE

nach SSW laufen, bestimmen, LEMKE, 2004. Diese sind auf dem Digitalen Geländemodell sehr gut zu beobachten, Abb. 8.

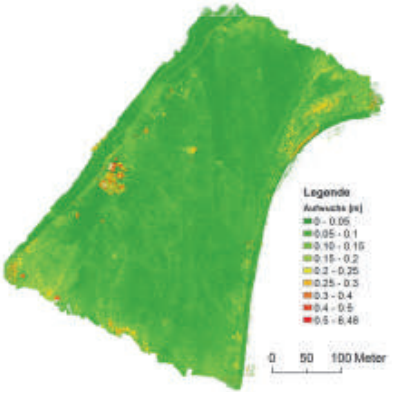


Abb. 6: Berechnete Aufwuchshöhen der Vegetation aus der 3D-Punktwolke

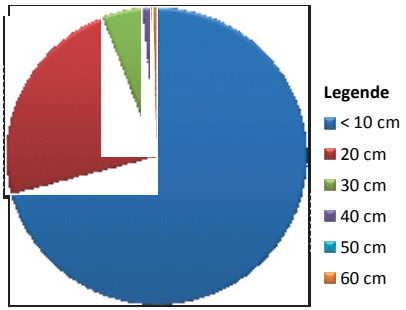


Abb. 7: Wuchshöhe der Vegetation am Standort der Vögel

Die aktuellen Küstenformungsprozesse sind etwas anders gelagert. Zum einen hat sich die Hauptwind- bzw. Hauptströmungsrichtung auf West geändert, so dass die Insel auf der südwestlichen Seite wächst.

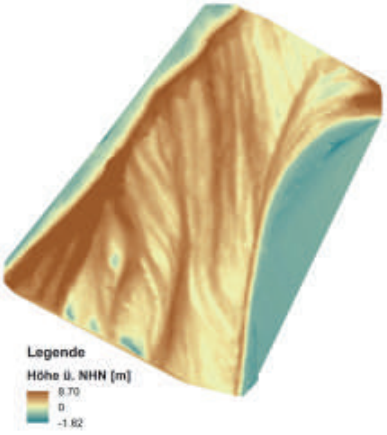


Abb. 8.: DGM der Vogelinsel Langenwerder

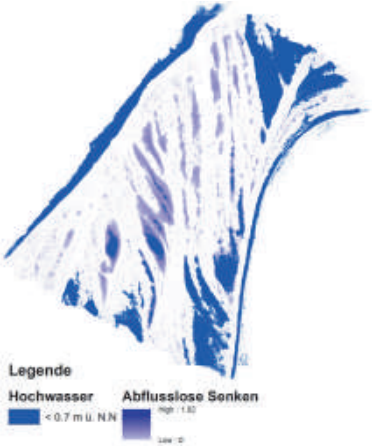


Abb. 9: Hochwasser- und vernässungsgefährdete Senkenbereiche

Zum anderen wird die Insel vor allem bei Sturmwetterlagen auf seiner nordwestlichen Seite angegriffen. Dieses Material wird zur nordöstlichen Seite verlagert und vergrößert den dortigen Haken.
Wie erwähnt suchen sich Vögel ihre Brutstandorte in trockenen, nicht überschwemmungsgefährdete Gebieten. Um dieses zu verifizieren wurden einerseits aus dem DGM abflusslose Senken berechnet, in denen es bei starken Niederschlägen zu Überschwemmungen kommen kann. Andererseits kann bei Hochwasser der Wasserstand kurzfristig 70 cm über Normal liegen und Überschwemmungen verursachen Abb. 9.

6.3 Zusammenfassung und Ausblick

Mit dem Beitrag konnte gezeigt werden, dass es grundsätzlich möglich ist Vögel mit Hilfe von UAS-Aufnahmen mit einer großen Zuverlässigkeit zu zählen. Allerdings lässt sich diese Methodik nicht unbedingt auf andere Vogelarten übertragen, da die untersuchten Möwen einen sehr guten Kontrast gegenüber ihrer Umgebung liefern. Leider wurde in den beiden Jahren keine Zählung der gesamten Sturmmöwenkolonie des Untersuchungsgebiets durch den Verein Langenwerder e.V durchgeführt. Daher können die Ergebnisse nicht mit der klassisch durchgeführten Zählmethode verglichen werden. Die Vogelzählungen mit dem UAS sollen im kommenden Jahr fortgesetzt werden. Schwerpunkt wird die Untersuchung der Ortstreue der Gelege sein.

7 Literaturverzeichnis

BIBBY J. C., BURGESS N. D., HILL D. A. (1995): Methoden der Feldornithologie; Bestandserfassung in der Praxis. Neumann Verlag GmbH

KULEMEYER, C., SCHULZ, A. WEIDAUER, A., RÖHRBEIN, V., SCHLEICHER, K. FOY, T., GRENZDÖRFFER, G. UND COPPACK, T. (2011): Georeferenzierte Digitalfotografie zur objektiven und reproduzierbaren Quantifizierung von Rastvögeln auf See,- Vogelwarte **49**, 2011: S. 105 – 110

LEMKE, M. (2004): Geodatenerfassung und -fortführung für landschaftsökologische und naturschutzkundliche Aufgaben am Beispiel der Indel Langenwerder.- unveröff. Diplomarbeit TU Berlin, 121 S.

SARDÀ-PALOMERA, F., BOTA, G., VIÑOLO, C., PALLARÉS, O., SAZATORNIL, V., BROTONS, L., GOMÁRIZ, S. AND SARDÀ, F. (2012), Fine-scale bird monitoring from light unmanned aircraft systems. Int. Journal of avian science (Ibis), **154**: S. 177–183.

SCHNEIDER, P. (2011): Hochmoderne luftbildgestützte Zählung von Vögeln mit Hilfe eines Multicopters - am Beispiel einer Sturmmöwenkolonie auf der Vogelinsel Langenwerder, Bachelorarbeit, Universität Rostock 39. S.

VON BLOTZHEIM U.N. UND BAUER, K.M. (2000): Handbuch der Vögel Mitteleuropas, Band 8/I, Charadriiformes (3. Teil), Schnepfen-, Möwen- und Alkenvögel, S. 440–493.

WATTS, A. C., PERRY, J. H., SMITH, S. E., BURGESS, M. A., WILKINSON, B. E., SZANTOI, Z., IFJU, P. G. AND PERCIVAL, H. F. (2010): Small Unmanned Aircraft Systems for Low-Altitude Aerial Surveys. The Journal of Wildlife Management, **74**: 1614–1619.

Different Ways to Process UAV Imagery – Results of Comparative Benchmark Tests

HARALD HERDA[1] & MICHAEL BREUER[1]

Abstract: The paper focuses on optimizing workflows to process UAV imagery. For this, one classic photogrammetric software (LPS from INTERGRAPH), two cloud services (Pix4UAV from Pix4D, 123D Catch from AUTODESK) and one CV-based software package (Bundler, Visual SFM, CMP-MVS) were investigated for pros and cons to develop useful combinations. The optimization of the traditional workflow (with LPS) is realized by determining the exterior orientation parameters. They are merely obtained from the raw images and known ground control points via classic bundle adjustment. To evaluate the output data, a reference area has been measured with terrestrial Laser-Scanning. An analysis of the data from the different software packages is made in comparison to that reference.

1 Introduction

UAV have reached the operational level. The user can choose from different commercial systems. In some fields of use (e.g. if the target area is small) UAV can be a valuable addition or even a substitute to (what we call here) "The Traditional Photogrammetric Workflow". The latter relies on capturing (classic) photogrammetric imagery with special photogrammetric sensors (like frame-cameras or line-scanners). Established commercial products are the UltraCAM from Vexcel or the DMC and ADS from INTERGRAPH. They all fly on aerial platforms like single engine aircrafts (e.g. Cessna airplanes). Digital Surface Models (DSM) and Ortho-Image Mosaics are typical geodata-products that can be derived from (classic) photogrammetric imagery using approved workflows. These workflows are more or less standardized, that means, commercial software products are well known and widely used by the photogrammetric community (e.g. products from INPHO, INTERGRAPH etc.). General requirements are well-known because of decades of experience. Hence interior and exterior orientation parameters are given or can be derived with sufficient accuracy. Traditional photogrammetry has reached a high degree of automation. Although UAV are operational, we assert that the workflows for image data processing are not. We conclude from different investigations that the classic photogrammetric workflow causes problems or even fails with UAV imagery. The reason is that both, the platform and the imagery, differ from the traditional approach. An UAV is much more susceptible to turbulences compared to single engine aircraft that are long established for aerial photography flights. The latter is naturally equipped with a gyro-stabilized platform whereas an UAV is usually not. The image frame size (or line length in case of line-scanners) is much bigger with classic aerial image sensors than the cameras that we use with an UAV. The sensors that we use to capture position and attitude (e.g. GPS/INS) are much less sophisticated in case of an UAV compared to a traditional photogrammetric flight. UAV may offer new areas of application, but they also cause new problems that especially arise when we try to turn UAV imagery into reliable geodata. On the other hand, it is important to note that the costs of an UAV

are up to fifty times less in comparison with the investment needed for data acquisition in traditional photogrammetric flights.

2 UAV-System

We use a quadrocopter md4-1000[*] for data acquisition within this project. The company GEO-METRIK Ingenieurgesellschaft mbH provided the system. Its main characteristics are:

- payload (max.): 1.2 kg
- flight time: up to 60 minutes
- waypoint-navigation

The UAV is equipped with the Micro Four Thirds (MFT[**]) camera Olympus Pen E-P2 that has a fix 17mm lens. These system specifications formed the guideline for further planning of project parameters. However we take into account that the focal length may vary up to some microns due to automatic focusing (we tested the camera with infinity focus and automatic focusing switched off. For unknown reasons images were blurred even for objects at infinite distance. Hence images were taken with active automatic focusing assuming that an almost constant flying height would result in almost constant focal length).

2.1 Flight Planning

The test field has an area of 10 ha and is located close to the city of Tangermünde (Saxony-Anhalt, Germany). Main reasons for choosing this area are the existence of various types of surfaces (e.g. federal highway, slopes, ordinary road) and existing flight permission up to 300 m.

First step is a rough definition of the project parameters. Due to LUHMANN (2010) these are:

- object size,
- accuracy verification,
- accuracy to be achieved in object space (tolerance).

An analysis of that rough planning is followed by specific solutions in detail.

- forward- and side overlap
- camera calibration approach
- camera positions in object space
- sensor selection (image format, focal length)
- achievable accuracy of image point coordinates
- determination of optical parameters (exposure, depth of field)

The following mission parameters are chosen based on references found in LUHMANN (2010), EISENBEIß (2009) and KRAUS (1990):

[*] Manufacturer: www.microdrones.com
[**] MFT is a standard developed by OLYMPUS and Kodak

- image overlap area (80%)
- ground sampling distance (GSD): 50 mm
- ground control points (GCP) homogeneously distributed (9)

2.2 Reference data acquisition

A terrestrial survey with a HDS 7000* laser scanner delivered a dense 3D point-cloud with a point density of approx. 100 points per m² (Figure 1). Special focus was put on getting an almost complete coverage of all road surfaces. They are estimated to remain unchanged over time and serve as test areas for the latter accuracy investigation. This assumption fails for the surrounding areas that are covered by meadows and low vegetation. A regular raster Digital Surface Model (DSM) with a ground spacing of 10 cm was interpolated from the 3D point-cloud. In compliance to KRAUS (1990) and LUHMANN (2010) the height accuracy was estimated to 24 cm (σ) and the planar accuracy to 5 cm (σ). Hence the DSM serves as an eligible reference data set to evaluate the subsequent photogrammetric products that we got from the different workflows based on the UAV imagery.

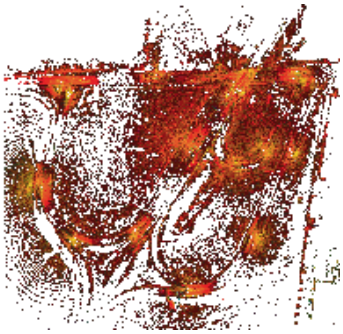


Figure 1 – Terrestrial 3D point-cloud

The field survey took an entire working day. Most of the time was spent on Laser-Scanning. Image capturing during the UAV flight was completed in 8 minutes. The flight campaign yielded 67 aerial images in 5 flying lines.

3 Software

The field of photogrammetry and computer vision overlaps comprehensively. Often different mathematical models are used to describe the same or similar phenomena. Linking the two disciplines can lead to improved results. In addition to commercial and non-commercial software packages also cloud services (cost-free and commercial) are evaluated. The objective is to create DSM and Ortho-Images with all programs. In the following section the different software and the associated workflows are described. Furthermore, strength and weaknesses are identified.

3.1 Traditional photogrammetric approach

The Leica Photogrammetry Suite (LPS), part of the INTERGRAPH product portfolio, is a software application that extracts information from images using traditional photogrammetric operations (point matching on interest points using epipolar geometry, aerial triangulation using bundle adjustment). Product development goes back to the year 1991. LPS is the only software solution in the project that supports the import of exterior orientation parameters. The results of processing are:

- 3D point-cloud

* Manufacturer: www.leica-geosystems.com

- Digital Surface Model
- Ortho-Images
- georeferenced Ortho-Mosaik
- improved parameters of interior and exterior orientation

3.2 Cloud services

In addition, two online services were examined to process UAV images. The main feature of the hereinafter presented cloud services is a fully automatic analysis of the input data. We understand "fully automatic" with respect to a relative orientation in a local coordinate system. A possible georeferencing to a world coordinate system requires manual intervention to pick GCPs.

The Swiss company Pix4D offers a service called Pix4UAV that is designed to process commercial UAV applications. Next to LPS it offers the opportunity to pick GCPs in the raw image data and upload according coordinates.

Autodesk 123D Catch is another fully automated web service that can generate textured surface models. Input data are the pure image data without additional information about the camera positions. There is no published documentation about the program workflow and used algorithms. Also the output data is merely a textured surface model.

3.3 CV-based Freeware

Freeware is defined as software that is provided by the author for free use. It is not synonymous with the terms "open source" or "free software". This means, among other things, that the source code should not be available and no change to the software is allowed. Additionally the author can grant specific rights like non-commercial usage.

The software Bundler[*] is a structure-from-motion (SFM) system developed by NOAH SNAVELY at Cornell University, USA. It is designed to produce a sparse 3D point-cloud out of a set of unordered images. This output in not scaled because the input is just the raw image data. Thus camera orientation and position are available in a local coordinate system.

Also VisualSFM[**] is a structure-from-motion system. It combines Bundler with a graphical user interface. In addition it includes CMVS/PMVS2 for generating dense point-clouds. It is written by CHANGCHANG WU at the University of Washington in Seattle, USA. In contrast to the Bundler stand-alone version it uses multicore CPU as well as multicore GPU for feature detection, matching and bundle adjustment. It is up to 10 – 30 times faster (WU et al., 2011) than current state of the art methods (AGARWAL et al., 2010). The multi-view reconstruction software CMP-MVS[***] is written by MICHAL JANCOSEK at the Czech Technical University in Prague. The input to this software is an undistorted set of images with the associated camera parameters (interior and exterior camera calibration). The necessary data can be exported from VisualSFM. The most popular feature next to the automatic generation of a point-cloud, a DSM and a video animation is the possibility to create an Ortho-Image.

[*] http://phototour.cs.washington.edu/bundler/
[**] http://homes.cs.washington.edu/~ccwu
[***] http://ptak.felk.cvut.cz/sfmservice/websfm.pl?menu=cmpmvs

3.4 Pros & Cons

Tests with different UAV image-sets were made to identify pros and cons of the different software products. These experiences are used to develop useful combinations that compensate or eliminate discovered weaknesses. Due to processing various criteria were defined to compare the different approaches (Table 1).

Table 1 - Results of Comparative Benchmark Tests

Criteria	Bundler	VisualSFM	CMPMVS	LPS	Pix4UAV	123D Catch
Ease of use	+	+	+	-	/	/
Range of functions	-	+	-	++	/	/
Keypoint matching	++	++	++	-	/	/
Image orientation	+	+	+	+	++	/
Georeferencing	--	-	-	+	++	/
Trouble shooting	--	+	--	+	-	--
Documentation	+	++	++	++	-	--
Costs	++	++	++	--	-	++

The appraisal reflects the end-user's perspective. This explains why the most expensive software (LPS) has a double minus for costs. We did not rate some criteria of the cloud services because they are hidden from the end-user. We discovered that the keypoint matching is the main weakness of the traditional LPS workflow. The latter provides an extensive diagnostic analysis and report framework. Nevertheless a suppression of numerous generated blunders failed. The Open Source Tools scored at keypoint matching but there is no accurate solution to provide a relation to the world coordinate system. A simple Helmert-Transformation showed promising results at processing the UAV images of the test area, but a bundle adjustment including known GCPs and standard deviations is essential to avoid possible deformations with irregular image datasets like those shown in Figure 2.

The almost fully automated workflow of Pix4UAV and Bundler & Co. compared to a high amount of manual intervention in LPS seems to be advantageous. Nevertheless – from a professional user's point of view – it is desired to be able to specify certain parameters e.g initial values and standard deviations of exterior orientation parameters, image and ground control point measurements. A visual check of the output DSM is necessary to identify blunders before generating Ortho-Images. The basis of a successful aerotriangulation is reliable feature point detection. The problems with processing UAV imagery with the traditional LPS workflow are dominated by blunders that arise from the automatic tie point measurement. This leads to erroneous results of the bundle adjustment and the image orientation. They appear at the following items:

- accuracy of adjusted exterior orientation
- residuals of GCPs
- height of reconstructed object-points

The reasons can be found in conjunction with the acquisition of reliable approximate values for exterior orientation parameters as well as the detection and elimination of blunders (LUHMANN, 2010). We specified different standard deviations of the GCPs and camera positions without getting any improvement of the results.

4 Optimization

Approaches for optimization were established to handle these issues. The aim is to provide interim results from low-cost 3D-reconstruction techniques for further processing in LPS. CV-based algorithms for feature-detection (SIFT) and robust feature matching (RANSAC) show promising results (ANDERSEN et al., 2006; LOWE 2004).We tried to get adjusted camera positions from Bundler to introduce them as initial values for bundle adjustment with LPS. This was done because we had assumed the above mentioned problems to be the main reason for obstructions detected during processing in traditional workflows.

4.1 Approach & Implementation

VisualSFM (see 3.3) determines the camera orientation solely on homologous feature-points. As the input data is just the raw UAV imagery, this information is un-scaled with no reference to any world coordinate system. A transformation of the rotation matrices to global roll and nick angles is unsolvable for principal mathematical reasons. Nevertheless a definite link between the reconstructed terrain-surface (point-cloud) and the exterior camera parameters (initially defined in a local coordinate system) is established. VisualSFM provides the camera-alignment in form of a quadruplet of points representing the four edges we got the idea to deduct new approximations for the exterior orientation parameters from the quadruplets. First we calculated the parameters of a 7-parameter Helmert-Transformation from a dense 3D point-cloud that we linked to a world coordinate system. The linkage was done by using control points with known coordinates in both coordinate systems. As soon as we got the transformation parameters we could transform each quadruplet from the local to the world coordinate system. We had to take into account that exterior orientation parameters refer to projection centers. Hence we deduced the location of each camera from the quadruplets. Finally the parameters have to be pre-processed to be conforming to individual LPS block-properties. As a result we discovered significant differences between the rotation angles from the microdrone's log-file and the calculated exterior orientation parameters. The average deviation has an amount of 10,4° in roll-axis, 4,1° in pitch-axis and 8,6° in yaw-axis.

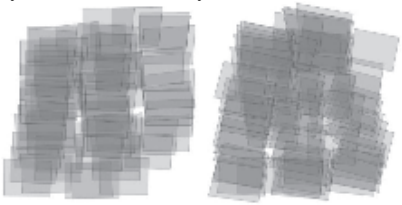


Figure 2 – Image-footprints from microdrone's log-file (left) and newly calculated exterior orientation parameters (right)

	roll [°]	pitch [°]	yaw [°]
min	-9.38	-14.19	-0.20
max	44.14	22.72	23.78
mean	10.44	4.13	8.62

Table 2 – Deviation between rotation angles from microdrone's log-file and calculated values

4.2 Results

With the newly calculated values of the exterior orientation the automatic tie point measurement had been performed again with the same set of images (see 2.2). For the first time the LPS bundle adjustment showed realistic results. Finally the generation of an Ortho-Image is functioning well. Besides, an Ortho-Image had been produced with the software-combination VisualSFM and CMPMVS, as well as with the web-service Pix4UAV. Figure 3 shows the differences in quality.



Figure 3 – Different Ortho-Image results



Figure 4 – Moving objects removal: CMPMVS (left), LPS (right)

The DSM of Pix4UAV is very noisy. Due to the fully automated creation of the Ortho-Image, there are distinct misinterpretations in the appearance of height. In LPS the figure corresponds to a typical manually selected single Ortho-Image where elevated objects are radially distorted. This effect is small in the given subset because the object is located close to the nadir. CMPMVS provides by far the best result in the test. Elevated objects (e.g. billboards, signs etc.) disappear. The same applies to moving objects such as vehicles (Figure 4). Both are ignored at further reconstruction.

The whole set of images has pictured an amount of 113 vehicles. The Ortho-Image generated by CMPMVS shows neither vehicles nor their shadows. Also details like road-signs are unconsidered at Ortho-Image generation. This result is a very good approximation of a so-called "True-Ortho-Image" where each single pixel is rectified correctly.

4.3 Analysis

For the evaluation of the output data we investigated the 3D-point clouds as result from the different workflows. We used the regular raster DSM as reference (see 2.2). We georeferenced each 3D point-cloud to a world coordinate system first. Then we interpolated a raster DSM with the same spacing of the reference data-set. Finally we subtracted the results produced with Bundler, VSFM, CMPMVS, Pix4D, 123Catch and LPS from the reference data. Table 3 shows the results. The quality assessment of differences in X, Y and Z illustrates that Autodesk 123D Catch obtains the most inaccurate georeferencing. Main reason therefore is the sparse point density of 88.000 in contrast to e. g. VSFM with 4.5 million points. For this reason it is not achievable to pick the GCPs with the same accuracy in each point-cloud.

Table 3 – Differences in X, Y and Z

mean error rmse	Bundler [m]	VSFM [m]	CMPMVS [m]	Pix4UAV [m]	123D Catch [m]	LPS [m]
X	0,064	0,036	0,027	0,010	0,068	0,015
Y	0,036	0,027	0,051	0,020	0,072	0,028
Z	0,018	0,027	0,044	0,030	0,063	0,037
σ	0,076	0,052	0,073	0,037	0,117	0,049

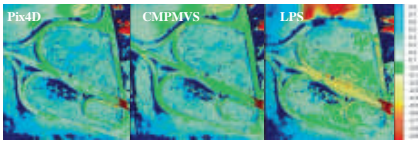


Figure 5 – DSM quality evaluation (green: good fitting)

No reference data exist in the dark red and dark blue areas. There were ten weeks between Laser-Scanning and the UAV flight. This explains the height differences in the light blue areas due to vegetation growth. As expected we observe the slightest deviations in the areas of road surfaces (green: +/- 5 cm). Bundler produced the best results, whereas 123D Catch was the worst (for details see Table 4).

Table 4 - Quality assessment of height differences in specific test areas

standard deviation Z	Bundler [m]	VSFM [m]	CMPMVS [m]	Pix4UAV [m]	123D Catch [m]
access north	0,029	0,052	0,021	0,032	0,079
access south	0,027	0,043	0,032	0,033	0,146
federal highway	0,027	0,052	0,037	0,032	0,087
federal road	0,044	0,074	0,046	0,058	0,239
cycle lane	0,06	0,091	0,049	0,033	0,121

5 Conclusions

The processing of UAV images with different applications shows a high potential in multidisciplinary combination of algorithms. Traditional photogrammetric software poses unexpected challenges with the typical unordered image collection obtained by UAV-flights. The image orientation through bundle adjustment on the basis of consistent feature matches requires considerable manual intervention. Approximate values of the exterior orientation parameters are fundamental for the creation of reliable geodata. The log file of the md4-1000 is useless for further processing because it describes the attitude of the drone and not of the camera itself. The paper shows the possibility to determine the camera position and alignment from just the UAV

image data and known ground control points. With the aid of these calculated parameters the traditional LPS workflow with a block of five stripes is functioning well. Within an eight minutes flight an area of 10ha was captured. At a flying height of 200 m an object accuracy of less than σ = 5 cm was achieved (see table 4). The expected accuracy in X and Y has been confirmed and corresponds well to the aerial image GSD of 5 cm The achieved accuracy in height is better than expected by a factor of 4.

Future work will focus on two main aspects. On the one hand it is planned to improve flight- and hardware-issues. Therefore tests with different flying altitudes, image overlap and other camera sensors will be made. Additionally an evaluation of an IMU, mounted directly on the camera, is envisaged in comparison with the developed calculation of the exterior orientation. On the other hand it is planned to enhance the quality and reliability of the generated output, especially the DSM. In detail there will be exercises with different methods of filtering to remove noise and blunders. In addition a statistical analysis of the reconstructed point-cloud without gridding (interpolation) is designated.

6 Literature

AGARWAL, S., SNAVELY, N., SEITZ S. & SZELISKI R., 2010: Bundle adjustment in the Large. In *ECCV10*, 2010. pp. 29–42.

ANDERSEN, V., PELLARIN, L. & ANDERSON, R., 2006: Scale Invariant Feature Transform (SIFT): Performance and Application", Copenhagen, Denmark. http://www.itu.dk/people/vedrana/reports/SIFT.pdf (01 Dec. 2012).

EISENBEIß, H., 2009: UAV PHOTOGRAMMETRY, DISSERTATION, ZURICH, SWITZERLAND. PP. 62 - 76. www.igp-data.ethz.ch/berichte/Blaue_Berichte_PDF/105.pdf (01 Dec. 2012).

KRAUS, K., 1990: Photogrammetrie - Grundlagen und Standardverfahren, Dümmler, Bonn, pp. 147

LOWE, D., 2004.: Distinctive image features from scale-invariant keypoints. In *International Journal of Computer Vision, 2004*. pp. 97-110.

LUHMANN, T., 2010: Nahbereichsphotogrammetrie - Grundlagen, Methoden und Anwendungen, Wichmann, Berlin, pp. 303, 529.

WU, C., AGARWAL, S., CURLESS, B. & SEITZ, S., 2011: Multicore Bundle Adjustment. CVPR 2011, Colorado Springs, USA.

6.1 Acknowledgements

Investigations and results of this paper emerged from the ongoing research project “Entwicklung von Methoden zur automatisierten Erzeugung von Geoinformation aus Bilddaten ferngesteuerter unbemannter Flugsysteme (UAS)” (Development of methods for automated production of geoinformation based on unmanned aerial systems’ imagery). This is a subproject of the research funding "Forschungsassistenz VI at Beuth Hochschule für Technik Berlin – University of Applied Sciences". It is founded and financed by the European Social Fund and the Berlin Senate Department for Economics, Technology and Research. It focuses on strengthening the regional economy and opens up the opportunity for new developments and further advancement of existing techniques, methods and procedures. For further information please follow the link:

http://www.beuth-hochschule.de/forschungsassistenz/

Forest Change Assessment and Corresponding Driver Analysis in the Magdalena Department, Colombia (1985-2010)

TEJA KATTENBORN[1]

Seit Jahrzehnten ist die alarmierende Rate der Tropenwaldverluste begleitet von empfindlichen Rückgängen der Biodiversität Gegenstand der Forschung. Jedoch wurde bisher weniger Focus auf trockene Tropenwaldgesellschaften gerichtet, den dominierenden Waldtyp der karibischen Ökoregion Kolumbiens.
Auf Grundlage von Landsat TM/ETM Daten wurden für die Jahre 1985, 2003 und 2010 über NDVI-Schwellwertbildung alle dicht mit Vegetation bedeckten Flächen ermittelt und landwirtschaftliche Flächen nach NN-Klassifizierung ausmaskiert.
Der immense Verlust an Waldflächen und damit Biodiversität wurde bestätigt und dokumentiert; tatsächlich beläuft sich die im zweiten Erhebungszeitraum 2003 bis 2010 fast verdoppelte Netto-Entwaldungsrate auf 126 km²/Jahr. Die Muster der Waldentwicklung konnten anhand sozioökonomischer Faktoren statistisch analysiert und erklärt werden.

1 Motivation und Zielsetzung

Kolumbien gehört zu den Ländern mit der höchsten Biodiversität, die seit Jahrzehnten durch Landnutzungsveränderungen hohem Druck ausgesetzt ist. Damit sind auch die Umweltleistungen der dortigen Waldbestände gefährdet sowie ihre sozio-ökonomische Nutzen.
Die Provinz Magdalena an der kolumbianischen Karibikküste wurde während der letzten Jahrzehnte in ihrer sehr dynamischen Entwicklung durch politische Instabilität, Drogenanbau, bewaffnete Konflikte und die für ein Entwicklungsland typische Armut geprägt. Die dortige Entwaldungsrate zählt zu den höchsten der Welt.

Im Rahmen eines Deutsch-Kolumbianischen Kooperationsprojektes mit dem Ziel, nachhaltige Forstwirtschaftskonzepte unter Betrachtung des Potentials von Lesser-Known Species (LKS) für diese Provinz zu entwickeln, war eine Fernerkundungspart vorgesehen, um Karten und Statistiken der Waldentwicklung seit 1985 zu erstellen, die maßgeblich durch die hohen Migrationsraten aufgrund langer gewaltsamer Konflikte geprägt ist. In den letzten Jahren haben sich in dieser Region die wirtschaftlichen Aktivitäten der meist kleinen “Stakeholder“ normalisiert. Damit ist die Gefahr der Waldumwandlung in Anbauflächen und Viehweiden für die verblieben Sekundärwaldflächen wieder aktuell.

Im Rahmen des Gesamtprojektes hatte die vorliegende Studie folgende Zielsetzungen:

- Feststellung der Waldflächenveränderung seit den politischen Unruhen der 80er und der nachfolgenden stabileren Jahre seit 2003
- Analyse von Infrastruktur und Populationsdaten als Parameter für Ursachen der Walddegradation

1) Teja Kattenborn, Abteilung für Fernerkundung und Landschaftsinformationssysteme, Albert-Ludwigs-Universität Freiburg, Tennenbacherstr. 4, 79104 Freiburg, teja.kattenborn@felis.uni-freiburg.de

2 Material und Methoden

2.1 Die Provinz Magdalena

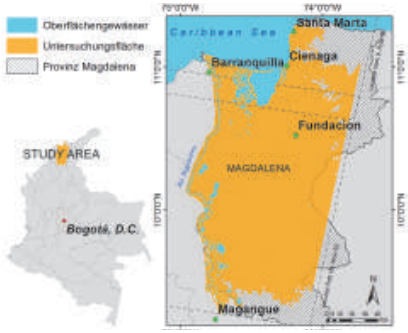


Abb. 1: Untersuchungsgebiet in der Provinz Magdalena, Kolumbien

Das Untersuchungsgebiet mit einer Größe von 20073,13 km² befindet sich im Norden Kolumbiens zwischen Lat. 9°16´ - 11°19´ und Long. -73°43´ - -74°57´ (Abb. 1) und ist von einem innertropischen Klima geprägt. Die Ebenen (0-500 m) der Provinz Magdalena werden von Semi Decidous Tropical Forest (kurz SDTF) dominiert, wobei nahe Gewässern Mangroven und andere hygrophile Waldarten vorherrschen (OLSEN & DINERSTEIN, 2002).

Die Hügel und Berge westlich von der Sierra Nevada de Santa Marta (kurz SNSM) westlich von Fundación bzw. Ciénaga zeichnen sich durch montane Varianten von Savannen und Tropenwäldern aus. Der Marihuana-Anbau der Campesinos in den 1970ern später auch Coca-Anbau hatten ausgedehnte Waldzerstörungen in den niederen Regionen der SNSM zur Folge. Auf einigen dieser Flächen konnte sich sekundärer SDTF entwickeln (CAVELIER et al., 1998; FPSNSM, 1991).

Die meisten Siedlungen – analog die höchste Bevölkerungsdichte - befinden sich an der Küste und entlang des Rio Magdalena. Letztere stieg zwischen 1985 und 2011 von 38 auf 52 Einwohner/km² (DANE, 2012). Die Bevölkerung setzt sich aus diversen Gruppierungen zusammen wie den Campesinos, welche aufgrund von politischen Unruhen während den 1950ern aus dem Landesinneren eingewandert sind, vier indigenen Kulturen in den SNSM sowie militärische und paramilitärische Gruppen (FPSNSM, 1991; DAVALOS & BEJERANO, 2008).

Die Wirtschaft der Region basiert hauptsächlich auf Tourismus, Handel und arbeitsintensiver Bewirtschaftung von Bananen- und Palmölplantagen. Aufgrund des anhaltenden Bedarfs an Agrarflächen sowie des Verbrauchs von Holz kommt es nach wie vor zu ausgedehnten und vor allem unkontrollierten Waldflächenverlusten sogar in Schutzgebieten wie der SNSM durch Coca-Anbau. (RESTREPO & SYVITSKI, 2012; DAVALOS & BEJERANO, 2008).

2.2 Feldkampagne

Eine Feldkampagne wurde durchgeführt, sowohl, um mit den lokalen Landnutzungsformen und Vegetationsverhältnissen vertraut zu werden, als auch um nötige Informationen für die Interpretation der Satellitendaten und Verifikation der Ergebnisse zu gewinnen. Dabei wurden die vorherrschenden Landnutzungen stichprobenweise beschrieben und mit einer GPS-Kamera dokumentiert.

Des Weiteren wurden Ansässige und lokale Experten befragt, um spezifische Einblicke und Daten der regionale Kultur und Situation sowie Triebkräfte der Waldentwicklung zu erhalten. Die Felderhebungen umfassen zusätzlich eine systematische Forstinventur von etwa 80 ha Savanne und SDTF im Kerngebiete des Rahmenprojektes (PAREDES, 2012).

2.3 Prozessierung der Satellitendaten zur Analyse der Waldentwicklung

Da die Waldvorkommen des Untersuchungsgebiets hauptsächlich aus SDTF bestehen und Strategien zu dessen Erhalt Hauptzielrichtung des Rahmenprojektes waren, wurde ein Change Detection-Ansatz entwickelt, welcher hauptsächlich auf Basis eines Vegetationsindexes Waldflächen sicher identifiziert. Auf weitergehende Verfeinerungen z.B. nach Waldtypen konnte im Rahmen der Zielsetzung verzichtet werden.

2.3.1 Datenakquise und Vorprozessierung

Aufgrund vegetationssensitiver Sensorspezifikationen sowie guter zeitlicher und räumlicher Abdeckung wurden Landsat/TM/ETM+-Daten als Grundlage für die Change Detection-Studie gewählt. Drei Landsat-Abdeckungen (Pfad 08, Reihe 52/53) der Jahre 1985 (TM), 2003 (ETM+) und 2010 (TM) wurden identifiziert, welche den Großteil der Provinz Magdalena abdecken, aufgenommen jeweils zwischen dem 24. und dem 29. Januar, so dass jahreszeitliche Dynamiken die Change Detection-Ergebnisse nicht beeinflussen. Da dieser Zeitraum im Zentrum der meteorologisch stabilen Trockenzeit liegt, sind die Daten zudem geringer von wechselnden atmosphärischen Bedingungen wie Nebel oder Dunst beeinflusst. (BRUCE & HILBERT, 2006)
Aus den so vorliegenden Rohdaten wurden zunächst anhand von post-Launch Gain and Offsets Top of Atmosphere-Reflektanzen berechnet. Zur Atmosphärenkorrektur wurde die Dark Object Subtraction angewandt, da diese keine in situ-Messungen benötigt, schnell implementierbar ist und dabei sehr gute Ergebnisse liefert (SONG et al., 2001).
Kleinflächig auftretende Cirrus- und Stratuswolken wurden auf Grundlage einer Orthogonal Subspace Projection (OSP) klassifiziert und anschließend maskiert.
Aufgrund der äquivalenten Aufnahmegeometrie der Landsat-Szenen und der überwiegend flachen Topographie wurde in der Auswertung auf eine topographische Korrektur verzichtet.

2.3.2 Normalized Difference Vegetation Index zur Walflächenanalyse.

In Tests verschiedener Vegetationsindizes (NDVI, SAVI, EVI, TCT) lieferte der Normalized Difference Vegetation Index (NDVI) die besten Ergebnisse. Damit wurden eine Reihe von ähnliche Studien in der Nutzung von Landsat-NDVI für die Analyse von SDTF-Formationen bestätigt (MARTINUZZI et al., 2007, BARREDA-BAUTISTA et al., 2011).
Um den direkten, quantitativen Vergleich in der Change Detection-Analyse zu ermöglichen, müssen die Datensätze der Zeitreihe radiometrisch angeglichen werden. Unter der Annahme, dass Wasser, Fels und dicht bestandene Waldflächen z.B. in Schutzgebieten oder unzugänglichen Tälern zeitlich stabile spektrale Reflexionseigenschaften aufweisen, wurden deren NDVI-Werte zum radiometrischen Angleich genutzt. Zur Anwendung der Psdeuo Invariant Feature (PIF)-Methode (BRUCE & HILBERT, 2006), wurden auf Basis der Feldaufnahmen 14 vegetationsfreie bzw. dicht bewachsene PIF ausgewählt. Um potentielle spektrale Veränderungen der PIF zwischen den Zeitperioden zu minimieren, wurde der mittlere Aufnahmezeitpunkt 2003 als Referenzspektrum gewählt, an das die NDVI-Spektra der Jahre 1985 und 2010 angeglichen wurden.
Im Anhalt an MARTINUZZI et al. (2007) wurde auf Basis von Ground Truth, panchromatischen ORBVIEW3-Daten und Landnutzungskarten des JRC/CIAT TREES-II-Projektes (ACHARD et

al., 2002) ein NDVI-Schwellwert bestimmt, der Waldbestände nach der FAO-Definition (FAO, 2005) von anderen Vegetationstypen trennt.
Um aufgrund ihrer hohen NDVI-Werte ebenfalls erfasste Bananen- und Palmenölplantagen auszuscheiden, wurden Maximum Likelihood-, Parallel Piped- und Neural Network-Klassifizierer mit originalen und synthetischen Landsat-Kanälen getestet, wobei die Kombination von Tasseled Cap-transformierten Landsat-Kanälen und Neural Network-Klassifikation die besten Ergebnisse lieferte. Somit konnte die aktuelle Waldfläche identifiziert werden, indem die aus der NDVI-Schwellwertbildung erhaltenen Vegetationsflächen mit jenen der klassifizierten Plantagen verschnitten wurden.
Die entsprechenden Waldflächen der Jahre 1985, 2003 und 2010 dienten folgend dem Prinzip von ELMQVIST et al. (2007) zur Erstellung der Waldentwicklungsklassen (Tab. 3).

2.4 Korrelation zwischen Waldentwicklung und potentiellen Triebkräften

Für eine Bewertung der potentiell für die Walddegradation bzw. -flächenveränderungen verantwortlichen sozio-ökonomischen Triebkräfte standen folgende Parameter zur Verfügung: A) Transportnetzwerk bzw. Zugänglichkeit, B) Entfernung zur nächsten Siedlung und damit Holzmarktnähe, C) Populationsdichte als Indikator für direkten Holzverbrauch D) Naturschutzgebiete E) Brandflächen (Tab. 1).

Tab. 1: Parameter zur Analyse der Beziehung zwischen Waldentwicklung und sozio-ökonomischer Triebkräfte, Provinz Magdalena, Kolumbien

Triebkraft	*Parameter*	*Räuml. Auflösung*	*Datenquelle*
A) Zugänglichkeit	Euklidische Distanz zum Transportnetzwerk [m]	90-m grid	Vorliegende Studie
B) Marktnähe	Euklidische Distanz zur nächsten Siedlung [m]	90-m grid	GRUMP
C) Holzverbrauch	Populationsdichte [Bewohner/km²]	5-km grid	UNEP-GRID
D) Schutzstatus	Naturschutzgebiete [binär 0/1]	90-m grid	SIGAC
E) Slash & burn	Klassifizierte Brandflächen in 1985, 2003, 2010	30-m grid	Vorliegende Studie

Die Beziehung zwischen diesen Parametern und den Waldentwicklungsklassen (Waldverlust, stabiler Wald und Waldzuwachs seit 1985) wurde sowohl für geeignete Sub-Regionen als auch für das gesamte Gebiet in R-Statistical Language untersucht. Dabei wurde Spearman´s Korrelationskoeffizent genutzt, da einige dieser Parameter keine Normalverteilung aufwiesen.

3 Ergebnisse

NDVI-Schwellwert und Maskierung von landwirtschaftlichen Flächen durch NN-Klassifikation ergaben die Verteilung der Waldflächen (Tab. 2), wobei deutlich wird, dass lediglich 22,6% des gesamten Untersuchungsgebietes in 2010 bewaldet sind. Anhand der erstellten Waldentwicklungsklassen (Tab. 3) lässt sich erkennen, dass 60% des gesamten Waldverlustes zwischen 1985-2010 seit 2003 stattfand. Hierdurch ergibt sich für den Zeitraum 2003-2010 eine Nettowaldverlustrate von 126 km², die sich im Vergleich zum Zeitraum 1985-2010 mehr

Tab. 2: Wald- und Plantagenfläche, Provinz Magdalena, Kolumbien, 1985/2003/2010

Jahr	*Wald (km)²*	*%*	*Plant. (km²)*	*%*
1985	6368,75	31,7	337,6	1,6
2003	5417,29	26,9	619,7	3,1
2010	4534,60	22,5	808,9	4,0

als verdoppelt hat.

Die Waldverluste hauptsächlich an SDTF fanden überwiegend in ländlichen und abgelegenen Gebieten statt aber auch in Hotspots nahe den klassifizierten Plantagen und in sekundären SDTF östlich von Fundación (Abb. 2).

Stabiler Wald wurde überwiegend in schwer zugänglichen Gebieten der SNSM über 450 m NN oder in den Sümpfen der Lagune Ciénaga de Grande de Santa Marta identifiziert. Kleinere Flächen befinden sich am Rande von Flussläufen oder an steilen Hängen und Tälern. Waldzuwachs fand verstreut auf ehemaligen Rodungsflächen und in Clustern nahe Baranquilla, Ciénaga und Santa Marta statt.

Die ermittelten Waldentwicklungsklassen und sozio-ökonomischen Faktoren weisen generell eine hohe Korrelation auf (Spearman´s rho).

Davon wies Zugänglichkeit (Abb. 3) die stärkste Korrelation auf. Die Beziehung zeigt, dass Waldflächenverluste vermehrt in leichter zugänglichen Gebieten auftreten. Entsprechend befindet sich das erste Quantil der Waldflächenverluste innerhalb einer Distanz von 212 m zum Verkehrsnetzwerk. Stabiler Wald hingegen befindet sich hauptsächlich in weniger zugänglichen Gebieten. Waldflächenzuwachs weist kein deutliches Muster.

Die Entfernung zu Siedlungen wies ein ähnliches Muster auf, wobei die gleichwohl starke Korrelation geringer ausfiel. Auffallend ist jedoch die leichte Zunahme von Waldflächenzuwächsen in Nähe zu Siedlungen.

Verlust an Waldflächen konzentriert sich in Gebieten mit geringen Populationsdichten (23 Einw./km²). Mit höheren Populationsdichten steigt der Anteil von Waldflächenzuwachs (32 Einw./km²) und stabilen Wäldern (46 Einw./km²), letztere

Tab. 3: Waldentwicklungsklassen nach Fläche und Prozent, Provinz Magdalena, Kolumbien, 1985-2010

Waldentwicklungsklasse	*Logisch. Ausdruck**	*in km²*	*in %*
Waldverlust 1985-2010	1985 ⌐(2003 ∨ 2010)	3019,7	15,0
Waldverlust 2003-2010	2003 ^ ⌐ 2010	1829, 3	9,1
Stabiler Wald 1985-2010	1985 ^ 2003 ^ 2010	2925,2	14,6
Waldzuwachs 1985-2010	2010 ^ ⌐ 1985	1185,5	5,9
Waldzuwachs 2003-2010	2010 ^ ⌐ 2003	946,6	4,7

*^ steht für "und"; ⌐ steht für "nicht"; ∨ steht für "oder"

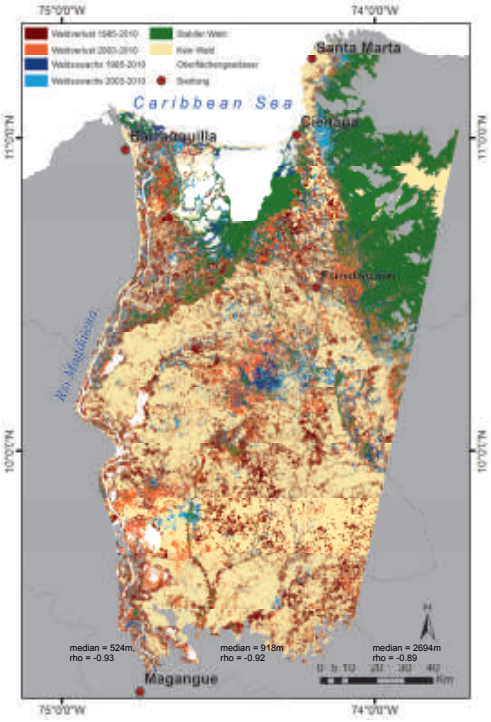


Abb. 2: Karte der Waldentwicklung, Provinz Magdalena, Kolumbien, 1985-2010

befinden sich überwiegend in dichter besiedelten Gebieten.

Schutzgebiete weisen einen deutlich erhöhten Anteil von stabilen Wäldern auf (Tab. 4). Dabei stieg der Anteil stabiler Wälder im Zeitraum 2003-2010 deutlich an.

Tab. 4: Relativer Anteil der Waldentwicklungsklassen in geschützten und nicht geschützten Gebieten, Provinz Magdalena, Kolumbien

	Zeitraum	*Seit 1985*	*Seit 2003*
Geschützt	*Verlust*	4,7%	0,9%
	Stabil	77,9%	90,1%
	Zuwachs	17,3%	8,9%
Ungeschützt	*Verlust*	47,1%	36,0%
	Stabil	36,3%	46,2%
	Zuwachs	16,5%	17,7%

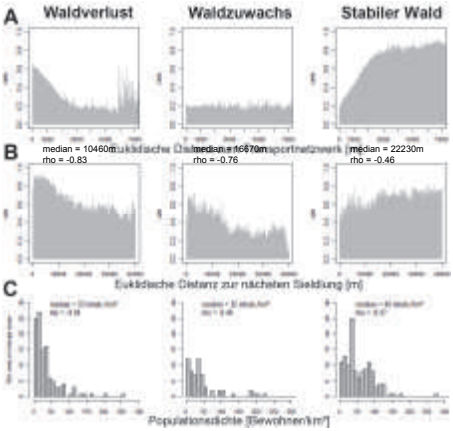


Abb. 3: Beziehung der Waldentwicklungsklassen zu Zugänglichkeit, Marktnähe und Populationsdichte, Provinz Magdalena, Kolumbien, 1985-2010

4 Diskussion und Schlussfolgerung

Unter operationellen Bedingungen wurde eine modulare, flexible Prozessierungskette entwickelt. Korrektur der Satellitendaten und deren Auswertung bezüglich der Waldentwicklung und der korrespondierenden Triebkräfte sind als Blaupause auf ähnliche Regionen übertragbar, wiederholbar und somit ein wertvolles Werkzeug für Monitoring und Erfolgskontrolle implementierter Maßnahmen. Überschlägig kalkulierte Kosten für das Verfahren belaufen sich auf 0.46 €/km². Damit konnten mit geringem Mitteleinsatz wertvolle, detaillierte Informationen in Form von Karten und Statistiken erstellt werden.

Die alarmierende Waldzerstörung und der damit einhergehende Verlust von Ökosystem-Leistungen wurde von vorherigen Studien auf nationalen und kontinentalen Skalen belegt (z.B. TREES II, FAO-FRA). Dies konnte in der vorliegenden Studie auf regionaler und lokaler Ebene für einen Untersuchungszeitraum von drei Jahrzehnten bestätigt und analysiert werden. Dabei wurden die Waldzerstörungstrends für die Provinz Magdalena zum ersten Mal kartiert und gleichzeitig deren sozio-ökonomischen Triebkräfte analysiert.

Der Großteil der Entwaldung fand in den leicht zugänglichen Ebenen (siehe Abb. 4) der Provinz Magdalena statt. Vergleichbar zu einer ähnlichen Studie von ELMQUIST et al., (2006) fand weitreichende und intensive Entwaldung in dünn besiedelten Gebieten mit großen Distanzen zu den Märkten statt. Diese Regionen zeichnen sich insbesondere durch veraltete Landwirtschaftspraktiken, Unsicherheit durch Paramilitär- und Guerillapräsenz, Armut begleitet von einer erhöhten Abhängigkeit von natürlichen Ressourcen z.B. Brennholz. Gleichzeitig ist von einer unzureichenden Durchsetzung von Landnutzungs- und Umweltverordnungen auszugehen, d.h. Entwaldung findet oft unkontrolliert statt (ETTER et al. 2006). Bis 2003 wurde

die Entwaldung vor allem durch die bedingt von politischen Unruhen hohen Migrationsraten katalysiert, indem Einwanderer illegal Grundstücke besetzten. Derartige Zustände verminderten besonders langfristige Investitionen (KERNAN et al., 2006)
Neben dem Bedarf an Flächen zur Ausweitung der Infrastruktur wurden Waldflächen zu Bananen- und Palmölplantagen transformiert, deren Fläche sich zwischen 2003 bis 2010 fast verdoppelt hat. Die bisherige Wirtschaftsweise ist geprägt von kurzfristigen Investitionen in Monokulturen, intensiver Pestizid- und Düngernutzung und kurzer Rotation von Anbauflächen. Langfristig führt dies zur Unfruchtbarkeit und Erosion von Böden und verminderten Erträgen Die vielfache Anwendung der Slash & Burn-Technik erhöht die Bodendegradation und reduziert das Nutzungspotential auf Weidebetrieb. Folgen sind eine stetige Rodung für neue Anbauflächen im Zuge einer stetigen Savannisierung (RESTREPO & SYVITSKI, 2006; CAVALIER et al., 1998), da SDTF auf den degradierten Flächen nur ein geringes Regenerationspotential aufweisen. Die separat durch MxL-Klassifikation identifizierten Slash & Burn-Flächen sind auf die gesamte Region verteilt und folgen in ihrer räumlichen Verteilung den Entwaldungsmustern.
Bis auf kleine Flächen nahe Flussläufen oder auf steilen Hängen befinden sich stabile Wälder hauptsächlich in Naturschutzgebieten. Die Analyse von Waldentwicklung und Schutzgebieten (Tab. 5) hat zeigt, dass deren Effektivität von Schutzmaßnahmen im Zeitraum 2003-2010 zugenommen hat. Die noch vorkommenden Waldverluste in Schutzgebieten (siehe Peaks in Abb. 4A) resultieren u.a. aus der Erschließung von Coca-Anbauflächen durch Guerilla, deren Präsenz nicht nur zur Entwaldung selbst, sondern auch zur Behinderung von Forstinventuren, Monitoring und Naturschutz beiträgt (DAVALOS & BEJERANO, 2008).
Waldflächenzuwächse weisen keine deutlichen Abhängigkeiten bezüglich der sozio-ökonomischen Variablen auf und erfolgt wie auch ETTER et al. (2006) bestätigt zufällig statt. Damit wird deutlich, dass in den ländlichen Gebieten der Provinz Magdalena bisher ungenügend Anstrengungen zur Aufforstung unternommen werden.
Zusammenfassend verdeutlicht die untersuchte Beziehung zwischen Waldentwicklung und sozio-ökonomischen Faktoren den Bedarf an nachhaltigem Forstmanagement und Schutzgebieten in ländlichen Gebieten, die leicht zugänglich und von direktem Konsum geprägt sind. Insbesondere Anreize zur Nutzung von Lesser-Known Species (LKS) könnten ein nachhaltiges Management fördern. Gleichzeitig müssen dafür nötige Rahmenbedingungen, wie Sicherheit, Bildung und Monitoring-Konzepte mit einbezogen werden.

5 Danksagungen

Dank gilt dem DAAD für finanzielle Unterstützung in Kolumbien. Weiteren Dank für technische Unterstützung an A. Figueroa (Universidad Distrital, Bogotá), G. Becker, A. Paredes, C.P. Gross, B. Porkorny, (Albert-Ludwigs-Universität, Freiburg), TREES-Projekt (JRC, Ispra)

6 Literaturverzeichnis

ACHARD, F., EVA H., STIBIG H.J., MAYAUX P., GALLEGO, J., GICHARDS, T. & MALINGREAU, J. P., 2002: Determination of the world's humid tropical deforestation rates during the 1990's –

methodology and results of the TREES-II research programme. Office for Of-ficial Publications of the European Communities, Luxembourg. 144 pp.

OLSON, D. M. & DINERSTEIN, E. 2002: Ecosystemas Terrestres de Colombia. In: The global priority ecoregions for global conservation. USDA Forest Service and U.S. Geological Survey: The Nature Conservancy **89**:125-126. -, ,

CAVALIER, J., AIDE T.M., SANTOS, C., EUSSE, A.M. & DUPUY, J.M. 1998: The savannization of moist forest in the Sierra Nevada de Santa Marta, Colombia. Journal of Biogeography **25**:901-912.

FPSNSM, FUNDACIÓN PRÓ-SIERRA NEVADA DE SANTA MARTA. 1991: Historia y geografia, Sierra Nevada de Santa Marta. Bogotá, CORPES; Fondo FEN; Bavaria; PROPAL; Banco de Occidente; Jessie Smith Noyes Foundation; Unión Mundial para la Conservación de la Naturaleza-UICN.

DANE, DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADISTICA. 2012: Estimaciones de población 1985-2005 y proyecciones de población 2005-2020. <http://www.dane.gov.co/files/investigacione s/poblacion/proyepobla06_20/Edades_Simples_1985-2020.xls> last visited 02. Jul. 2012.

DAVALOS, L.M. & BEJARANO, A.C. 2008: Conservation in conflict: Illegal drugs versus habi-tat in the Americas. State of the wild 2008-2009: A global portrait of wildlife, wildlands and oceans (2):218-225.

RESTREPO, J.D. & SYVITSKI, J.P.M. 2006: Assessing the effect of natural controls and land use change on sediment yield in a major Andean river: The Magdalena drainage basin, Colombia. Ambio **35**(2):65-74.

PAREDES, A. 2012: Assessing utilisation of lesser-known tree species in secondary semi-deciduous tropical forest in Colombia: A contribution to ecosystem rehabilitation and sustainable management. Promotionsschrift, Albert-Ludwigs-Universität, Freiburg, Deutschland.

SONG, C. WOODCOCK, C.E., SETO, K.C., LENNEY, M. P. & Macomber, S.A. 2001: Classifica-tion and change detection using Landsat TM data: When and how to correct atmos-pheric effects? Remote Sensing of Environment 75:230-244.

MARTINUZZI, S., GOULD, W.A., RAMOS GONZALEZ, O.M., MARTINEZ ROBLES, A., MALDONADO, P.C., PÉREZ-BUITRAGO, N. & FUMERO CABAN, J.J. 2007: Mapping tropical dry forest habitats integrating Landsat NDVI, Ikonos imagery, and topographic information in the Carib-bean island of Mona. Rev. Biol. Trop. **56**(2):625-639.

BARREDA-BAUTISTA, B., LÓPEZ-CALOCA, A.A., COUTURIER, S. & SILVÁN-CÁRDENAS, J.L. 2011: Tropical dry forest in the global picture: The challenge of remote sensing-based change detection in tropical dry environments. In: Planet Earth – Global warming challenges and opportunities for policy and practice: 232-256.

ETTER, A., MCALPINE, C., PULLAR, D. & POSSINGHAM, H. 2006: Modelling the conversion of Colombian lowland ecosystems since 1940: Drivers, patterns and rates. Journal of En-vironmental Management 79:74-87.

KERNAN, B.S., MONJE, C. C. A. & VON HILDEBRAND, M.P. 2006: Report on tropical forests and biological diversity. USAID, Colombia country strategy management. FY 2006-2010, Bogota D.C., Colombia. 98 pp.

BRUCE, C.M. & HILBERT, D.W. 2006: Pre-processing methodology for application to Land-sat TM/ETM+ imagery of the wet tropics. Cooperative Research Centre for Tropical Rainforest Ecology and Management. Rainforest CRC, Caims. 44 pp.

About the impact of CCSDS image data compression on Image Quality

KRISTIAN MANTHEY[1]

Zusammenfassung: Zur Bewertung der Güte eines verlustbehafteten Bildkompressionsverfahrens werden häufig einfache Qualitätsmaße wie PSNR oder der besser an das menschliche Auge angepasste MSSIM-Index verwendet. Diese Qualitätsmaße quantifizieren den durch eine verlustbehaftete Kompression entstandenen Qualitätsverlust. Für die Fernerkundung eignen sich diese Verfahren jedoch nicht, da Fernerkundungsbilder zunehmend automatisch und mittlerweile weniger häufig durch den Menschen ausgewertet werden. Zur Bewertung der Bildqualität eines optischen Systems in der Fernerkundung werden u.a. die MTF und das SNR des Systems untersucht.
In dieser Arbeit wird der Einfluss aktueller verlustbehafteter Kompressionsverfahren mit Schwerpunkt auf dem Verfahren CCSDS-IDC auf die Bildqualität in untersucht. Werden die Bilddaten automatisch verarbeitet, so ist die Bildqualität letztlich hochgradig abhängig von den eingesetzten Algorithmen. Es wird gezeigt, dass eine verlustbehaftete Kompression selbst bei hohen Anforderungen an die Bildqualität möglich ist und diese Algorithmen bei geeigneter Wahl der Kompressionsparameter nicht negativ beeinflusst werden.

1 Introduction

In spaceborne applications, there are challenging requirements on optical systems. In the last years, the spatial as well as the spectral resolution of the image data has increased, resulting in a tremendous increase in data rate. Besides this, there are requirements to image quality and constraints, resulting from the environment in which the system is to be used. An optical system for spaceborne application has to be highly reliable, low in power consumption as well as low in weight. The system also has to be radiation tolerant and able to operate in vacuum and in a high temperature range.
One major problem of satellite imagery is onboard storage and transmission of image data. This is caused by limited memory resources at the spacecraft as well as by limited download capacity and time to the ground station. So, it is necessary to compress image data. Depending on the requirements of the mission, lossless or lossy compression schemes can be used.
When using image compression, new problems arise. On a satellite we usually have very limited resources like space, time and power. A typical lossless image compression ratio is around 1:2 which is most often not sufficient. The loss of information, which comes due to the quantization stage in lossy compression, can be critical for some application, which results in a common assumption that lossy image compression is not appropriate for this particular application. This is almost always the case in scientific mission. On the other side, lossy compression suppresses noise in the image what can cause, that other algorithms run better CHOONG et al. [2006].

1) DLR, Institut für Robotik und Mechatronik, Rutherfordstraße 2, 12489 Berlin, Germany

It is a fact, that lossy image compression introduces errors between the original image and the decompressed image. Depending on the compression ratio this might lead to visible distortions and artifacts (e.g. JPEG block artifacts) which results in degradation of the visual quality of the image. In applications where visualization for human beings is the main intention of producing images, the human visible system (HVS) has to be taken into account.
Correctly used, the term image quality describes the image degradation compared to an ideal image. On the way from the light source to the image, several components like the environment, the lenses or the electronics of the imaging system can introduce some amount of distortion. The term image quality is sometimes used wrongly. In the context of image compression, the performance of an image compression algorithm is evaluated with so called rate-distortion curves. Traditional distortion metrics used in the research field of image compression are mean square error (MSE) and peak signal-to-noise ratio (PSNR). As shown in PONOMARENKO et al. [2010], these metrics have a weakness in predicting visual quality. In WANG et al. [2004] the structural similarity (SSIM) was presented, which uses luminance, contrast and structure comparisons to get a better visual quality prediction. In this context it should be noted, that possible noise of the original image most often is not taken into account. When using lossy image compression, knowledge about noise can be exploited to improve compression ratio.
In remote sensing, the HVS is not a good reference to determine image quality degradation because the main attention of producing images is often to process these automatically by algorithms. Furthermore, from the engineering point of view, determining the final image quality of the system or designing a system for a given image quality is a desired goal. A very useful tool for describing the image quality is the modulation transfer function (MTF), which relates input and output of a linear (imaging) system. Using the MTF has the advantage, that the overall system MTF can be determined as the product of the individual component MTFs. The general image quality equation (GIQE) is an empirical formula for predicting the image quality in terms of the National Imagery Interpretability Rating Scale (NIIRS) IRVINE [1997]; LEACHTENAUER et al. [1997]; THURMAN & FIENUP [2008].
The purpose of the paper is to get a better understanding of the impact of lossy image compressing using the CCSDS-IDC algorithm. The author currently works on a fast hardware implementation on a FPGA which is capable to compress data rates of 1 GByte/s and more. In future project, it is planned to integrate an image compression module directly on the sensor platform. So, it is absolutely necessary to predict the achievable image quality for given output data rates. CCSDS-IDC was chosen because of several aspects presented in section 4.1.
The paper is organized as follows. Section 2 gives a short overview of used visual quality metrics. Section 3 presents some basics of image quality and MTF measurement. In section 4, the CCSDS-IDC algorithm is presented in a very short term and the datasets are presented. In section 5, the results of the investigations are shown. Section 6 gives a short discussion on the results and outlook of this work.

2 Visual Quality Assessment

The quality of an image compression algorithm is described with rate-distortion curves. These curves describe the dependency between bitrate (in bits per pixel, bpp) and the distortion of the

reconstructed image in comparison with the original image. Most often, the mean square error (MSE) or peak signal-to-noise ratio (PSNR) are used. The definitions for these metrics are

$$MSE(X,Y) = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}[X(i,j) - Y(i,j)]^2$$

and

$$PSNR(X,Y) = 10 \cdot log_{10}\left(\frac{L^2}{MSE(X,Y)}\right) dB,$$

where M and N denote the width and the height of the images $X(i,j)$ and $Y(i,j)$ and L is the dynamic range of the pixel values (for n bit unsigned integer it is $2^n - 1$). Although they are almost always used, MSE and PSNR produce not an adequate description of the visual quality. The problem is that different distortions can produce the same MSE or PSNR although they have very different visual quality WANG et al. [2004]. Structural similarity (SSIM) is a metric, which involves structural, luminance and contrast information in the images. There are a lot of other visual quality assessment metrics which were compared against a mean opinion score (MOS) VIF, VSNR or PSNR-HVS, just to name a few PONOMARENKO et al. [2009]. All these metrics are full-reference metrics as they compare in undistorted with a distorted version of the image.

3 Image Quality Assessment

Image quality is influenced by properties of the scene and by the image acquisition system. Some of the image quality factors are dynamic range, signal-to-noise ratio (SNR) and the modulation transfer function (MTF). Dynamic range is the information of the smallest and the highest possible value of a signal (or image). SNR is a measure that gives to relationship of the signal, usually containing the information, to the noise. The MTF describes the loss of contrast of the edges of an object compared to the contrast in the image.

Methods for the determination of the MTF are shown extensive in BURNS [2009]. Image compression algorithms can be understood as a component of the overall imaging system HADAR et al. [2001]. However, they are typically neither linear, i.e. not all the output frequencies correspond with the same input frequencies due to the irreversible quantization, nor isoplanar, i.e. they are space-variant because they use some sort of blocking of image pixel or coefficients. In HADAR et al. [2001], lossy compression was assumed to be linear to solve that problem. Nevertheless, the MTF gained from compressed images must be carefully assessed, because image compression can introduce edges and artifacts.

The MTF can be determined by several methods presented in BLANC & WALD [2009]. The method used in this paper is based on edge targets. The edge spread function (ESF) is the systems response to a high dark-bright contrast edge. With the slanted edge algorithm used in KOHM [2004] it is possible to create a super resolution version of the ESF. The derivative of the ESF is the line spread function (LSF). The LSF is then normalized to 1. The Fourier transform of the LSF gives the MTF.

4 Algorithms and Datasets

4.1 Used Compression Algorithms

To compress images, the standards JPEG, JPEG2000, SPIHT PEARLMAN et al. [2004] and CCSDS-IDC CCSDS [2007] were used. Some of these compression algorithms have recently been used in space applications YU et al. [2009]. There are a lot of other compression algorithms in the academic research are: e.g. EZW SHAPIRO [1993] or SPECK, just to name two. The standard CCSDS 122.0-B-1 is a recommendation of the Consultative Committee for Space Data Systems (CCSDS). It is a wavelet-based image compression algorithm and often just called CCSDS-IDC (image data compression, IDC). This standard can operate in lossless as well as in lossy mode. As denoted in CCSDS [2007], the standard differs from JPEG2000 in several aspects:

- it specifically targets high-rate instruments used on board of spacecraft;
- compression performance has been traded off against complexity, with particular emphasis on spacecraft applications;
- the lower complexity of the recommendation supports fast and low-power hardware implementation;
- it has a limited set of options, supporting its successful application without in-depth algorithm knowledge.

The compressor typically consists of two parts: a discrete wavelet transform (DWT) module and the bit-plane encoder, which is an entropy coding stage. To have a well understood code basis, the encoder and the decoder was implemented in C++. For all tests, the encoder was set to frame mode, which means, that the whole images were compressed in a single segment. The wavelet was set to the lossy bi-orthogonal CDF 9/7 wavelet. Quality control was achieved by setting segment size to the desired file size.

4.2 Datasets

To investigate in the impact of image compression on image noise, Gaussian noise images with certain variance and mean were generated. In the case of compression on noisy images, Gaussian noise was added to images with a certain variance, but zero mean. For this paper, noise was added to the recently used image “Lena” and the image “Marstest” taken from CCSDS [2007]. To investigate in the impact of image compression on feature matching, test images “boat” and “graf” of the university Leuven LEUVEN [2007] were used. The datasets contain several images taken from different perspectives and transformation matrices to project the images on each other. For the investigation in the impact of image compression on stereo matching, images from the Middlebury Stereo Vision Page SCHARSTEIN & SZELISKI [2002] were used. For stereo-matching, the StereoSGBM matching implementation in OpenCV was used. To determine the MTF, tools and test charts from BURNS [2009] and the recently used image “boat” were used.

5 Results

5.1 Impact of compression on noise

A first investigation on the impact of compression was done on Gaussian noise images with $\sigma^2 = \{10, 100\}$ and $\mu = 127$. A similar research was done in PONOMARENKO et al. [2010] were the authors corrupted the test image Lena by Gaussian noise (see also subsection 5.2). Depending upon compression algorithm, the compression ratio can be varied by compression rate (CR) or by a quality factor. In the case of JPEG, the compression rate is controlled by a quality factor in the range $[0, 100]$. Modern algorithms like JPEG2000, SPIHT and CCSDS-IDC support to select a desired bitrate. Figure 1 shows the results. As expected, at low bitrates, all compression algorithms produce images with low noise variance due to the fact, that lossy image compression has low-pass characteristics. One can see, that these rate-noise-curves have an overshoot between 2 and 3 bpp, what can be explained by signal over- and undershoot at high frequencies and in turn, this increases the dynamic range of the image. It is also obvious, that the SPIHT and the CCSDS-IDC algorithm have a similar curve except they have a small spatial shift. This can be explained by the fact, that both algorithms use the bi-orthogonal CDF 9/7 wavelet and a similar hierarchical coding scheme.

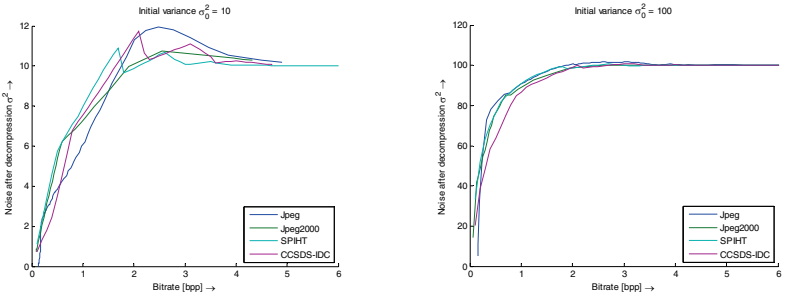


Figure 1: The noise variance before compression σ_0^2 and after decompression σ^2 for $\sigma_0^2 = 10$ (left) and $\sigma_0^2 = 100$ (right).

5.2 Impact of compression on noisy images

In order to investigate the impact of lossy image compression on noisy images, the images Lena and Marstest (I_{nf}, noise free) were corrupted by Gaussian noise with $\sigma^2 = \{50, 200\}$ and $\mu = 0$ as it was done PONOMARENKO et al. [2010]. The resulting image is I_n which contains additive Gaussian noise. The image I_n is now compressed and then decompressed to image $I_{\tilde{n}}$.
Figure 2 shows the results. One can see that for noisy images with known noise characteristics there is some optimal bitrate which in this case is for $\sigma^2 = 50$ at around 0.35 bpp and for

$\sigma^2 = 200$ at around 0.2 bpp. For the image Marstest, results look some different. For $\sigma^2 = 50$

there is clear maximum of the PSNR curve, but there is an optimal bitrate at around 0.75 bpp whereas the optimal bitrate for $\sigma^2 = 200$ is at around 0.4 bpp. These results show that there is a dependency between image noise and optimal compression bitrate. The higher the image noise is, the lower is the optimal bitrate for the tested images. Unlike in PONOMARENKO et al. [2010], the PSNR value of the rate-distortion curve of the noise-free image is always higher or equal to the corresponding value in the curve of the noisy image. Furthermore the results show that the dependency between image noise and optimal bitrate also depends on the image scene. In the case of image Lena there is a clear maximum of the PSNR curves whereas in the image Marstest there is no such point.

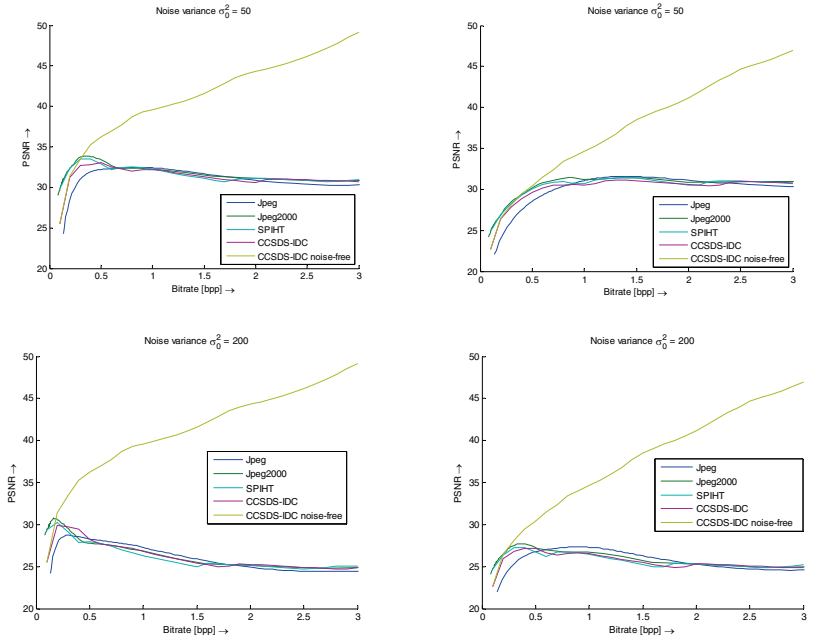


Figure 2: Dependency between bitrate and $PSNR(I_{nf}, I_d)$ for the image Lena (left) and Marstest (right).

5.3 Impact of compression on feature matching

Features are often used to match a current image against another image (or images) to find corresponding image points. With these correspondences it might be possible to detect and track objects, to reconstruct the stereo information of two cameras or to orientate images to each other. It is conceivable that image features are matched against features derived from compressed images. In order to investigate the impact of image compression on feature matching, scenes from the database Leuven where used. The database contains eight scenes each including 6 images and 5 transformation matrices from the first to every other image of the scene.
For each test scene, sift LOWE [2004] features ware calculated the first image of the dataset.

Then, images 2 – 6 were compressed by CCSDS-IDC at different compression rates. After

decompression, again sift features were calculated and matched against the features from the first image. With the transform matrix it is possible to decide whether a matching is good or bad: For each matching, the point in the first image is transformed to the position in the second image. Then, the distance from this position to the position of the matching feature was calculated. If the distance is higher than a threshold, the matching was set to bad. A threshold of 10 pixels was chosen for these tests.
Figure 3 shows the results. One can see that on both plots, there are very different percentages of wrong matches for the different images. For the scene “boat” on the right side, the rotation and zoom of the scene increase from image 2 to 6. This leads to a higher percentage of wrong matches. The upper curves show a high percentage of wrong matches at all bitrates, which is caused by high perspective distortions between the images. According to the compression rate, one can see that there is a rise in wrong matches below 1 bpp for both scenes.

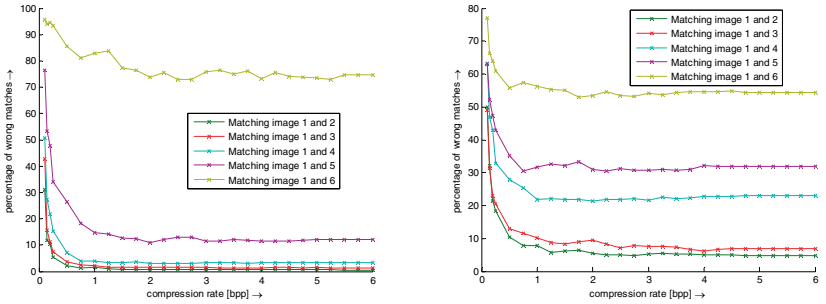


Figure 3: The percentage of erroneous sift feature matches in dependence of the compression rate for the scene wall (left) and for the scene boat (right).

5.4 Impact of compression on stereo-matching

Beside feature matching, stereo-matching algorithms are used more and more in the interpretation of remote sensing images. The datasets from the Middlebury Stereo Vision Page SCHARSTEIN & SZELISKI [2002] and the StereoSGBM matching implementation in OpenCV were used. The database contains four image pairs (cones, teddy, tsukuba and venus), the corresponding ground truth and three masks which are needed for stereo-matching evaluation. For these tests, the mask "all" was chosen containing all pixels including half occluded pixels. Because the OpenCV implementation of StereoSGBM cannot calculate disparities for all pixels, some more pixels were excluded from error calculation. To measure the matching error, the percentage proportion of faulty pixels in the disparity map was determined with the mask. See SCHARSTEIN & SZELISKI [2002] for more details on this calculation.

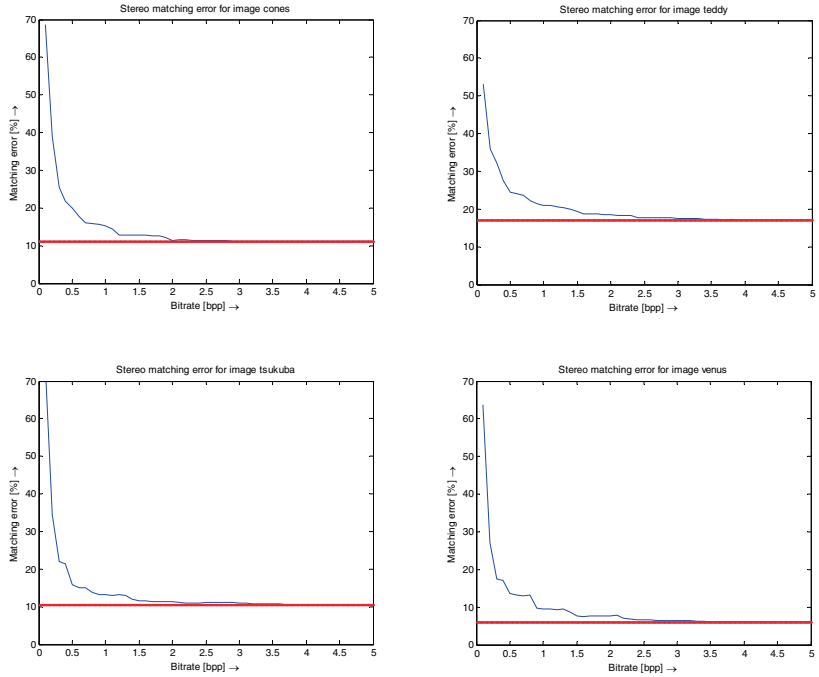


Figure 4: The matching error in dependence of the compression bitrate chosen for both images. The horizontal red lines show the matching error without compression. The evaluation database contains four images: cones (top left), teddy (top right), tsukuba (bottom left) and venus (bottom right).

Figure 4 shows the results. The blue curves describe the percentage of erroneous disparity values. The red lines show the error value when no compression was used. One can see that at 4 bpp there is almost no error for all images. The matching error rises at around 2-2.5 bpp. At 1 bpp, the matching error is around 5 % higher than without compression. At 0.5 bpp, it is around as twice as high as without compression. At around 0.25 bpp, stereo-matching the number of bad matches increases dramatically.

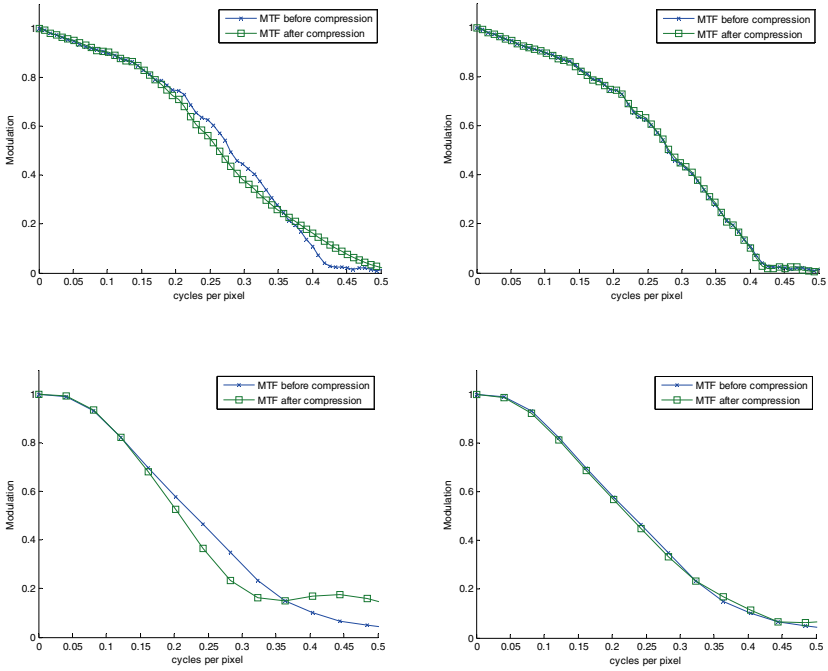


Figure 5: The MTF before compression and after decompression on the same region of interest. The top curves show the MTF curves for the testchart at 0.5 bpp (left) and 1 bpp (right). The bottom curves show the MTF curves for the boat image at 1 bpp (left) and 2 bpp (right).

5.5 Impact of compression on image MTF

To determine the MTF, tools and test charts from BURNS [2009] and the image "boat" were used. The tool sfrmat3 uses the slanted edge method KOHM [2004]. Two images were chosen for evaluation. Regions of interest for MTF determination were selected for both images. In the image "ISO DSC 300 mono", this area in $x = [515,731]$ and $y = [337,454]$ and for the image "boat", the area is in $x = [337,378]$ and $y = [255,279]$. The compression ratios were 0.5 bpp and 1 bpp for the testchart, and 1 bpp and 2 bpp for the boat image. Figure 5 shows the results.

One can see that for 1 bpp, the MTF of the Testchart, and for 2 bpp, the MTF of the boat are nearly unaffected. When increasing the compression ratios, both MTF curves degrade. But because of the non-linearity of the image compression, results must be interpreted with caution. The dependence of MTF and compression rate was also investigated with other bitrates and the MTF curves, which are not shown in this paper, show, that the MTF is stable at higher bitrates than 2 bpp for the boat image and 1 bpp for the testchart.

6 Discussion and Outlook

In this paper, the impact of lossy image compression with CCSDS-IDC on image quality was investigated on selected examples. The results have shown that a compression ratio of 1:4 (2 bpp bitrate for 8 bit images) has almost no impact on the quality of the results. For the author, this study is not yet complete and there are many more relationships to be examined. Furthermore, the already obtained results will be examined in the future and must be proofed with more images. The complexity of the image scene is another topic which has to be included in this investigation. The influence of the selected wavelet on the application performance and MTF measurements is also an interesting and part of the author's current research.

7 Acknowledgements

Kristian Manthey would like to acknowledge financial support of the Helmholtz Research School on Security Technologies.

8 Literature

BLANC, P. & WALD, L., 2009: A review of earth-viewing methods for in-flight assessment of modulation transfer function and noise of optical spaceborne sensors. Working Paper.

BURNS, P.D., 2009: sfrmat3: SFR analysis for digital cameras and scanners.

CCSDS, 2007: Image data compression - Green book.

CHOONG, M. K.; LOGESWARAN, R. & BISTER, M., 2006: Improving diagnostic quality of MR images through controlled lossy compression using SPIHT. J. of Med. Sys., 30, S.139-143.

HADAR, O.; STERN, A.; HUBER, M. & HUBER, R., 2001: Modulation transfer function as a quality measure for compressed images transmitted over a lossy packet network. Optical Engineering, 40(10), S. 2134-2142.

IRVINE, J.M., 1997: National imagery interpretability rating scales (NIIRS): overview and methodology. Proc. SPIE 3128, Airborne Reconnaissance XXI, 93: S. 93-103.

KOHM, K., 2004: Modulation transfer function measurement method and results for the orbview-3 high resolution imaging satellite. ISPRS Archives - Volume XXXV Part B1.

LEACHTENAUER, J. C.; MALILA, W.; IRVINE, J.; COLBURN, L. & SALVAGGIO, N., 1997: General image-quality equation: GIQE. Applied Optics, 36: S. 8322-8328.

LOWE, D. G., 2004: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 60: S. 91-110.

KATHOLIEKE UNIVERSITEIT LEUVEN, 2007: Affine covariant features.

http://www.robots.ox.ac.uk/vgg/research/affine

PEARLMAN, W. A.; ISLAM, A.; NAGARAJ, N. & SAID, A., 2004: Efficient, low-complexity image coding with a set-partitioning embedded block coder. IEEE Transactions on Circuits and Systems for Video Technology, 14(11): S. 1219-1235.

PONOMARENKO, N.; LUKIN, V.; ZELENSKY, A.; EGIAZARIAN, K.; ASTOLA, J.; CARLI, M. & BATTISTI, F., 2009: Tid2008 - A database for evaluation of fullreference visual quality assessment metrics. Advances of Modern Radioelectronics, Vol. 10: S. 30–45.

PONOMARENKO, N.; KRIVENKO, S.; LUKIN, V.; EGIAZARIAN, K. & ASTOLA, J., 2010: Lossy compression of noisy images based on visual quality: A comprehensive study. EURASIP Journal on Advances in Signal Processing, 2010(1): S. 976436.

SCHARSTEIN, D. & SZELISKI, R., 2002: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision, 47(1-3): S. 7-42.

SHAPIRO, J. M., 1993: Embedded image coding using zerotrees of wavelet coefficients. Signal Processing, IEEE Transactions on, 41(12): S. 3445-3462.

THURMAN, S. T. & FIENUP, J. R., 2008: Analysis of the general image quality equation. In Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, volume 6978.

WANG, Z.; BOVIK, A. C.; SHEIKH, H.R & SIMONCELLI, E. P., 2004: Image quality assessment: from error visibility to structural similarity. IEEE Trans. on Image Proc., 13(4): S.600-612.

YU, G.; VLADIMIROVA, T. & SWEETING, M. N., 2009: Image compression systems on board satellites. Acta Astronautica 64, S. 988-1005.

Objektbasierte spektrale Analyse und Klassifikation von ölkontaminierter Landoberfläche in Westsibirien mit hochaufgelösten RapidEye Daten

RENÉ MICHAELIS[1] & SÖREN HESE[2]

Zusammenfassung: Erdölreserven sind auf Grund des starken Abhängigkeitsverhältnisses der Weltwirtschaft von den Erdölprodukten, von geopolitischer und wirtschaftlicher Bedeutung. Russland hat in Westsibirien sehr große Erdölvorkommen und zählt zu den größten Exporteuren der Welt. Die massive Erdölförderung führt zu großflächigen Umweltverschmutzungen in der Nähe der Erdölindustrie. Mit Hilfe von Fernerkundungsdaten können die kontaminierten Gebiete schnell und effektiv kartiert werden. Die verwendeten RapidEye Aufnahmen eignen sich auf Grund der hohen temporalen und räumlichen Auflösung hervorragend für Monitoringanwendungen. In dieser Arbeit wird ein objektbasiertes Klassifikationsschema verwendet, welches an einem kleinen Testgebiet parametrisiert und dann auf große Flächen übertragen wird. Ziel ist es, ein übertragbares Klassifikationsschema zu entwickeln welches ohne großen Aufwand auch auf andere Aufnahmen übertragen werden kann.

1 Einleitung

Erdölprodukte haben auf dem Weltmarkt eine einzigartige Bedeutung für die wachsende Weltwirtschaft. Das Wirtschaftswachstum steht in einem starken Abhängigkeitsverhältnis zur Verfügbarkeit von Energie. Die steigende Nachfrage und die begrenzten globalen Ressourcen machen die Erdölförderung und den Export von Erdöl und Erdgas zu einem wichtigen Wirtschaftsfaktor. Russland hat in Westsibirien große Erdöl- und Erdgasvorkommen und zählt zu den größten Öl- und Gasexporteuren der Welt. Die seit Jahrzenten ausgeweitete Erdölförderung birgt aber auch die Gefahr der massiven Umweltverschmutzung durch Ölkontamination in den Gebieten der Erdölförderung und der Transportwege. Die natürlichen Verwitterungsprozesse für das Rohöl werden durch das vorherrschende subarktische Klima verlangsamt oder gestoppt, wodurch die Erdölkontamination das Ökosystem über mehrere Jahrzehnte belastet (FINGAS 2001:179). Da die klimatischen Bedingungen Pipelinebrüche begünstigen und die Verwitterung des Rohöls verlangsamen, sind die Folgen der Ölkontamination in Sibirien besonders problematisch. Durch die lange Verweildauer im Ökosystem breitet sich das Erdöl auch im Boden aus, bis es auf Sperrschichten trifft. Im Falle einer Grundwassersperrschicht können die Folgen der Erdölkontamination weitreichender sein als der oberflächliche *oil spill*, da ganze Grundwasser Reservoirs kontaminiert werden (WIDGET 1997:o.S.). Erdöl ist eine große chemische Belastung für die Umwelt weshalb solche *oil spills* so schnell wie möglich entdeckt und auch beseitigt werden müssen. Durch räumlich und zeitlich immer hochauflösendere Fernerkundungssensoren wie das RapidEye System ist es möglich Erdölförderregionen zu Überwachen und Veränderungen im Ökosystem schnell zu erkennen.

1) René Michaelis, FSU Jena, Kastanienstr. 14, 07747 Jena; rene.michaelis@uni-jena.de
2) Sören Hese, FSU Jena, Grietgasse 6, 07745 Jena; soeren.hese@uni-jena.de

2 Grundlagen

Der Fokus der *oil spill* Kartierung liegt hauptsächlich in maritimen und Küstenbereichen, die durch Tankerunglücke oder Plattformunfälle wie der Deepwater Horizon *oil spill* (2010) betroffen sind. Auf dem Meer konnte mit Hilfe von Synthetik Aperatur Radar (SAR) Daten verschiedener Systeme (TerraSAR-X, EnviSat ASAR, ERS RADARSAT) Methoden entwickelt werden *oil spills* in ihrer Ausdehnung und ihrer Bewegung zu kartieren (MARTINIS et al. 2012, BREKKE & SOLBERG 2005, GIRARD-ARDHUIN et. al. 2005, FISCELLA et al. 2000). Diese Methoden basieren auf der höheren Viskosität von Erdöl gegenüber Wasser, wodurch kleine Wellen an der Wasseroberfläche gedämpft werden und die Oberflächenrauhigkeit reduziert wird (GIRARD-ARDHUIN et al. 2003:4). Durch die glattere Oberfläche der Ölschicht gibt es weniger Radarrückstreuung als über den Wasseroberflächen, weshalb *oil spills* im Radarbild tief schwarz erscheinen (BREKKE & SOLBERG 2005:4).

Auf dem Land lassen sich Erdölkontaminationen nur sehr schwer mit Fernerkundungssensoren kartieren, da das Öl nur selten flächendeckend an der Landoberfläche bleibt. Das Erdöl verteilt sich entsprechend der Topographie im Gelände, wobei die leichten Bestandteile sehr schnell durch Evaporation und Verdunstung in die Atmosphäre übergehen (BRADOCK et al. 2003:11).

Die verbleibenden Rohölreste bilden mit den vorhandenen Landbedeckungen (gesunder Vegetation, offenen Böden und toter oder kranke Vegetation) spektrale Mischsignaturen, wodurch eine saubere Klassifikation erschwert wird (HESE & SCHMULLIUS 2009:130). HÖRIG et al. (2001) untersuchen mit Hilfe von hyperspektralen HyMap-Daten die spektralen Eigenschaften von Öl und zeigen anhand von Erdölprodukten (Plastik) und ölkontaminiertem Sand, dass es spektrale Absorbtionsmerkmale von Kohlenwasserstoffen gibt (z.B. bei 1730nm), die sich auch auf die kontaminierten Sandflächen übertragen. Des Weiteren, zeigt sich, dass die Versuchsflächen mit ölkontaminiertem Sand über die gesamte spektrale Signatur geringere Rückstreuwerte erzeugen als die Versuchsflächen mit reinem Sand. LI et al. (2005) verwenden ebenfalls Hyperspektraldaten (AVIRIS) und konzentrieren sich auf Vegetationsstress der durch *oil spills* verursacht wird. Im Fokus steht dabei der *red edge inflection point* (REIP) der im Gegensatz zu anderen Vegetationindizes weniger durch Störungen aus Atmosphären oder Bodensignal beeinflusst wird und deshalb sensitiver für die frühzeitige Stressdetektion ist (LI et al. 2005:3). Die Grundlage dafür ist der Zusammenhang von Chlorophyl- und Stickstoffgehalt und dem Anstieg zwischen rotem und NIR Spektrum, wobei mit zunehmender Vitalität einer Pflanze der Anstieg im Spektrum in Richtung NIR verschoben wird (PINAR & CURRAN 1996).

Normalerweise sind sehr enge spektrale Messungen nötig, um Aussagen über den Anstieg zwischen rotem und NIR Kanal, oder dem REIP zu treffen. Das RapidEye Satellitensystem besitzt einen Kanal direkt im Bereich der *red edge* und ermöglicht so, ohne hyperspektrale Auflösung, Pflanzenstress über die Informationen des *red edge* Kanals schneller zu erkennen als herkömmliche Vegetationsindices (EITEL et al. 2007:3645). Im Gegensatz dazu zeigen HESE & SCHMULLIUS (2009) unter Verwendung eines objektbasierten Ansatzes und sehr hochauflösenden Quickbird-Daten, dass ölkontaminierte Gebiete auch ohne hohe spektrale Auflösung kartiert werden können. Dabei kommen Kontextinformationen wie Distanzbeziehungen zwischen *oil spills* und industrieller Infrastruktur zum Einsatz.

3 Datengrundlage und Untersuchungsgebiet

Das Testgebiet liegt in der Khanty Mansi Region in Westsibirien im subarktischen Klima. Im Süden befindet sich Taiga die nach Norden von der Tundrenvegetation abgelöst wird. Die dominanten Bäume in der Taiga sind Fichten und Kiefern (WEIN 1999:36), wobei der Bewaldungsgrad von 40% im Süden (beim Ob) nach Norden immer weiter abnimmt. Der geringere Baumbestand in diesem Gebiet ist durch die Waldsumpfzone begründet, in der Seen, Sümpfe und Moore bis zu 50% der Fläche bedecken (WEIN 1999:30). Die Ölförderanlagen und die Sandstraße bilden eine sehr charakteristische Struktur, bei der Betrachtung mit Fernerkundungssensoren (Abbildung 1). Die Klassen umfassen industrielle Gebiete, Straßen, ölkontaminierte Gebiete, Tundren und vitale Tundrenvegetation. Einige Gebiete scheinen von einer leichten Schneedecke bedeckt zu sein. Dies verkompliziert die Analyse der spektralen Signaturen der tatsächlichen Landbedeckung erheblich. Zu den typisch dunklen Wasserflächen wurden im Südwesten des Testgebietes Wasserflächen mit hoher Reflektanzen gefunden. Als Datengrundlage diente eine RapidEye Datenaufnahme aus dem Sommer 2011 (Tabelle 1) des RESA (RapidEye Science Archive) Projektes 472. Vor allem durch die hohe temporale Auflösung sowie den zusätzlichen *red edge* Kanal eignet sich das RapidEye Satellitensystem besonders gut für vegetationsbezogenes Umweltmonitoring. Die RapidEye Daten wurden im L1B Zustand von RESA (DLR) bezogen, auf UTM WGS84 umprojiziert und mittels ATCOR atmosphärenkorrigiert und in „scaled reflectance“ umgerechnet. Referenzdaten für die Lage der Ölkontaminierten Flächen im Kernuntersuchungsgebiet wurden aus dem OSCAR Projekt (HESE & SCHMULLIUS 2009) übertragen.

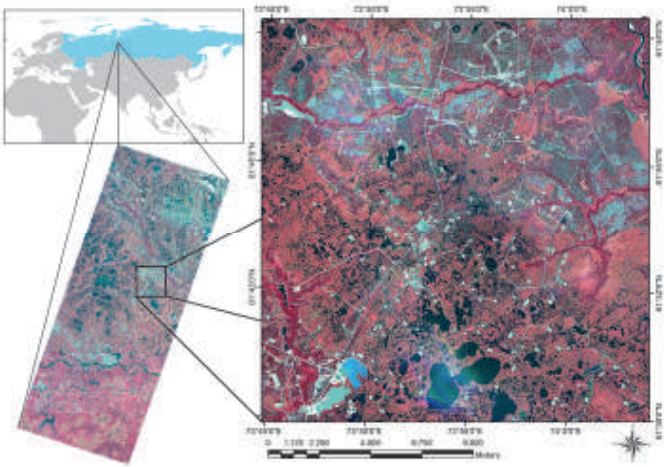


Abbildung 1: RapidEye Szene (RESA Projekt 472) mit Testgebiet in der Khanti Mansi Region (Westsibirien) in der Kanalkombination (r:NIR, g:rot, b:grün)

Tabelle 1: Datengrundlage und Spezifikation

		Kanal	Wellenlänge (nm)
Sensor	RapidEye	Blau	440 - 510
Anzahl Satelliten	5	Grün	520 - 590
Aufnahmedatum	28.06.2011	Rot	630 - 685
räumliche Auflösung	6,5m	*Red Edge*	690 - 730
Athmosphärenkorrektur	ATCOR2	NIR	760 - 850

4 Methode

Die Kartierung der ölkontaminierten Flächen wurde mit Hilfe eines objektbasierten Ansatzes durchgeführt, wobei die spektralen Klassifikationsregeln im Vorfeld aus einer spektralen Analyse abgeleitet wurden. Im Fokus der spektralen Analyse standen vor allem Vegetationsindices, da diese Aufschluss über gestresste Vegetation geben. Über die Distanzabhängigkeit der *oil spills* von industrieller Infrastruktur sollen die ölkontaminierten Gebiete kartiert werden.

4.1 Spektrale Trennbarkeitsanalyse mittels Ratios

Mit Hilfe von Trainingsobjekten wurden 100 repräsentative Objekte mit einheitlicher Größe (2x2 Pixel) für jede Klasse gesammelt, die dann statistisch untersucht wurden. In einem Vorverarbeitungsschritt wurden für alle Klassen die Ausreißer über die dreifache Standartabweichung entfernt. Abbildung 2 zeigt die Mittelwertspektren der gesuchten Klassen. In der spektralen Untersuchung wurden fünf Eingangskanäle und bekannte Vegetationsindices wie *Difference Vegetation Index* (DVI), *Enhanced Vegetation Index* (EVI) (MATSUSHITA et al. 2007:2639), *Soil Adjusted Vegetation Index* (SAVI) (HUETE 1988), *Normalized Difference Vegetation Index* (NDVI) (MATSUSHITA et al. 2007:2639) und der *Normalised Difference Red Edge Index (*NDRE) (EITEL 2011:3642) auf *red edge* und rot Basis (NDRE_rot) und *red edge* und NIR Basis (NDRE_NIR) untersucht. Ebenfalls wurden zwei spektrale Objektmerkmale aus der eCognition Software untersucht (*Brightness* und *maximum Difference*) (TRIMBLE 2012:235, 237). Des Weiteren wurde ein weiteres Merkmal entwickelt (N/diff_re), welches die Informationen aus rotem, *red edge* und NIR Kanal enthält. Die Grundlage dafür ist eine lineare Funktion (Formel 1) und die Differenz (diff_re) des gemessenen r*ed edge* Grauwertes zu dem Wert der linearen Funktion an der Stelle des red edge Kanals (Formel 2).

$$f(x) = mx + y \tag{1}$$

mit:

$$m = \frac{(NIR - red)}{(805 - 656.5)}$$

$$b = (NIR - (m * 805))$$

$$y = m * 710 + b$$

$$diff_re = red\ edge - y \quad (2)$$

Die lineare Funktion kann aus den Messwerten für rot und NIR sowie den dazugehörigen spektralen Zentrum der Kanäle berechnet werden. Durch das Verhältnis von diff_re und dem NIR Wert können die Differenzen zur linearen Funktion bezüglich der Pflanzenvitalität interpretiert werden und bei der Klassentrennung von Vegetationsklassen helfen (Formel 3).

$$NIR/diff_re = \frac{NIR}{diff_re} \quad (3)$$

Mit Hilfe von Boxplotdiagrammen (Abbildung 3) wurde die statistische Verteilung der Klassen für jedes Merkmal visualisiert. Zusätzlich wurde mit Hilfe der Jeffries Matusita Distanz (JM) (Formel 4) die Trennbarkeit bewertet. In die Jeffries Matusita Distanz geht die Bhattacharyya Distanz ein, welche bei der Betrachtung von nur einem Merkmal mit der Standardabweichung berechnet wird statt mit den Kovarianzmatrizen (Formel 5) (GUNAL & EDIZKAN (2008:3718).

$$JM = \sqrt{2(1 - e^{-B})} \quad (4)$$

$$B = \frac{1}{8}(\mu_i - \mu_j)^T \left(\frac{\Sigma}{\sigma_i^2 + \sigma_j^2}\right) + \frac{1}{2}\ln\left(\frac{\sigma_i^2 + \sigma_j^2}{2\sigma_i\sigma_j}\right) \quad (5)$$

Nach PEREIRA et al. (2004:306) liegt die Fehlerwahrscheinlichkeit einer Klassifikation bei 10% wenn die JM Distanz 1.09 beträgt. Mit einer JM Distanz von 1.24 beträgt die Fehlerwahrscheinlichkeit einer Klassifikation nur noch 5%. Das Maximum liegt bei der Wurzel aus 2 (1.414).

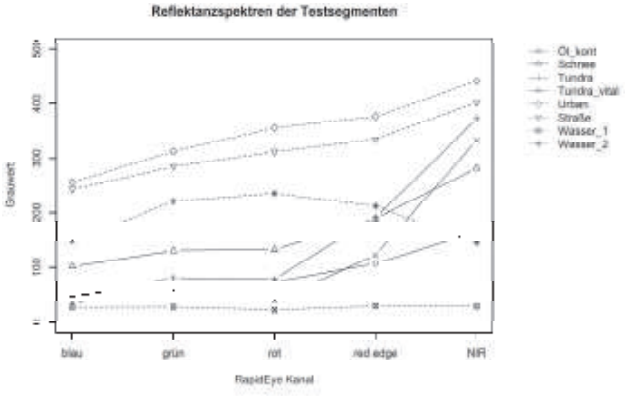


Abbildung 2: links Mittelwertspektren für alle Klassen.

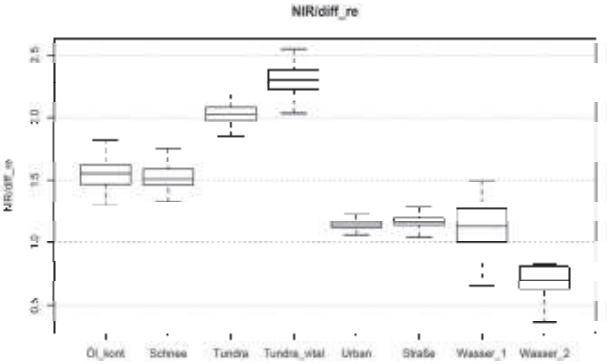


Abbildung 3: Beispiel eines Boxplotdiagramms von dem N/diff_re Merkmal für alle Klassen

Tabelle 2 zeigt die höchsten Jeffries Matusita Distanzen und die dazugehörigen Merkmale, wobei ersichtlich wird, dass nach den Bewertungskriterien von PEREIRA et al. (2004:306) alle Klassen bis auf Straße und Industrie sehr gut zu trennen sind. Die besten Trennbarkeitsmerkmale wurden dann in der Klassifikation berücksichtigt.

Tabelle 2:Die besten Trennbarkeitsmerkmale nach Jeffries Matusita und die dazugehörigen Distanzwerte

Max JM Merkmal	Öl_kont	Schnee	Tundra	Tundra_vital	Industrie	Straße	Wasser_1	Wasser_2
Öl_kont		NIR|1.306	DVI|1.413	NIR_diff|1.410	Rot|1.412	Blau|1.370	Mean.Layer.5|1.386	DVI|1.414
Schnee			Max..diff.| 1.408	NIR_diff|1.412	Rot|1.399	SAVI|1.346	Mean.Layer.5|1.414	DVI|1.414
Tundra				EVI|1.345	SAVI|1.414	SAVI|1.414	DVI|1.414	DVI|1.414
Tundra_vital					NIR_diff|1.414	NDVI|1.414	Max..diff.| 1.414	NIR_diff|1.414
Industrie						EVI|0.542	NIR|1.414	DVI|1.414
Straße							Brightness|1.396	DVI|1.411
Wasser_1								DVI|1.414
Wasser_2								

4.2 Klassifikation

Im Laufe des Klassifikationsprozesses kommt es zu mehreren Klassifikationen die auf zwei verschiedenen Objektebenen und über unterschiedliche Merkmale der Segmente (spektral, klassenbezogen, Formeigenschaften) ablaufen (Abbildung 4). Im Level 1 wurden bei der Segmentierung kleine Objekte gebildet mit geringer Formberücksichtigung (Abbildung 4). Die ersten Klassifikationen im Level 1 erfolgen deshalb auf Basis von spektralen Merkmalen. Die zwei Wasserklassen sind in einer Klasse zusammengefasst und werden als erstes über NIR und SAVI Schwellwerte maskiert.

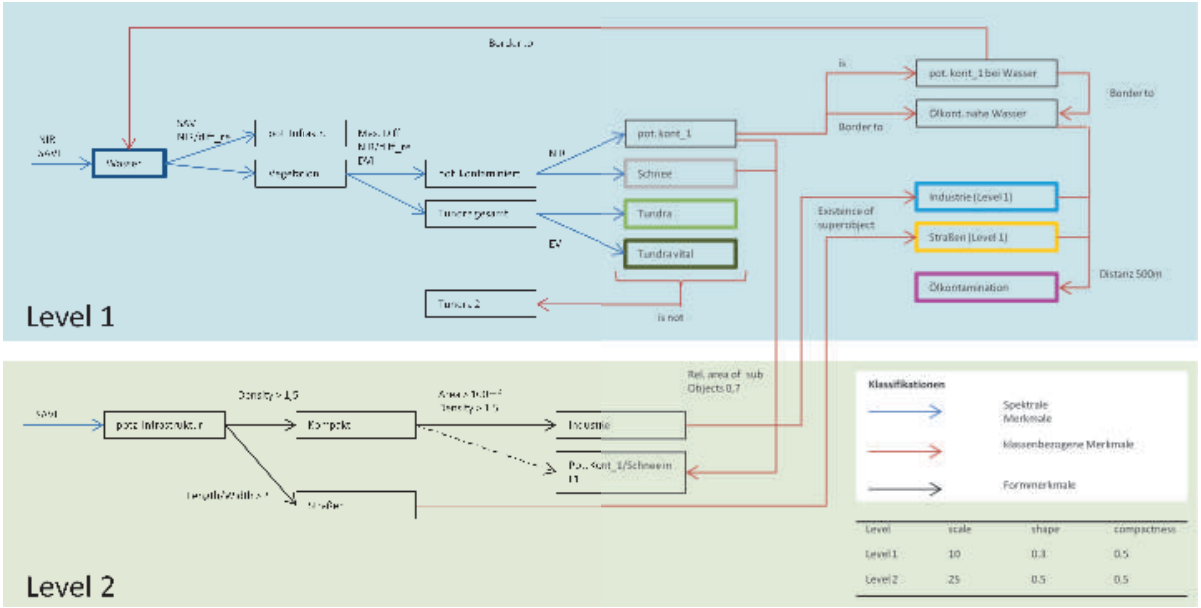


Abbildung 4: Klassifikationsschema für die Kartierung der ölkontaminierten Gebiete

Die Stichproben der übrigen Klassen erzeugen Überlappungen für die meisten Merkmale weshalb weiche Klassenübergänge über Zugehörigkeitsfunktionen definiert werden. Da die Klassenstichproben nicht normalverteilt sind, wurde der maximale Zugehörigkeitswert nicht bei dem Mittelwert einer Klasse gesetzt, sondern bei dem Median. Die Spannweite der Zugehörigkeitsfunktion wird mit Hilfe von weiteren statistischen Parametern festgelegt, wobei die Zugehörigkeitswerte an den statistischen Parametern für alle Klassen gleich sind. In einem zweiten Level wurde die charakteristische Infrastruktur der Ölförderregionen klassifiziert. Das Straßennetz und die Industrieanlagen zeichnen sich vor allem durch robuste Formmerkmale aus, weshalb bei der Neusegmentierung relativ größere Objekte erzeugt wurden (Abbildung 4). In einem ersten Schritt werden über den SAVI potentielle Infrastrukturbestandteile klassifiziert, die dann über *Area, Length/Width und Density (*TRIMBLE 2012:261,264,271*)* in Industrie und Straßenobjekte aufgeteilt werden. *Density* beschreibt die Dichte eines Objektes, wobei die höchste Dichte bei quadratischen Objekten erreicht wird (TRIMBLE 2012:271). Die Infrastrukturobjekte aus Level 2 werden dann auf das Level 1 übertragen. Die ölkontaminierten Gebiete werden abschließend aus den potentiell ölkontaminierten Gebieten abgeleitet, wobei eine Distanz von 500 m zu Straßen oder Industrieanlagen gegeben sein muss.

5 Ergebnisse

Für die Validierung wurden neue Testobjekte für die Zielklassen gesammelt. Die Wasserklassen wurden dafür zusammengefasst. Um die Übertragbarkeit des entwickelten Klassifikationsschemas zu testen wurde es ohne Anpassungen auf zwei eigenständige Validierungsgebiete (Abbildung 5) übertragen. In den Konfusionsmatrix von Tabelle 4 sind die Ergebnisse der Validierung für das Testgebiet dargestellt, sowie die *User, Producer und Overall Accuracy* von den Beiden Validierungssubsets. Es kann festgestellt werden, dass die Klassifikationsgenauigkeiten von den Validierungssubsets für die meisten Klassen in etwa gleich sind wie bei dem Testgebiet auf dem das Klassifikationsschema entwickelt wurde.

Tabelle 4: Validierungsergebnis für das Testgebiet und die zwei Validierungssubsets (*User, Producer, Overall Accuracy*)

TESTSIDE								
User \ Reference Class	Strassen	Industrie	Tundra Vital	SChnee	Tundra	Wasser	Ölkontaminieirt	Sum
Strassen	1641	0	0	110	0	0	0	1751
Industrie	351	1667	0	313	0	0	456	2787
Tundra Vital	0	0	3822	0	388	0	69	4279
SChnee	147	0	0	3547	145	0	505	4344
Tundra	0	0	27	0	5400	15	57	5499
Wasser	0	0	0	0	0	10286	191	10477
Ölkontaminieirt	47	0	0	0	256	0	4247	4550
unclassified	0	0	0	0	0	0	0	0
Sum	2186	1667	3849	3970	6189	10301	5525	
Producer	0.750686185	1	0.992985191	0.893450882	0.872515754	0.998543831	0.768687783	
User	0.937178755	0.598134194	0.893199346	0.816528545	0.981996727	0.981769591	0.933406593	
Overall Accuracy	0.908659127							
Producer	0.78098571	0.993510707	1	0.729064039	0.92953125	1	0.772560773	Subset 1
User	1	0.782319877	0.820323488	0.494156928	0.936407996	0.974493133	0.987654321	
Overall Accuracy	0.900209038							
Producer	0.879880932	0.972321963	0.992403781	0.533858727	0.99919363	1	0.727260006	Subset 2
User	0.935231714	0.689402779	0.991901468	0.829478458	0.947705545	0.846309112	0.934544147	
Overall Accuracy	0.889733909							

Die Nutzergenauigkeit für die ölkontaminierten Flächen liegt sehr hoch für alle Validierungsgebiete. Da *oil spills* die Reflektanz der ursprünglichen Landbedeckung im Allgemeinen senken, können fast alle Klassen (wenn diese sehr dunkel sind) fälschlicher Weise aus ölkontaminiert klassifiziert werden. Die Schneeflächen schneiden in der Klassifikation am schlechtesten ab, da es sich nicht um eine geschlossene Schneedecke sondern nur um eine dünne Schneedecke handelt bei der die unterliegenden Landbedeckungen eine Mischsignatur erzeugen. Dabei kommt es vor allem zu Fehlklassifikationen mit den Klassen Straße und Industrie.

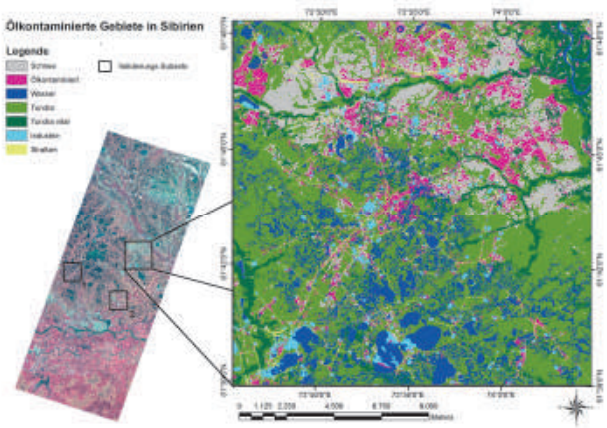


Abbildung 5: Klassifikationsergebnis im Testgebiet und Übersicht über die Validierungsgebiete

6 Diskussion und Zusammenfassung

Die Merkmale auf deren Grundlage die *oil spill* Klassifikation durchgeführt wurde basieren nicht auf spektralen Eigenschaften von Rohöl (Kohlenwasserstoffen), sondern auf deren Wirkung auf die Vegetation. Das bedeutet, dass andere Faktoren die sich negativ auf die Vegetation auswirken zu Fehlklassifikationen führen können. Aus diesem Grund wurde in dieser Arbeit ein maximales Distanzkriterium für *oil spills* eingesetzt, wonach vermeintlich kontaminierte Gebiete die mehr als 500m von der Infrastruktur entfernt sind nicht in die finale Klasse der ölkontaminierten Gebiete übergehen. Bei der Bewertung der Trennbarkeitsmerkmale auf Grundlage der Jeffries Matusita Distanz muss berücksichtigt werden, dass eine normalverteilte Stichprobe vorausgesetzt wird. Dadurch können Distanzmaße für die Jeffries Matusita Distanzen verfälscht sein und es kann im schlimmsten Fall zur Auswahl falscher Merkmale gekommen sein. Es konnte gezeigt werden, dass die Übertragung eines objektbasierten Klassifikationsschemas mit kontextbezogenen Merkmalen (Distanz) für die *oil spill* Kartierung möglich ist.

Wir danken dem RapidEye Science Archive (RESA) des DLR für die freundliche Bereitstellung der RapidEye Daten (RESA Projekt 472 „Large Scale Oilspill Mapping in West Siberia“).

7 Literaturverzeichnis

BRADOCK, J. F., LINDSTROM, J. E. & Prince, R. C., 2003: Weathering of a subarctic oil spill over 25 years: the caribou-Poker Creeks Research Watershed experiment. Cold Regions Science and Technology 36, 11 - 23.

EITEL J.U.H., VIERLING, L.A., LITVAK, M.E., LONG, D.S., SCHULTHESS, U., AGER, A.A., KROFCHECK, D.J. & STOSCHECK, L., 2011: Broadband, red-edge information from satellites improves early stress detection in a New Mexico conifer woodland. Remote Sensing of Environment 115, 3640-3646.

FINGAS, M., 2001²:The basics of oil spill cleanup. Boca Raton: CRC Press LLC.

FISCELLA, B., GIANCASPRO, A., NIRCHIO, F., PAVESE, P. & TRIVERO, P., 2000: Oil spill detection using marine SAR images. International Journal of Remote Sensing 21 (18) 3561- 3566.

GIRARD-ARDHUIN, F., MERCIER, G., & GARELLO, R., 2003: Oil slec detection by SAR imagery: potential and limitation. Occeans 2003. Proceedings (1), 164-169.

GIRARD-ARDHUIN, F., MERCIER, G., COLLARD, F. & GARELLO, R., 2005: Operational oil-Slick Characterization by SAR Imagery and Synergistic Data. Journal of Oceanic Engineering, 30 (3), 487-495.

GUNAL S. & R. EDIZKAN 2008: Subspace based feature selection for pattern recognition. Information Science 178, 3716-2726.

HESE, S. & SCHMULLIUS, C., 2008: High spatial resolution image object classification for terrestrical oil spill contamination mapping in West Siberia. International Journal of Applied Earth Observation and Geoinformation 11, 130-141.

HÖRIG, B. & KÜHN, F., 2001: HyMap hyperspectral remote sensing to detect hydrocarbons. International Journal of Remote Sensing, 22 (8), 1413-1422.

HUETE, A.R. (1988): A Soil-Adjusted Vegetation Index (SAVI). Remote Sensing of Environment 25, 295 - 309.

LI, L., USTIN, S.L. & LAY, M., 2005: Application of AVIRIS data in detection of oil-induced vegetation stress and cover change at Jornada, Noe Mexico. Remote Sensing of ENvironment 94, 1 - 16.

MARTINIS, S., GÄHLER, M. & TWELE A., 2003: A multi-scale Markov model for unsupervised oil spill detection in TerraSAR-X data. IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2012, S. 923-926.

MATSUSHITA, B., YANG, W., CHEN, J., ONDA, Y. & QIU G., (2007): Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to Topographic Effects: A Case Study in High-Density Cypress Forest. Sensors, 7, 2636-2651.

PEREIRA, J.M.C., MOTA, B., PRIVETTE, J. L., TAYLOR, K. K., SILVA J. M. N., SÁ A. C. L. & NI-MEISTER, W., 2003: A simulation analysis of the detectability of understory burns in *miombo* woodlands. Remote Science of Environment 93, 296 - 310.

PINAR A. & CURRAN P.J., 1996: Technical Note Grass chlorophyll and the reflectance red edge. International Journal of Remote Sensing, 17 (2), 351-357.

TRIMBLE, 2012: Reference Book. München: Trimble German GmbH.

WIDGET, A., 1997: Black Snow: Oil and the eastern Khanty. <http://www.nmsu.edu/Academic_Progs/CAS/english/public_html/hc/IMPACTOIL.html> (Stand: 1997-08)(Zugriff: 2012-012-10).

Evaluation of high resolution digital surface models for single tree extraction approaches in mixed forests

MOHSEN MIRI[1], STEVEN BAYER[2] & TILMAN BUCHER[3]

Abstract: High resolution digital elevation models (DEM) are utilized in automatic extraction of single-trees in mixed forests. In this paper the digital surface models (DSM) provided from aerial images with resolutions of 8 cm and 20 cm and airborne laser scanning ALS data are investigated. The results showed a relative good representation of the tree crowns even in image-based DSM. However, increasing the resolution of imagery does not inevitably lead to better quality in DSM of forests. To evaluate the characteristics of deciduous and coniferous trees on such models, two features are introduced. The Inversed Quasi-Flatness (IQF) and the Averaged Height Variation (AHV) of the DSM are implemented on some automatically selected samples at the highest levels of elevation on nDSM to demonstrate the geometric and morphologic characteristics of the different DSMs.

1 Introduction

Knowledge of tree-crown parameters such as height, shape, and crown closure is desirable in forest and ecological studies, though those parameters are difficult to measure on the ground (SHENG et al., 2001). Thanks to rapid developments in high density Airborne Laser Scanning (ALS) data and High Resolution Digital (HRD) images, both in technical and processing points of view, providing valuable geospatial information in forestry applications has increasingly become available (LECKIE et al. 2003). The necessity and the great interest in more detailed information, concerning single tree delineation and classification, especially for mixed forests, has emerged. More accurate quality and quantity information of individual trees requires detailed three dimensional positioning and accurate attributing of single trees. For this end, a normalized digital surface model (nDSM), also called canopy height model (CHM), has to be provided by subtracting DTM from DSM, which shows the importance of both DTM and DSM data in forestry areas. ALS data have proven their capabilities in capturing such elevation data. On the other side, image-based techniques, like Semi Global Matching (SGM, HIRSCHMÜLLER, 2005), can also provide good results in DSM generation in forestry. Due to the fact that the repetition rate of statewide coverage with airborne imagery is performed every 3 years, these data pose a huge and cost efficient potential for forest monitoring. For this reason LIDAR and photogrammetric DSMs are evaluated, whether or up to which extend the image-based digital surface models fulfill the forestry requests. This paper evaluates both laser-based and image-

1) Mohsen Miri, Free University of Berlin & Beuth University of Applied Science Berlin, FB III, Photogrammetry lab., Luxemburgerstr. 10, 13353 Berlin, Germany; Email: mohsen.miri@fu-berlin.de
2) Steven Bayer, Institutefor International Urban Research (InUrban), INA gGmbHat Free University of Berlin, Malteser Str. 74-100, 12249 Berlin, Germany; Email: steven.bayer@fu-berlin.de
3) Tilman Bucher, German Aerospace Center (DLR), department of Sensor Concepts and Applications, Rutherfordstr. 2, 12489 Berlin, Germany; Email: tilman.bucher@dlr.de

based DSM data at single tree level over closed mixed forests in Germany. Additionally the characteristics of the canopy height models are evaluated on automatically selected samples on the top level of the forests, using two geometric and morphologic features.

2 Image-based and laser-based data acquisition in forestry

As laser scanners can provide 3D point clouds from different heights of forest, employing the ALS data in forestry has been reaching satisfying results, even at single trees levels. In (HYYPPÄ et al. 2004) the algorithms and methods used in this area are summarized in several aspects. Along with airborne laser scanners, high resolution aerial images, as one of the basic and rich sources of detailed geospatial information for forest inventory, has been used for the relevant applications in large extend. However, image-based methods are confronted with the limitations of matching algorithms. Table 1 shows an overview of the pros and cons of both data acquisition methods in extracting geospatial attributes in forestry.

Capabilities and Potentials	**ALS**	**Imagery**
Direct measurement of the surface	✓	✗
Textural contents	!	✓
DSM generation	✓	✓
DTM generation in dense forests (or in leafy season)	✓	!
DTM generation in sparse forests (or leafless season)	✓	✓
Independence from illumination conditions	✓	✗
Costs for repeatable measurement	✗	✓
Independence from texture of the object	✓	✗
Spectral information	!	✓

Table 1: Comparison of laser-based and image-based data potential for forestry applications (potential level: ✓ high; ! medium; ✗ low).

Considering the above table, one may prefer to employ the ALS data to generate the digital elevation models, required in forestry. However, regarding the high costs of laser-based techniques, in one hand, and the richness of information in image-based methods, in the other, an evaluation of the DSM obtained from image-based is needed. It should be mentioned, that the capability of the laser-based methods to produce the point cloud of the ground under the crowns should not been unseen. As the ground surface remains more unchanged than the crown closure, typically, laser-based DTM can be used as a reference for nDSM over several years.

3 Dataset

In this paper, the image-based DSMs are generated by SGM. The first area of interest, near Hoppengarten, contains DSM generated from UltraCam XP (8 cm and 20 cm GSD) and the rasterized laser-based DSM with 40 cm GSD, acquired by LMS-Q560 from Riegl. The feature analysis is performed on the second area near the city Steinfurt using a DSM of a DMC-1 camera from the company Intergraph with 20 cm resolution and additionally the DSM of ALS data of LMS-Q560 from Riegl with 20 cm GSD. Both areas are located in state North-Rhine-Westphalia (NRW) in Germany.

4 Image-based DSM from SGM

To evaluate the image-based DSM, two sets of aerial images captured by UltraCam XP from Microsoft with 8 cm und 20 cm GSD are used. Both datasets are compared separately with DSM of LIDAR data.
Figure 1 shows a true-orthophoto of the Hoppengarten test area.



Figure 1: True-orthophoto of the test area Hoppengarten; small subset see Figure 2

For a better comparison of the tree forms in image-based and laser-based datasets, digital surface models are demonstrated in Figure 2. Although the laser data are free from noises and flying points, because of their measuring characteristics, they represent the rough and complex tree surfaces rather heterogeneous. As can be seen, the results of the SGM method, leads also to an appropriate representation of the canopy surface, though with smoother surfaces, especially in lower resolutions.

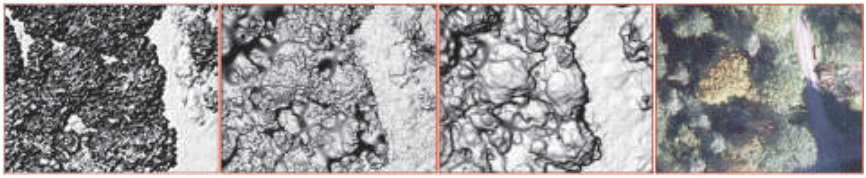


Figure 2: Test area Hoppengarten; DSM Laser 40cm, photogrammetric DSM 8cm and 20cm, RGB 8cm

A comparison between the image-based and laser-based DSM is carried out on the 8 cm and 20 cm images in this area. The differentiated images, shown in Figure 3, demonstrate well fit results (green colors) on the ground and also at the highest levels of elevation, in both resolutions. The interpolated areas, mostly demonstrated in blue colors, confirm the typical problems of stereo matching algorithms in shadowed and occluded areas. This comparison shows that although, in many cases, an increase in spatial resolution of aerial images leads to more detailed DSMs, it does not necessarily result in better surface models. Complexity of the texture on the trees, different representation in overlapping images and the illumination conditions can be named as some problematic aspects, which influence the performance of matching algorithms in forestry areas.

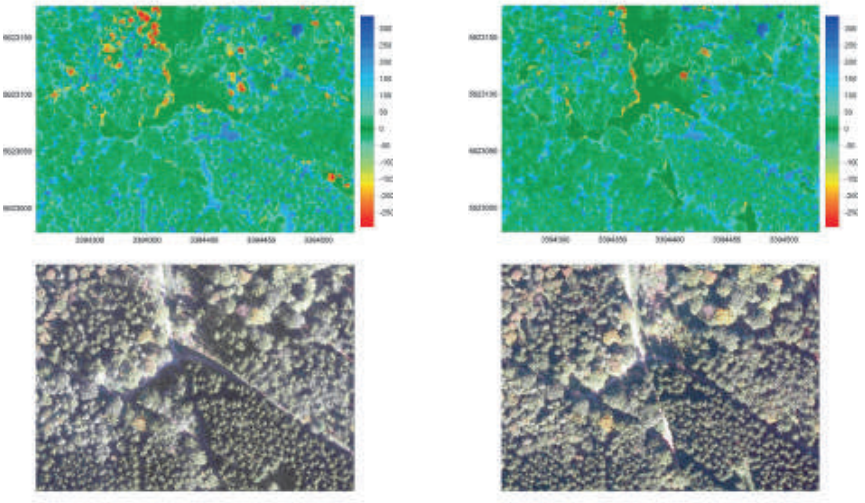


Figure 3: Top: difference images with colorized height differences in decimeter; DSM_{image_8cm} – DSM_{LIDAR_40cm} (left), DSM_{image_20cm} – DSM_{LIDAR_40cm} (right). Bottom: True orthophotos 8 cm (left) and 20 cm (right).

For the further evaluations on the second test area, the photogrammetric images with 20 cm resolution have been used. This dataset is, in our opinion, representative for a standard data type, which can be provided using SGM or similar matching algorithms.

5 Geometric and morphologic features

Due to the irregular shape of the tree crowns and the fine structure of the branches, DSMs of forested regions are difficult to generate and challenging to extract geometric and morphologic features. Coniferous and deciduous trees react differently in laser-based and image-based DSM

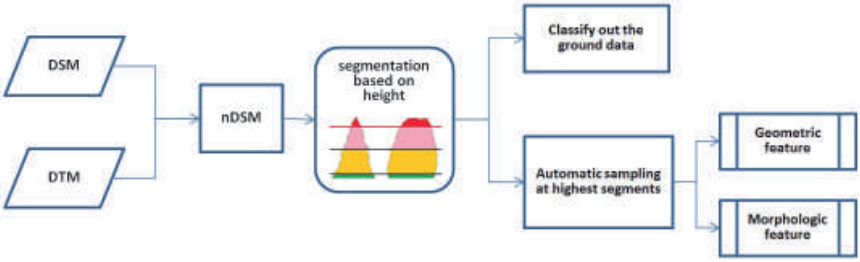


Figure 4: Processing flowchart of the feature-based evaluation of the DSM data at the level of THS (red)

data. These features can be measured using statistical parameters in order to recognize deciduous and coniferous trees; those should not be dependent on the segment height. As the locational information of the trees in remotely sensed datasets, especially on the crown, focuses mostly on their position at local maxima, the evaluation will be done at the highest levels of these datasets. The concentration on local maxima segments also avoids the critical parts of the DSM, like shaded or occluded areas between the trees. In this paper two features are defined and classified to evaluate the Top Hat Slice (THS) of the trees in several segmentation scales using a combination of pixel-based and object-based methods. Figure 4 represents the flowchart of this procedure.

In the first step, the nDSM is extracted from the DSM with 20 cm pixel size and the laser DTM of one meter (DGM-1, GEOBASIS-NRW). The next step consists of an automatic segmentation, based on the elevation data of the nDSM. This step is followed by the selection of the segments including the most significant local maxima; the top hat slices are generated. Areas below 5 m in the nDOM are classified and attributed as ground data.

For the top hat slices, two hypotheses (deciduous / coniferous) are studied, introducing two features: the *Inversed Quasi Flatness (IQF)* as geometric feature and the *Average Height Variations (AVH)* as morphologic feature. To study the reaction of these features on nDSM data the correctness of the selected samples are assessed in coniferous and deciduous stands using true-orthophotos and the terrestrial data is based on the formula employed by STRAUB (2003):

$$\text{Correctness}\,(\%) = \frac{\text{truely estimated segments}}{\text{truely estimated segments} + \text{false estimated segments}}, \quad [0]$$

A detailed assessment of the accuracy and correctness of the image-based DSM has been performed by BAYER et al. (2013), based on terrestrial laser scanning and photogrammetric reference data.

5.1 Inversed Quasi Flatness

Approximately, a tree can be modeled with known geometric 3D shapes by spherical, ellipsoidal or conic parameters. The IQF describes the flatness of the trees at the highest segments of the nDSM. At this level the deciduous trees are expected to be flatter than coniferous ones, which yield higher values in IQF. Figure 5 shows the color-coded results of this feature. The estimations in the laser nDSM (left) show good results for coniferous trees but relatively poor recognition of the deciduous ones. This is mainly due to the small segment size at THS level, which overestimates the IQF in this tree type. In contrast, the results in the photogrammetric nDSM show relatively good recognition of true segments in deciduous trees and lower correctness for coniferous ones (see also Table 2).

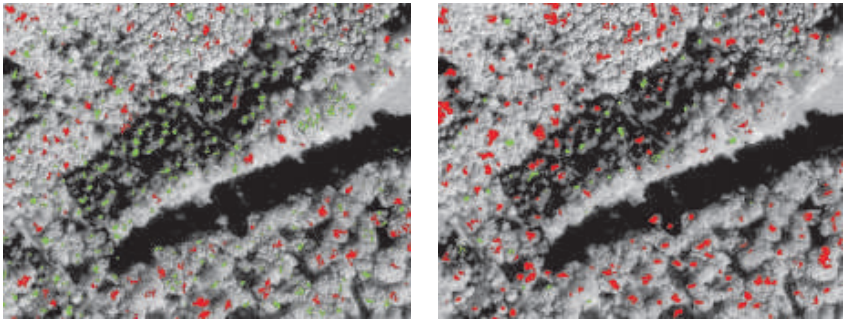


Figure 5: Comparison of IQF on laser-based nDSM (left) and image-based nDSM (right) at low segmentation level for ■ coniferous and ■ deciduous tree types.

5.2 Averaged Height Variations

The second feature, AHV, is provided to measure the averaged variation of height at the THS level. Comparing the coniferous and deciduous trees in leafy season, it is expected that the latter react with more scattered values on the upper surfaces but less height variation among the segment.

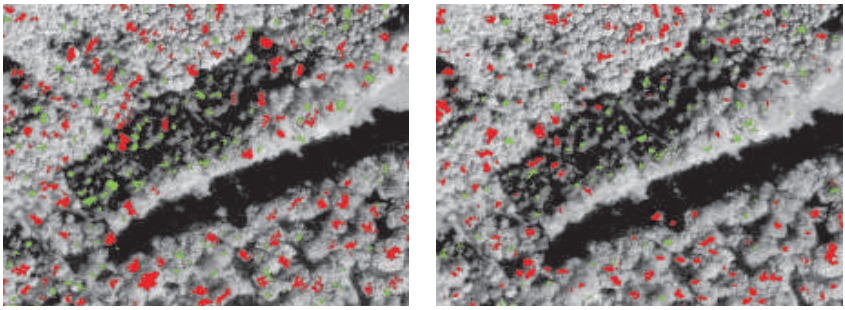


Figure 6: Comparison of AHV on laser-based nDSM (left) and image-based nDSM (right)at low segmentation level for ■ coniferous and ■ deciduous tree types.

As seen in Figure 2 LIDAR data exhibit strong local variations. These are reduced in the first steps of segmentation on ALS data using a 3x3 median filter to avoid the oversegmentation. The results show that coniferous trees are correctly detected and decreasing the segment size leads to better detection of these tree types. On the other side, due to the characteristics of the laser data, more detailed segments decreases the generally low detection probability of deciduous trees. In image-based nDSM the capability of this feature to separate deciduous trees seems to increase with smaller segment sizes.

nDSM	Segmentation detail level	Stand type	IQF (%)	AHV (%)
Laser-based	Low	Coniferous	88.09	85.71
		Deciduous	63.79	50
	High	Coniferous	98	98.43
		Deciduous	37.62	25.74
Image-based	Low	Coniferous	52.38	78.57
		Deciduous	94.74	24.56
	High	Coniferous	70	91.25
		Deciduous	68.478	41.30

Table 2: suitability of the IQF and AHV on laser-based and image-based data

6 Conclusions and Outlook

High resolution DSM generated from laser-based and image-based techniques are evaluated in different resolutions. Comparing two photogrammetric DSMs, a higher resolution lead to a more detailed DSM, however due to the limitations in stereo matching algorithms such as occlusion and shadows, the results especially in forests did not improve significantly. Therefore the standard resolution 20 cm has been used for further investigations. To evaluate the characteristics of deciduous and coniferous trees in the models, two features have been introduced to measure the flatness and roughness value of the DSM data in automatically selected segments at top head slices. The results show that both features work properly in coniferous stands. However the deciduous trees react differently. Because of the different morphology of the two tree types, in on hand, and the stand density, in the other, separation of the trees into different feature-classes is necessary. This can exemplarily be seen in Table 2, where the accuracy of the deciduous trees is much higher than for the coniferous trees, which depends on the segment size. This effect can be used for the classification of deciduous trees without spectral information. After classification of deciduous ones, the remaining trees can be extracted with smaller segment sizes and smaller search radius for detecting the local maxima. This results in a higher accuracy in recognition of the deciduous trees, too. Further works will be done to optimize the parameters for these features in and also the on the preprocessing steps of laser and photogrammetric DSMs.

7 Acknowledgment

The authors would like to acknowledge the great support of RWHT Aachen and RIF e.V. in the project “Virtual Forest”. The remote sensing data was recorded for this project, which is co-financed by the European Union and North-Rhine-Westphalia – European Regional Development Fund (EFRE). We would like also thank Geobasis NRW for preparing the datasets and their technical support. We dedicate our great thanks to Ingenieurbüro GILAN in Berlin and ZeDe3D Ltd. in Brandenburg for their personal and technical support in field data acquisition steps.

8 References

BAYER, S., MIRI, M. & BUCHER, T., 2013: Automatisierte Einzelbaumerfassung auf Basis von hochaufgelöstenTrueOrtho-Mosaiken und photogrammetrisch abgeleiteten Oberflächenmodellen. D-A-CH conferece, 2013, Proceedings (current volume).

HIRSCHMÜLLER, H., 2005: Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. – Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 20-26 June 2005, **2**, S.807-814.

HYYPPÄ, J., HYYPPÄ, H., LITKEY, P., YU, X., HAGGRÉN, H., RÖNNHOLM, P., PYYSALO, JUHO, U., PITKÄNEN, J. & MALTAMO, M., 2004: Algorithms and Methods of Airborne Laser-Scanning for Forest Measurements, ISPRS proceedings, October, Page(s) 8.

LECKIE, D.; GOUGEON, F.; HILL, D.; QUINN, R.; ARMSTRONG, L. & SHREENAN, R., 2003: Combined high-density lidar and multispectral imagery for individual tree crown analysis; Canadian Forst Service Publications, Vol. 29, pp. 633–649.

SHENG, Y.; GONG, P. & BLGLNG, G. S., 2001: Model-Based Conifer-Crown Surface Reconstruct ion from High-Resolution Aerial Images. PE&RS, August, P. 957–965.

STRAUB, B.-M., 2003: Automatische Extraktion von Bäumen aus Fernerkundungsdaten. Wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover, Nr. 249.

Change detection by means of multitemporal satellite imagery for the Indian Mega-City Hyderabad

MAIK NETZBAND[1]

Abstract: This paper describes the temporal urban monitoring of a large Indian agglomeration by comparing satellite images. The emerging megacity of Hyderabad (India) serves as an example of a rapidly growing agglomeration on the Indian subcontinent with one of the most extremely increasing number of inhabitants during the past decades. The strength of satellite remote sensing is to structure and classify such large and complex urban areas as well as evaluating their spatial development by change detection studies. Since there are hardly or only very few GIS based map bases for these large cities available in India, the satellite-based remote sensing represents a chance to structure and map such expanded urban residential areas at least approximately into different land coverage or land use classes.

1 Introduction

India's urban population has grown tremendously in the last four decades from 79 million in 1961 to 285 million in 2001. This fast rate of increase in urban share is mainly due to large scale migration of people from rural and smaller towns to bigger cities in search of better employment opportunities and good life style. The rapid population pressure has resulted in unplanned growth in the urban areas to accommodate these migrant people which in turn lead to urban sprawl. Large slum areas exist from uncontrolled and unplanned immigration into the megacities of Asia, partly appearing even within the inner parts of the city, in which humans are living under catastrophic conditions. Additionally strong economic and industrial growth, often in direct neighborhood to populated areas, causes health problems for the adjacent residents within the respective cities.
In this paper an attempt has been made to use temporal change detection as well as Shannon's entropy model to assess the urban sprawl for one of the fastest growing city of South India (Hyderabad-Secundarabad) and its surrounding area.

The leading research questions are:

- Is it possible to track the growth of a megacity using high to medium resolution satellite images over a given period of observation?
- Can we make by these observations general statements about certain types of growth, such as particularly strong growth of settlement space, loss of green space, etc.?
- What are the strengths and weaknesses of the used satellite data and the applied classification approaches and results?

1) Maik Netzband, Ruhr-Universität Bochum, Geographisches Institut, Universitätsstraße 150, 44801 Bochum; E-Mail: maik.netzband@rub.de

2 Data and Method

2.1 Study Area and Data

To study the urban sprawl in an emerging megacity of South Asia, Hyderabad-Secundrabad, the twin capital city of the state Andhra Pradesh in India and its surrounding area, has been chosen because it is one of the 33 million cities of India (CENSUS OF INDIA 2001) which has one of the highest urbanization rates (27%). Hyderabad city is situated on the Deccan Plateau region in the Krishna river basin. River Musi is a tributary of river Krishna that physically divide Hyderabad city into two half, the north of Hyderabad which is the new development called Secundrabad and the south of Hyderabad which is the old city. The Musi River used to flow over its rocky bed with mere trickle of water for more than 8 months in a year but today it is a major sewage drain. With the urban sprawl after 1950 this area has changed considerably, forests and grasslands have been cleared and reclaimed by new settlements colonies. Inadvertently, this is resulting in the increase in the built-up area and associated changes in the spatial urban land-use patterns. Consequences are the loss of other land cover and land uses such as productive agricultural lands, forest, green spaces, surface water bodies, beyond, depletion in ground water aquifers and increasing levels of air and water pollution.

For this analysis ASTER (=„Advanced Spaceborne Thermal Emission and Reflection Radiometer“) satellite data with a resolution of 15x15 m per pixel from the years 2001 and 2009 have been used. The area that has been covered by the data extends over the largest part of the city.

2.2 Multitemporal and Multispectral Analysis

In detail the following analysis steps have been conducted:, i.e.

- Multitemporal urban Land Cover/Land Use surface and structure classification analysis, basic urban parameters: degree of imperviousness, extent and quality of green and open areas, urban structure types and brownfields/potential development areas; Examining whether a spatial correlation exists between the results of the different thematic land-use/ land-cover analyses
- Identification of land-use patterns combined with a vegetation index analysis (NDVI) and Urban Structure Types (UST); and
- Estimation of spatial indicators for quality of life and vulnerability to natural hazards such as flooding.

2.3 Modeling Urban Sprawl by Calculating the Entropy

For all 35 wards (boundary defined by the Municipal Corporation, Hyderabad) and the four outer zones 36-39 on each corner i.e. NE, NW, SE and SW zone (the remaining portion beyond the Municipal boundary i.e. within the 20’x20’ grid), built-up area was clipped in ILWIS GIS software from both 1971 and 2005 classified land use/land cover (LU/LC) maps. This was done because the entropy model takes into account the relationship between built up area and the open

area. Since entropy is used to measure the distribution of a geographical phenomenon and the difference between two time periods (t) and (t+1) i.e. (1971 and 2005), the entropy change was also calculated that can indicate the change in the degree of dispersal of land development or urban sprawl in two time period using equation (1).

$$\text{Entropy change } \Delta E = (t+1) - (t) \qquad (1)$$

3 Results

The physical expansion of Hyderabad was very slow till recently, i.e. in 1501 the city covered an area of about 15 km² and population of few thousands. But gradually city expanded and by the end of 1944 its area reached to 75 km² which further expanded to 179 Km² by 2005. In the North of Hyderabad, Husain Sagar Lake provides water for the growing urban population so the further expansion of this twin city took place mainly in the North which is the newer part called Secundrabad.

3.1 Multitemporal Analysis

The scientific results are change maps of the urban agglomeration of Hyderabad between 2001 and 2009 using temporal transects of remote sensing data and applying up-to-date change detection techniques at different spatial and temporal scales. This is used for quantifying urban and peri-urban processes (land use / land cover changes of settlements, agriculture, industry, and landscape) in this urban agglomeration (growth rates, urbanization) and additionally, for predicting the development of urban fringes, rural settlements, informal settlements, urban and peri-urban agriculture. Data of macroeconomic and demographic development, instruments of urban planning, and socio-economic settings have been integrated as data were available.

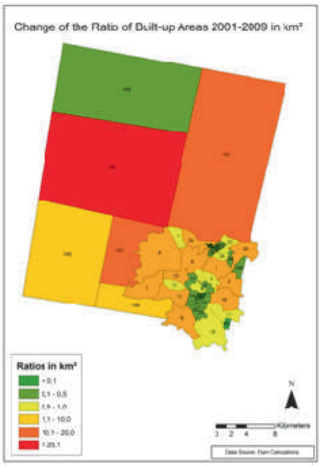


Figure 1. Example Change Map 'Built-up Area' Hyderabad 2001-2009

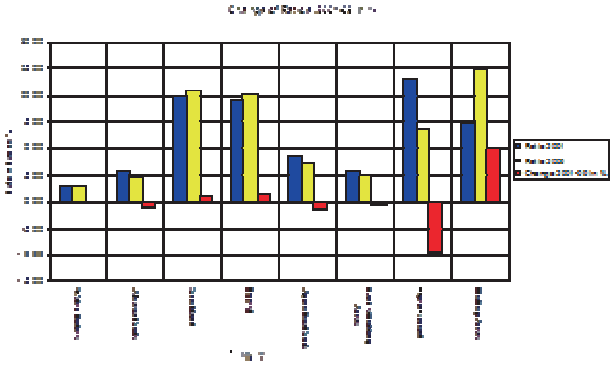


Figure 2. Change Ratios for all classes 2001-2009

3.2 Entropy Calculation

Table 1 shows the area in different land use/land cover classes in 1971 and 2005, agricultural land reduced from 280.60 to just 60.59 km² whereas the waste land have reduced from 529.60 to 385.53 Km2 during 1971 to 2005 which mainly due to urban sprawl. Shannon's Entropy value (Hn) was calculated for 35 wards and also for four fringe zones (no. 36-39) i.e. SE, NE, NW and SW. The result shows that urban sprawl has occurred in all the wards of the twin city but not at the same rate. Sprawl has been felt more in the fringe wards but the wards which are in the city centre have also experienced development i.e. vertical expansion. Some of the lands, which were open / vacant till last few years, are now occupied with high rising buildings.

TABLE I. Change in Entropy values in Hyderabad

Year	Built-up Area (km²)	Increase in built-up area %)	Entropy value (En)	Entropy change(Δ E)
1971	135	Base year	0.627	Base year
1989	214	58.52 (71/89)	0.712	13.6%
2001	288	34.58 (89/01)	0.794	(71/89)
2005	370	28.47 (01/05)	0.918	26.6%

The proportion of the total population in a region to the total built-up is a measure of quantifying sprawl. A comparative picture was drawn to assess the change in the built-up area (1971-2005) in different wards/zone. The biggest change in the built-up area can be noted in the outer four zones i.e. 38-40. Comparing built-up (urban expansion) vs. population, the rate of development of land in the Hyderabad-Secundrabad region is far outstripping the rate of population growth. This implies that the land is consumed at excessive rates and probably in unnecessary amounts as well. Between 1971 and 2005, population in the region grew by about 124% (CENSUS OF INDIA, 2001) while the built-up area increased by about 174%.

This means that the per capita consumption of land has increased markedly during three and half decades. The per capita land consumption is the utilization of all the land development initiatives like the commercial, industrial, educational, and recreational establishments along with the residential establishments per person. This is also due to the increase in the per capita income of the people, because this twin-city is one of the Hi-Tech cities of India, where people are mainly engaged in highly paid jobs. So because of increase in income and purchasing power multi-story houses (apartment complexes) are being built on the low cost fertile agricultural land and that in turn leads to urban sprawl.

3.3 Built-up land as an Indicator of Urban Sprawl

The class 'Built-up land' is a composition of different surface types, e.g. impervious built-up like buildings, roads and also smaller unsealed open spaces. This class is the most important in the inner parts of the city and covers the most of the city area. The percentage of an area covered by impervious surfaces such as asphalt and concrete is a straightforward measure of development (BARNES et al, 2001). The built-up is generally considered as the parameter of quantifying urban sprawl (TORRENS & ALBERTI, 2000). It can be quantified by considering the impervious or the built-up as the key feature of sprawl, which is delineated using top sheets or through the data acquired remotely (Epstein et al., 2002). So, the developed areas have greater proportions of impervious surfaces, i.e. the built-up areas as compared to the lesser-developed areas. Since the sprawl is characterised by an increase in the built-up area along the urban and rural fringe, this attribute gives considerable information for understanding the behaviour of such developments. In this study to quantify this attribute since most of the expansion is taking place on the agricultural land so, land transformation map was generated. The figure shows that urban expansion has taken place all around the twin-city during (1971-2005) and 262.02% of fertile agricultural land has lost mainly to built-up land (table 2).

TABLE II. CHANGE IN ENTROPY VALUE FOR CLASSES

S. No.	Landuse/land cover	class Gain (%)	Loss (%)
1	Built up lands	165.50	-66.37
2	Agriculture	47.01	-262.02
3	Forests	109.68	-79.35
4	Water bodies	3.55	-42.54
5	Waste lands	150.63	-294.76
6	Others	305.17	-36.51

4 Conclusions

The multitemporal change detection study with mid-resolution ASTER data from 2001 and 2009 reveals the following results:

- Rate of development of land in the Hyderabad-Secundrabad region far outstripping the rate of population growth
- Land is consumed at excessive rates and probably in unnecessary amounts
- Per capita consumption of land increased steadily over three and half decades.
- 'Where and how urban India is developing to in the next decades?'

The study also demonstrates the usability of entropy calculation to identify measure and monitor urban sprawl in Hyderabad-Secundrabad, twin-city and its environs, using remote sensing and GIS techniques. The entropy method can be easily implemented using GIS to facilitate the measurement of urban sprawl. The study suggests that entropy is a good indicator for identifying the spatial dispersion and land development. The increase in entropy value from 0.63 in 1971 to 0.98 in 2005 confirms that urban sprawl and land consumption has taken place. There is a significant entropy change (29%) in the last 35 years. The sprawl is normally uneven with highest value being in the NW of the city that can be seen with its high entropy value of this zone no. 38, because of development of Hi-Tech computer software companies. The increase of entropy values indicates that there is an increase in urban sprawl and the urban growth tends to be more dispersed over a period of time. The entropy values in sub-urban areas are much higher than in the urban core indicating rapid urbanization process in the fringe areas of the city.

5 References

BRUMMITT, N. & LUGHADHA, E. N., 2003: Biodiversity: Where's Hot and Where's Not. Conservation Biology, **17** (5), S. 1442-1448.

CENSUS OF INDIA 2001, New Delhi.

BARNES K. B., MORGAN III J. M., ROBERGE M C., & LOWE S. 2001. Sprawl development: Its patterns, consequences, and measurement, Towson University, http://chesapeake.towson.edu/landscape/urbansprawl/downlo:ad/Sprawl_white_paper.pdf

EPSTEIN, J., PAYNE, K & KRAMER, E., (2002). Techniques for mapping suburban sprawl. Photogrammetric Engineering and Remote Sensing, Vol. 63 (9): pp 913 - 918.

TORRENS P. M., & ALBERTI M., (2000). "Measuring sprawl", Working paper no. 27,Centre for Advanced Spatial Analysis, University College London. Available online: http://www.casa.ac.uk/working_papers

A Concept for Guiding the Learning of Conditional Random Fields for the Classification of Urban Areas in SAR Images

TESSIO NOVACK[1], OLIVER MAKSYMIUK[1] & UWE STILLA[1].

In dieser Arbeit wird ein Konzept zur Klassifikation urbaner Strukturen in SAR Intensitätsbildern vorgestellt. Die Methode basiert dabei auf der Verwendung von Conditional Random Fields (CRF), deren Netzwerkstruktur und Parametrisierung automatisch durch das Optimieren einer Zielfunktion gelernt werden soll. Allerdings handelt es sich um ein schlecht gestelltes Problem, was deutliche Herausforderungen an die Algorithmen stellt. Wir zeigen einen effektiven Weg, um dennoch eine Optimierung erreichen zu können und welchen Vorteil regelbasierte Verfahren, die durch den Nutzer definieret werden, in diesem Zusammenhang haben.

1 Introduction

Regarding the classification of high-resolution SAR imagery, the consideration of the spatial context is practically indispensable. This can be done by (1) calculating texture attributes, (2) possibly using these and/or other attributes as inputs to an image segmentation procedure and (3) considering contextual attributes that relate neighbouring image regions. Probabilistic Graphical Models (PGMs) are powerful probabilistic frameworks that allow the modelling of complex contextual relations between neighbouring image regions (or pixels) regarding their attributes and their possible classes. As a specific PGM type, Conditional Random Fields (CRFs) focus directly on estimating the posterior probability distribution of the possible classifications of the scene given the observed images attributes (KOLLER & FRIEDMAN, 2009). The automatic learning of the CRF network's structure and its parameters based on sample data is known to produce models of better performance and potentially reveal unsuspected dependencies between the observed and unobserved variables (GANAPATHI, 2008). Nevertheless, the automatic learning of CRFs models very easily becomes computationally intractable, due to the necessity of running inference over the current CRF network at each step of the optimization procedure and, if discriminative training is carried out, for each data sample. This problem is worsened by the fact that very often remote sensing image regions may contain a large number of attributes with large cardinality. In this paper, we show a framework for the automatic learning of the CRF network's structure and parameters in a tractable way. This framework was customized aiming its application on remote sensing image classification and having two guidelines in mind: (1) reduce as much as possible computational costs and (2) allow the inclusion of image interpretation knowledge in the learning process. We give a brief explanation on the theoretical background of our framework and exemplify how it would be applied for the classification of Urban Structure Types based on high-resolution (In)SAR imagery.

1) Photogrammetrie & Fernerkundung, Technische Universität München, Arcisstrasse 21, 80333 München, http://www.pf.bv.tum.de

2 Some theory behind the proposed framework

2.1 Probabilistic Graphical Models and CRFs

Probabilistic Graphical Models (PGMs) enable the representation of complex joint distributions with considerably fewer parameters. Such representation is given in the form of a network whose nodes are random variables and the links between them represent the probabilistic dependencies and conditional independencies among the variables of the network. A fully connected sub-network is called a factor. A potential is associated to every possible assignment of the variables values in a factor. It should be noted that in most cases the variables are modelled as being discrete. Most effective inference algorithms only work for discrete variables. A PGM is fully defined when its network structure and potentials are defined. The problem of defining the structure and the parameters of a PGM is the problem of defining how the joint distribution factorizes and which the potentials from each factor are. Each factor of a CRFs model must contain at least one unobserved variable. In the context of remote sensing image classification, the observed variables are the image attributes and the unobserved variables are the classes of interest.

2.2 Log-linear Models

Log-linear models are a compact representation of undirected PGMs and CRFs which are defined in terms of a set of feature functions $f_k(X_k)$. The variables X_k in every feature k are exactly the ones in one of the factors of the network. Given a set of feature functions $F = \{f_k\}$, the parameters of the log-linear model are weights $\theta = \{\theta_k : f_k \in F\}$. The overall distribution is then defined as:

$$P_\theta(x) = \frac{1}{Z(\theta)} \exp\left(\sum_{f_k \in F} \theta_k f_k(x_k)\right) \tag{1}$$

where x_k is the assignment to X_k within x, and $Z(\theta)$ is the partition function that ensures that the distribution is normalized. The term $f_k(X_k)$ can be any function that defines a numerical value for each assignment x_k. A commonly used function is the indicator function, which returns value 1 in case of certain x_k assignment and 0 otherwise. The weight θ_k becomes than the potential for this assignment. Commonly, one such indicator function is created for each possible assignment from each factor.

2.3 Parameter Learning

A few approaches have been proposed for learning the parameters of undirected PGMs (KOLLER & FRIEDMAN, 2009), the maximization of the log-likelihood function being the simplest variant of this task. The log-likelihood objective function has the form:

$$l(M, \theta : D) = \sum_{f_k \in F} \theta_k \left(\sum_m f_k(\xi[m])\right) - M \log Z(\theta) \tag{2}$$

where M is the number of samples and m is a single sample from the training data set. This function is a convex one, which means it has no local optima but a global one. However, it has

no closed-form solution and because of that its global optimum has to be found through an iterative optimization process (like gradient descent, for example) which uses the gradient between the empirical counts of the features (calculated directly from the sample data set) and their marginal probability given a parameter set θ to guide the optimization. Formally we want that for each feature *k*,

$$(\sum_{m} f_k(\xi[m])) - ME_{x\sim P_\theta}[f_k(x)] = 0 \quad (3)$$

In order to calculate the second term of this expression (i.e. the marginal probability of a certain feature), inference should be performed over the current model. Because usually several iterations are needed until convergence is reached, inference is run many times, what makes this process computationally very costly.

2.4 Structure Learning

The problem of finding the best structure of the CRF network is equivalent to the problem of finding the best set of features *F* given the set of observed and unobserved variables. Including the L1-regularization term to equation (2) and maximizing it has the effect of forcing many features to have weight zero when convergence is achieved. At the end of the optimization process, the features that have weights different than zero represent the factors of the CRF network, i.e. the edges and nodes of the CRF structure. This L1-regularized log-likelihood objective function has the form:

$$\max_{\theta} = \left[\sum_{f_k \in F} \theta_k (\sum_{m} f_k(\xi[m])) - M \log Z(\theta) - \sum_{k} \beta_k |\theta_k| \right] \quad (4)$$

where the first two terms are exactly the ones at equation (2). The third term is the L1-regularization term, which is nothing more than the sum of the model's parameters multiplied by the hyper-parameter β. This objective function is also a convex one, so when the convergence point is achieved, we are guaranteed that the CRF network and its parameters are the best ones[2]. As equation (2), this objective has no closed-form solution and its maximum is also achieved by reducing the gradient of each feature to zero. Formally, we want that for each feature $f_k(X)$

$$((\sum_{m} f_k(\xi[m])) - ME_{x\sim P_\theta}[f_k(x)]) - \tfrac{1}{\beta} sign(\theta_k) = 0 \quad (5)$$

The effect of biasing many of the parameters to zero happens to integrate the network learning and the parameter optimization problems into a single objective function. Features with weight zero represent discarded edges in the network, whereas features with non-zero weights represent variables (i.e. nodes) that are connected with each other in the network. Hence, we are solving

[2] Although this solution is the optimal one, it may not be unique. Several redundant parameterizations might be at the convergence point.

both the parameter and the structure learning problems by optimizing a single objective function at a computational cost that is not much higher than learning the parameters alone.

2.5 Discriminative training

Equations (4) and (5) are used for learning a CRF model in a generative way. We may learn a CRF model in a generative way as long as each of the possible features involves at least one of the unobserved variables. Another way to train such a model is in the discriminative way. In this setting, our training set consists of pairs $D = \{(y[m], x[m])\}_{m=1}^{M}$ specifying assignments to Y (unobserved variables) and X (observed variables). Here we want to optimize the likelihood of each $y[m]$ given the corresponding $x[m]$. This is also a convex function (actually a sum of convex log-likelihood functions, i.e. one for each data sample) with a convex region of global optima, i.e. redundant optimal parameterizations. The gradient in the discriminative training case has the form:

$$\sum_{m=1}^{M} (f_i(y[m], x[m]) - \mathrm{E}_{\theta}[f_i \mid x[m]]) \tag{6}$$

Whereas in the generative training each gradient step required only a single execution of inference, training a model in the discriminative way is more cumbersome because we have to execute inference for every data sample conditioned on $x[m]$. On the other hand, the inference is executed on a simpler model, since conditioning on evidence in a Markov network can only reduce the computational costs. Inserting the L1-regularization term in the computation of the gradient of each feature leads to the form:

$$\sum_{m=1}^{M} (f_i(y[m], x[m]) - \mathrm{E}_{\theta}[f_i \mid x[m]]) - \frac{1}{\beta} sign(\theta_k) = 0 \tag{7}$$

Discriminative training is considered to be better for image classification tasks, whereas on the other hand models can be trained with fewer samples in the generative way (KOLLER & FRIEDMAN, 2009).

2.6 The Learning Procedure

As mentioned, inserting the L1-regularization term in the parameter learning optimization process allow us to also learn the CRF network structure using the same objective function and at a not much higher computational cost. In theory, all the possible features can be submitted at once to this optimization process and at convergence the features that have weight zero (the vast majority of them) are discarded and the features that have weight non-zero induce probabilistic dependencies between the variables they involve, representing in this way the CRF network. However, inserting all possible features at once in the optimization process would give rise to a highly dense network, what makes parameter inference imprecise and prohibitively costly. Because of that, we must resort to a feature introduction scheme, where the most pertinent features are introduced first in order to achieve convergence faster. As the objective function is a convex one, any feature introduction scheme leads to the global optimum, but inserting unimportant features first increases the computational costs (even if they are eventually discarded) and decreases the accuracy of the calculated gradients (LEE et al., 2007). A few

feature evaluation measures have been proposed (DELLA PIETRA et al., 1997; PERKINS et al., 2003) and can be used for deciding which feature to introduce next in the optimization process. These measures evaluate quantitatively each individual features, what increases computational costs.

The learning proceeds in the following way: (1) among all possible features, a few of them are allowed to have weight non-zero (the so-called active features), (2) we optimize the weights of these features reducing their gradients to zero (equation 7); many of the features continue to have weight zero, due to the sparseness effect of the L1-regularization term; following, (3) each feature with weight zero (the so-called inactive features) is evaluated by some quantitative measure; (4) the best feature is inserted in the active features group and steps 2 to 4 are repeated until convergence is achieved. Each time the procedure finishes step 2, the convergence criteria are checked. The criteria are that the gradients of all active features should equal zero and the gradients of all inactive features should be in module smaller than $1/2\,\beta$ (LEE et al. 2007).

3 Guiding the introduction of features

As mentioned, the global optimum is not a point representing a unique solution, but a region with possibly many equivalent optimal solutions. We want to achieve this region as fast as possible and, more specifically, we want to pick a model that represents our expert knowledge regarding the interpretation of high-resolution SAR imagery. We now show the building blocks of our framework for guiding the CRF learning towards this goal. This framework is a flexible expert system (COWELL et al. 1999) for the gradual introduction of log-linear features in the optimization of the L1-regularized log-(conditional)-likelihood objective function. The expert system is a rule-based one composed of *if... then...* rules and has as inference engine the quantitative measure proposed by DELLA PIETRA et al. (1997) for ranking the gain of inserting each of the available log-linear features in the optimization process.

3.1 Context Range Definition

As the first step, we must define the distance range of the possible contextual relations between image segments. This can be defined in Euclidean distance or as *k*-nearest-neighbours (*k*-NN). The range is then divided into distance categories and the segments in each category are labelled accordingly. In case of the *k*-NN approach, we divide the maximum *k* into *n* categories and label the segments accordingly. Figure 1 shows an example of such a categorization for the two above mentioned cases. The centroids of the segments are considered for the distance computations.

3.2 Categorization and semantic association of log-linear features

Once we defined the context's distance range, we automatically define all the possible log-linear features that can be created with this universe of image segments (given that the observed image attributes are also defined). All these available log-linear features can be categorized regarding, for instance, their number of observed variables, number of unobserved variables, number of observed and unobserved variables in it, number of observed variables in the distance (or *k*-NN) category, number of unobserved variables in the distance (or *k*-NN) category, number of variables *lengthness*, number of variables *squareness*, number of variables of *texture*, number of

variables of *intensity* etc. These categories are not mutually exclusive and they can be sub-divided into many other categories. The types and number of categories are defined by the user. After the creation of all possible features and their categorization, the semantic association of the feature categories takes place. This can be represented as a semantic net where the arrows indicate the association of log-linear feature categories. Figure 2 (a) shows an example of such a semantic net when the features are grouped into nine categories.

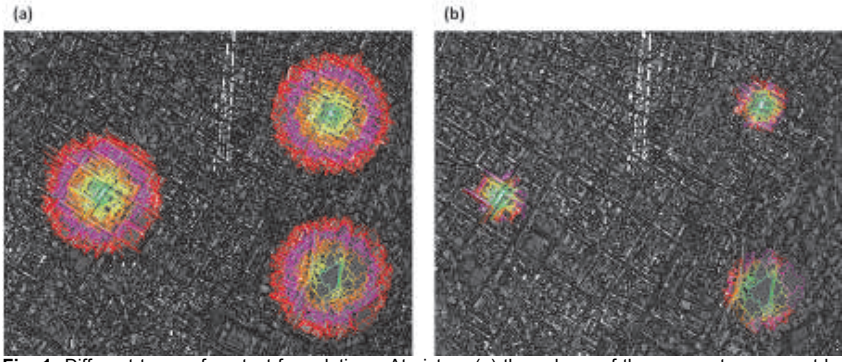


Fig. 1: Different types of context formulations. At picture (a) the colours of the segments represent how distant they are (in meters) from the central segment. At picture (b) the colours of the segment represent nearest-neighbours (NN) categories, e.g. 5-NN, 6 to 10-NN, 11 to 20-NN.

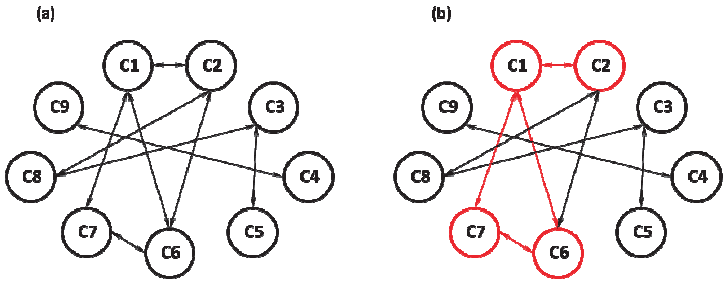


Fig. 2: Semantic net representing the association of log-linear feature categories (a) and graphical representation of the categories whose features should be evaluated next in case we insert a feature from category C1 at a certain stage of the optimization/feature introduction process (b).

Once we have a net representing the semantic association of features, we can use it to decide which ones should be evaluated at each point of the optimization process. Relying on this semantic model avoids us to evaluate each of the features each time the current model is optimized and we have to insert a new feature into the active group. At the first optimization

iteration we can evaluate all possible features and detect to which category the winning feature belongs to. Then, at the next iteration we will only evaluate features from these categories. So that the process does not get stuck on certain categories, we can at some point evaluate all features again. If the assertions and actions, i.e. if the semantic net and the rules are good, we will achieve an optimum solution soon, otherwise it will take longer. However, we will never need a longer time as we would by inserting features blithely relying on the convex property of the objective function.

3.3 Inference for computing the gradients

At the end of each optimization step we must check the converge criteria in order to know if an optimal solution has been found, what requires the computation of the gradients from the active and the inactive features. According to LEE et al. (2006), the best way to do that when using the Loopy Belief Propagation (MURPHY et al. 1999) inference algorithm to calculate the gradients (what almost always is inevitable) is to create a cluster graph which contains all the variables that the model may have. In this way, it is possible to extract clique trees out of the calibrated cluster graph and use them to infer the marginal distributions of the inactive features. The marginal distributions of the active features can be computed directly out of the calibrated cluster graph. Figure 3 shows what would be the simplest set of active log-linear features (factors of the graph) when starting the optimization process. This network structure gets altered according to what features get inserted and maintained in the active group. However, its basic structure must persist so that the gradients of all possible features can be calculated.

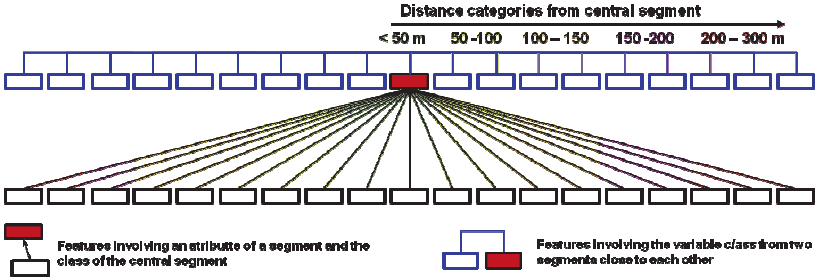


Fig. 3: Initial CRF network structure. This is the initial set of active log-linear features. Constructing a cluster-graph out of this network enables the computation of the gradient from each possible log-linear feature that can be considered with this set of variables.

4 Learning CRFs models for the classification of Urban Structure Types

Classifying Urban Structure Types (UST) (HEIDEN et al., 2012) requires considering long-range contextual relations between the observed attributes and the possible classes of the image segments. Hence, the context's distance range (section 3.1) should be defined accordingly. The mean or maximum area of the city's urban blocks may be considered for defining the range.

Also, log-linear feature categories and its semantic net might be edited so that features containing the observed variables *area*, *squareness* and *lengthness* at different range categories be privileged in the search for the optimum solution. USTs are complex and their complexity is inevitably reflected on the complexity of the features and the network. Furthermore, SAR image regions have to be classified based on its neighbouring regions classes besides its observed attributes. These conditions threat tractability. Good remedies are (1) the radical reduction of the observed variables cardinalities and (2) to avoid as much as possible over-segmentation of the SAR scene.

5 Summary and future work

In this paper we summarized how the network structure and the parameters of a CRF model can be learned in a tractable way by optimizing the L1-regularized log-likelihood objective function. We showed how the training can be carried out in a generative or discriminative way. Besides, the main elements of a rule-based system for guiding the insertion of log-linear features in the optimization process were shown. This system enables the insertion of expert knowledge in the CRF learning procedure and is expected to accelerate the convergence to a global optimal solution. In the next works we will test these hypotheses by applying this methodology for classifying UST using high-resolution space-borne InSAR data.

6 References

COWELL, R.G.; DAWID, A. P.; LAURITZEN, S.L. AND SPIEGELHALTER, D.J., 1999: Probabilistic Networks and Expert Systems. Springer Verlag: New York. P.321.

DELLA PIETRA, S.; DELLA PIETRA, V.J. & LAFFERTY, 1997: Inducing features of random fields. IEEE Transactions on Pattern Analysis and Machine Learning, **19**(4), P.380-393.

GANAPATHI V, VICKREY D, DUCHI JC, KOLLER D., 2008: Constrained Approximate Maximum Entropy Learning of Markov Random Fields. In: UAI 2008, Proceedings of the 24th Conference in Uncertainty in Artificial Intelligence, Helsinki, Finland, July 9-12, 2008: P.196-203.

HEIDEN, U., HELDENS, W., ROESSNER, S., SEGL, K., ESCH, T. & MUELLER, A., 2012: Urban structure type characterization using hyperspectral remote sensing and height information. Landscape and Urban Planning, **105**, P.361-375.

KOLLER, D. & FRIEDMAN, N., 2009: Probabilistic Graphical Models – Principles and Techniques. The MIT Press: Cambridge: USA. P.1231.

LEE S, GANAPATHI, V. & KOLLER, D., 2007: Efficient structure learning of Markov networks using L1-regularization. In: NIPS (Neural Information Processing Systems).

MURPHY, K., WEISS, Y. & JORDAN. M., 1999: Loopy belief propagation for approximate inference: an empirical study. In UAI, 1999.

PERKINS, S.; LACKER, K. & THEILER, J., 2003: Grafting: Fast, incremental feature selection by gradient descent in function space. **3**, P.1333-1356.

Point Clouds from Airborne Laser Scanning and Aerial Images for Forestry Application

IVAN SAČKOV[1] & MIROSLAV KARDOŠ[2]

Zusammenfassung: Der Beitrag befasst sich mit der Problematik von Erstellung und Qualitätsanalyse der Punkwolke aus dem Laserscanning und Luftaufnahmen (UltraCamX, UAV - Gatewing X100). Gleichzeitig präsentiert er ein paar Möglichkeiten von Anwendung der Punkwolke in der Forstwirtschaft.
Aus Sicht der Punktdichte erreichte die Punkwolke die besten Ergebnisse aus dem Luftaufnahmen (>150 b/m^2). Die vertikale Genauigkeit aller Verfahren war ungefähr gleich (RMSE <±20 cm), wobei war die Auswertung auf der Fläche ohne (oder mit nur geringer) Vegetation ausgeführt.
Als die beste Alternative für die forstwirtschaftliche Zwecke zeigte sich die Verbindung von beiden Datenquellen an die Fernerkundung der Erde. Diese Methode findet die Anwendung namentlich in der Kartographie, der Verkehrsuntersuchung, der Forstinventur und bei der Identifikation von Grenzen des Eigentums.

1 Introduction

Point cloud (PCL) is a set of data defined by the X, Y and Z coordinates, or other attributes in a three dimensional space of the coordinate system. Currently, the most used source for obtaining PCL data is the technology of laser scanning. This system works on the principle of emitting laser pulses, reception of reflected pulses and recording of the necessary information during this activity (time, position, tilt). An alternative for obtaining PCL data is digital photogrammetry. The condition is that photos have known exterior orientation parameters and capture the same area for at least 60 % of endlap and sidelap. The exterior orientation parameters can be determined with sufficient accuracy using GNSS/IMU unit located directly on the aircraft, or calculated by the bundle block adjustment process (PAVELKA 2009, CRAMER et al. 2000). Currently, there are several image matching techniques to identify identical points on the images that form the basis of dense digital elevation models generation. The most popular techniques include "Feature Based Matching" and "Least Squares Matching" (WOLF & DEWITT 2000), „Semi Global Matching“ (HIRSCHMÜLLER 2005, 2008) or „Cost Based Matching“ (INPHO 2011), whose output is a dense point cloud.
The resulting PCL, which is usually stored in a standard binary format LAS (ASPRS 2012) has a wide range of utilization. In forestry applications it is mainly in the field of forestry cartography, transportation surveys as well as determination of forest stand and tree characteristics. The advantage of laser scanning technology in these applications is the ability to record the objects that overlay. Thus it is possible to capture the space below the tree crowns. Aerial photos do not

1) Ivan Sačkov, National Forest Centre, Forest Research Institute Zvolen, Department of Forest Inventory and Management, T. G. Masaryka 22, Zvolen, 960 92, Slovakia, sackov@nlcsk.org
2) Miroslav Kardoš, Technical university in Zvolen, Faculty of Forestry, Department of Forest Management and Geodesy, T. G. Masaryka 24, Zvolen, 960 53, Slovakia, miroslav.kardos@tuzvo.sk

offer this benefit, but allow generation of point clouds with higher density as well as record a wide range of qualitative information about the surface (i.e. colour spectrum). From the data set it is then possible to filter out points that represent various object classes and create from them digital models in different data structure.
The primary aim of the paper is to examine the possibility of generation and the resulting quality of the point cloud obtained from a variety of sources of the remote sensing. The source of the data was the technology of laser scanning and aerial photos in two versions - a wide-angle camera and unmanned aerial vehicle. A secondary aim of the paper is to present several application possibilities of the generated point clouds utilization in forestry.

2 Point Clouds generation

2.1 Point Clouds generation from Airborne Laser Scanning

Data were obtained on 16/04/2012 by the Riegl Q680i laser scanner (Riegl Laser Measurement Systems GmbH). Flight parameters were set on the average flying height of 700 m, scan angle of 60˚ (FOV), average overlap of 40 % and frequency 120 Hz.
Processing of the point cloud was carried out in the relevant software of the laser scanner, which also served for the adjustment of laser strips based on the raw data and flight trajectories (relative adjustment) and then based on the ground control surfaces (absolute adjustment). Such adjusted data were exported to LAS format 2.1.

2.2 Point Clouds generation from Aerial Images

2.2.1 UltraCamX

Data were obtained on 15/09/2011 by the UltraCamX digital aerial camera (Vexcel Imaging GmbH). Both RGB and CIR photos were captured with the ground sample distance (GSD) 10 cm, 80% of endlap and 60% of sidelap. Processing of the point cloud was based on automatic aerotriangulation, which was first performed in the program MATCH-AT using bundle block adjustment and with support of the GNSS/IMU data. The result of this process were adjusted exterior orientation parameters, which were then used as input into the MATCH-T DSM to generate clouds of points using image matching technique "Cost-based matching." This technique ensured final generation of the dense point clouds for every third pixel (UCX_3px), or when editing parameters for each pixel of the aerial photo raster (UCX_1px).

2.2.2 Unmanned Aerial Vehicles

Data were obtained on 08/08/2012 by the Gatewing X100 unmanned aerial vehicle (Trimble). Flight parameters were set on the average flying height of 130 m, scan angle of 65˚ (FOV), average overlap of 75 %. Point cloud processing was done through software Gatewing Stretchout (Trimble). This process was fully automated, and it was only necessary to select the desired quality of the resulting point clouds in LAS format 1.2.

2.3 Elevation models generation from Point Clouds

A digital elevation model (DEM) is a representation of a continuous Earth surface through a large number of selected points with known coordinates in any coordinate system. Point cloud is therefore an ideal alternative for creating digital models of all types.
Creating digital models from point geo-objects P(x, y, h) was carried out through interpolation methods in ArcGIS 10 for Desktop environment (ESRI). The results were digital models in a raster data structure and a spatial resolution of 0.5 m.

3 Geometric quality of Point Clouds

3.1 Point Clouds density

The density of points per surface unit (Pt/m^2) is the primary characteristic of a point cloud, which, together with the method of classification, filtering or interpolation of points significantly influences the overall geometric quality (SKOLAUD 2012).
Presentation of achieved point density from analysed sources is shown in tab. 1. Most dense point cloud was acquired by the UAV technology - Gatewing X100, which reached 710 pt/m^2. Using the digital camera UltraCamX, the density reached 8.2 pt/m^2 (UCX_3pix: point clouds generated for every third pixel) to 100 pt/m^2 (UCX_1pix: interpolation of each point from the point cloud generated for every third pixel). Fewer points were acquired by the airborne laser scanning (ALS), where the density reached 7.6 pt/m^2.

Tab. 1: Point Clouds statistics

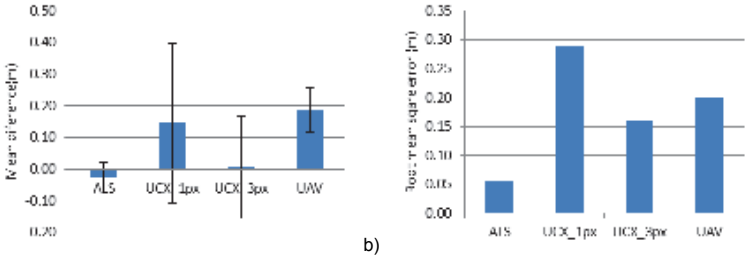
Technology	Pt_Count*	Pt/m^2
UAV	7 100 000	710
UCX_1pix	996 000	99,6
UCX_3pix	82 000	8,2
ALS	76 000	7,6

UAV – Unmanned aerial vehicle (Gatewing X100); UCX - UltraCamX; ALS – Airborne laser scanner;
*Point count/10 000m^2

3.2 Vertical accuracy

Overview of the various differences between the elevation values extracted from digital elevation model created by the inverse distance interpolation method and elevation achieved from the terrestrial survey is shown in fig. 1.
The average difference, shown in the histogram in fig. 1a, describes the systematic components of error and the results of the test of statistical significance are shown in tab. 2. From the results we can see, that the laser scanning systematically undervalued the terrain elevation ($\alpha = 0.05$). Aerial photos on the other hand, reached positive height differences of terrain height, which in the case of the UAV and UCX_1pix were with the 95 % reliability statistically significant. The lowest mean difference (+1 cm) was found in the case UCX_3pix, which also contained systematic error component.
Mean error, shown by range in fig. 1a, describes the variability of deviations around the mean difference. It means that 68 % of differences in the case of the ALS is in the range from -2 to 8

cm, for the UCX_1pix in the range from -10 to 40 cm, for the UCX_3pix in the range from -15 to 17 cm and for the UAV in the range from 12 to 26 cm.
The mean square error, shown by the histogram in fig. 1b describes the total deviation with a certain probability. It is valid that with 68 % reliability, the difference between the estimated terrain height and the real value will be found in the range of ± 6 cm (ALS) to ± 29 cm (UCX_1pix).



a) b)

Fig. 1: Vertical accuracy of digital elevation model a) Mean difference; Standard error; b) Root mean square error;
UAV – Unmanned aerial vehicle (Gatewing X100); UCX - UltraCamX; ALS – Airborne laser scanner;

Tab. 2: Statistical test

ALS	UCX_1px	UCX_3px	UAV
Normality test of diferences			
0,182*	<0,001	<0,001	**0,538***
Significance test of differences			
<0,001	<0,001	**0,730***	<0,001

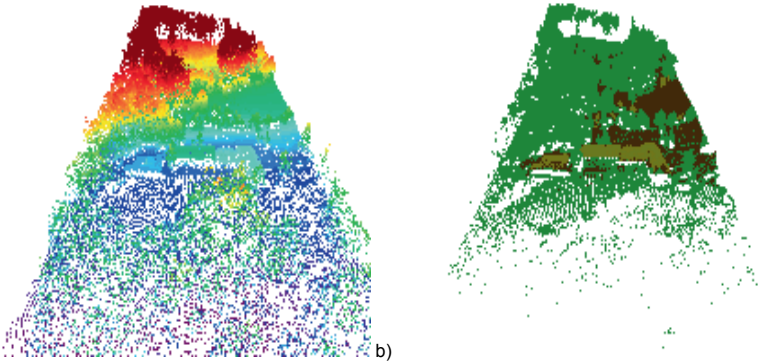
UAV – Unmanned aerial vehicle (Gatewing X100); UCX - UltraCamX; ALS – Airborne laser scanner;
*α=0,05

4 Forestry Application

4.1 Cartography

Visualization of the part of the Earth surface into the map plane has particular importance in the forestry and cartography is now a common part of it. Within the wide range of its products dominate digital aerial photos with the ground sample distance of 0.2 m in CIR and RGB colour representation, which are used for example in the determination of forest land ownership boundaries (ŽÍHLAVNÍK 2007, 2008), assessment of forest ecosystem health status as well as forest management planning.
Automated delineation of forest land boundaries based on a point clouds was carried out on the area of interest with an area of 750 ha. The source of the data was aerial laser scanning, and validation of the results was verified using aerial photos. Processed and filtered point cloud was classified and filtered into point classes ground, vegetation and buildings. This activity was

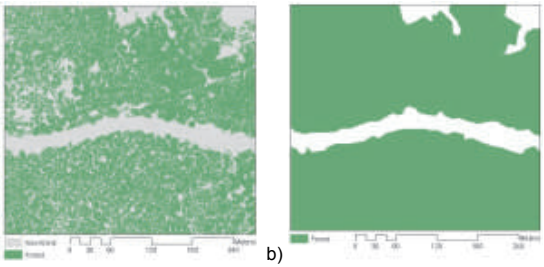
carried out by the progressive morphological filtering in the DTMaster (INPHO) environment and the result is presented in fig. 2.



a) b)

Fig. 2: Point cloud from ALS: a) Point Cloud with attribute „z"; b) Point Cloud with attribute „class" Brown color – Ground; Green color – Vegetation; Ocher color – Building;

From the classified point cloud was possible to create a raster or vector representations of forest and non-forest landscape (fig. 3). It is possible to enter the conditions according to standard definitions of forest (i.e. FAO considers as a forest land with the canopy greater than 10%, an area greater than 0.5 hectares and a height above 5 m). An example of the use of simple restrictive rule is a figure 3b, where as a forest is considered land with forest stand of an area greater than 0.5 hectares and a height more than 5 m.



a) b)

Fig. 3: Forest mask: a) Grid representation – vegetation (green color); b) Polygon representation – forest (green color);

4.2 Forest roads mapping

Level of the forest stands accessibility significantly determines the forestry transport logistics, the concept of fire and flood protection as well as the general public land use. Forest compartments are available mainly through forest roads, which are thus major aboveground line objects in the forest. The primary way to gather information about the state and level of the

transportation system is the field transportation survey, which is costly and time consuming. An alternative is therefore to present transportation survey carried out on the remote sensing data, while the best source of data for this purpose can be considered airborne laser scanning.
Use of the point cloud from airborne laser scanning for transportation survey purposes was tested on the model area with an area of 150 ha. From recorded points were created three alternative terrain models with a resolution of 50 cm, which served as a reference layer for manual digitization of forest roads. The first option was raster created by interpolation of the multi-point objects by inverse distance method (IDW) with attributes of mean terrain elevations. The second alternative was a raster, which represented the average slope of the terrain based on a percentage and it came out from the previous grid. The third option was raster, which was created by aggregation of the values of all points that fall on a certain grid cell. Aggregated attribute was the value of the maximum intensity of the signal impulse.
The terrain model with attribute values containing the height and slope showed as the optimal terrain model for manual identification and digitization of line objects (fig. 4a), which allowed to clearly identify (67 %) the majority of roads with the help of created cross profiles (fig. 4b).

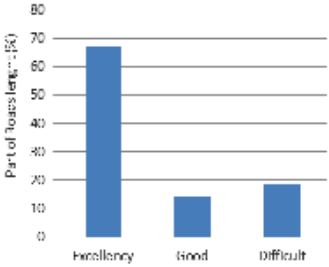
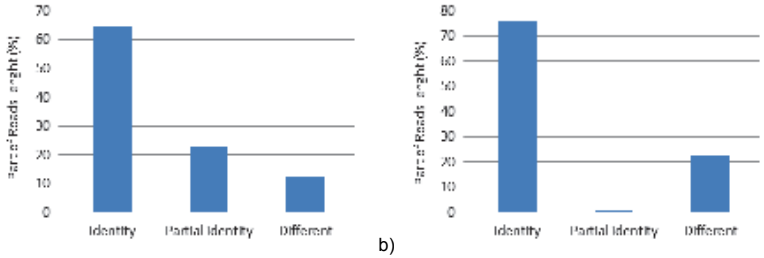
a)  b) 

Fig. 4: Qualitative characteristics of the road network in the study area acquired from ALS data:
a) Digital terrain model; b) The exactness of identification (Excellent – road easily recognisable within DTM; Good – road sufficiently recognisable within DTM; Difficult – road not easily recognisable within DTM); ALS – Airborne laser scanning; DTM – Digital terrain model;

In the area of interest from all permanent roads registered during the field survey, only about the 12 % of them were not registered by the contactless survey (fig. 5a). The roads were used only for skidding, situated under the dense forest vegetation and had only natural surface in poor condition and had not built a system of drainage, slope embankments, or excavations. At the same time about 23% of the roads were registered by the contactless method, which were not registered during the field survey of the permanent forest transportation network (fig. 5b). They were usually a temporary forest and field roads, which are characteristic by not strengthened surface without drainage elements and all were situated under the tree crowns.



a) b)

Fig. 5: Accuracy of identification: a) Terrain data vs. ALS data; b) ALS data vs. terrain data
Identity – the entire road is identified in both data sources, in either the same or a different position; Partial identity – part of the road is identified in both data sources, in either the same or a different position; Different – the road is not found in both data sources
ALS – Airborne laser scanning;

4.3 Forest inventory

Current and representative information about forest and landscape are the basis for the professional management of forests, as well as for effective nature protection or urban planning. At present, although there is a comprehensive set of methods for the determination of forest state, the experience indicates that common taxation survey systematically underestimates some of the data from a variety of causes, and the standard procedures are labour and cost quite challenging.

From these reasons, it is now necessary to look for the alternative solutions. One of the possibilities is obtaining new data about the state of the forest through remote sensing technology in combination of aerial photography and laser scanning. This alternative is in the Slovakia only in the research and so own applications and results are not available yet. Conclusions from abroad show the potential of such solution. They show, for example, the tree height can be determined with an accuracy of ± 0.2 m to ± 1.6 m (FRANKEN 2005). The thickness of the trees, according to some authors, can be determined with an accuracy of ± 3 to ± 7 cm (BOTCHER 2003, HOLOPAINEN et al. 2011). The possibility of estimating forest stand volume has been proved, where have been achieved results within the precision ± 6-35% (VAUHKONEN et al., 2009 TONOLLI et al. 2010). The research also showed that the accuracy of individual characteristics is largely conditioned by the vertical and horizontal variability of forest, terrain slope and density of the recorded pulses. It meant, for example, that in the forest stands with a complex structure can be identified only approximately 38-44 % of trees (HEURICH et al. 2004, MOFFAT et al. 2005, MONNET, 2010). Furthermore, this meant that the height of the tree was in the range of 1-5 m generally underestimated (CROAK 2007, IOKI 2009, HOLLAUS 2011), while the increasing slope caused the opposite trend of the systematic error.

5 Conclusion

For the purpose of the point clouds generation can be used data from laser scanning and aerial photography. While during the scanning is the point cloud basic data output, from aerial photos it is needed to create point clouds additionally. With the necessary technical infrastructure and convenient aerial photos this process doesn't represent significant additional problems. For

example, software for the UAV technology Gatewing X100 allows performing this process fully automatically.
The resulting point cloud from the alternative data sources reach qualitatively different attributes. The evaluation of the density of points in the point cloud showed as the best UAV technology Gatewing X100 (> 700 pt/m^2). Significantly fewer points were obtained from aerial photos acquired by the UltraCamX camera (about 8-100 pt/m^2) and then from laser scanning data (about 8 pt/m^2). It was also found that in the area of interest with the 68 % reliability the difference of terrain height achieved by the remote sensing techniques from it's real value was situated within the interval of ± 6 cm (ALS) from ± 16 to ± 29 cm (UCX) and ± 20 cm (UAV).
The use of point cloud, or products based on it, have a wide application in forestry. In Slovak conditions, until now there are proved applications in the field of high-quality digital terrain models (DTM) and surface models (DSM), in mapping of the boundaries of forest lands as well as the identification of line structures in forest and non-forest environment. Also currently runs research of the remote sensing data utilization within the forest state determination. In general, the data from laser scanning and high resolution digital aerial photography have been used in forestry for relatively short period and therefore it is necessary to ensure continuous research of possibilities and limits of their application in Slovak conditions.

Acknowledgement
This publication is the result of the implementation of the project: Centre of Excellence "Decision support in forest and country", ITMS: 26220120069, supported by the Research & Development Operational Programme funded by the ERDF.
Special thanks belong to SGS Holding a. s. company for lending Gatewing X100 unmanned aerial vehicle technology.

6 References

ASPRS, 2012: LAS Specification version 1.4-R12. The American Society for Photogrammetry & Remote Sensing. 27 S.

BÖTTCHER, K., KLEINN, C., 2003: Using Crown Surface Models for the Extraction of Forest Inventory Information: First Experiences With Study Areas in State of Mecklenburg-Vorpommern (Germany). In Proceedings of ScandLaser Scientific Workshop on Airborne Laser Scanning of Forest.

CRAMER, M., STALLMANN, D., HAALA, N., 2000: Direct georeferencing using GPS/Inertial exterior orientations for photogrammetric applications. International Archives of Photogrammetry and Remote Sensing, V. 33, part B3, S. 198-205.

FRANKEN, F., 2005: Neue Techniken in der Walddatenerhebung. LÖBF NRW. 6 S.

HIRSCHMÜLLER, H., 2005: Accurate and Efficient Stereo Proces-sing by Semi-Global Matching and Mutual Information. Proc. IEEE Conference on CVPR, New York, New York.

HIRSCHMÜLLER, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 30, No. 2.

HEURICH, M., GÜNTHER, S., SCHRÖDER, S., 2004: Baumhöhenmessung mit flugzeuggetragenen Laserscannern. In Allgemeine Forst und Jagdzeitung. ISSN 0002-5852, vol. 59, nu. 17, S. 945-947.

HOLLAUS, M., EYSN, L., BAUERHANSL, CH., RICCABONA, F., MAIER, B., 2011: Accuracy assessment of ALS-derived stem volume and biomass maps. In Proceedings of the EARSeL SIG Forestry workshop.

HOLOPAINEN, M., 2011: Effect of airborne laser scanning accuracy on forest stock and yield estimates: doctoral dissertations. Supervisor K. Viitanen. Helsinky: Department of Surveying Aalto university. 160 S.

IOKI, K., IMANISHI, J., SASAKI, T., 2009: Estimating stand volume in broad-leaved forest using discrete-return LiDAR: plot-based approach. In Landscape and Ecological Enginnering. ISSN 1860-1888, vol. 6, no. 1, S. 29-36.

KVAK, D. A., LEE, W. K., LEE, J. H., 2007: Detection individual trees and estimation of tree height using LiDAR data. In Journal of Forest Reserch. ISSN 1610-7403, vol. 12, no. 6, S. 425-434.

MOFFIET, T., MENGERSEN, K., WITTE, C., 2005: Airborne laser scanning: Exploratory data analysis indicates potential variables for classification of individual trees or forest stands accordings to species. In ISPRS Journal of Photogrammetry and Remote Sensing. ISSN 0924-2716, vol. 59, S. 289-309.

MONNET, J.-M., MERMIN, E., CHANUSSOT, J., 2010: Estimation of forestry parameters in mountainous coppice stands using airborne laser scanning. In Proceedings of the 10th International Conference on LiDAR Aplications for Assessing Forest Ecosystems, SilviLaser 2010.

PAVELKA, K., 2009: Fotogrammetrie 1. 1st edition. Prague: Czech technical university in Prague, ISBN : 978-80-01-04249-6.

SKOLAUD, J., SCHAER, P., 2012: Automated Assesment of Digital Terrain Models Derived From Airborne Laser Scanning. In Photogrammetrie Fernerkundung Geoinformation. ISSN 1432-8364, Heft 2, S. 105-114.

TONOLLI, S., DALPONTE, M., VESCOVO, L., 2010: Mapping and modeling forest tree volume using forest inventory and airborne laser scanning. In European Journal of Forest Reserch. ISSN 1612-4677, vol. 130, no. 4, S. 569-577.

VAUHKONEN, J., TOKOLA, T., MALTAMO, M., 2009: Applied 3D texture features in ALS-based forest inventory. In European Journal of Forest Reserch. ISSN 1612-4677, vol. 129, no. 5, p. 803-811.

WOLF, PAUL. R & DEWITT, BON. A., 2000: Elements of photogrammetry with application in GIS. 3rd edition, McGraw-Hill, Boston, 608 S., ISBN 0-07-292454-3.

ŽÍHLAVNÍK, A., 2007: Current changes and new features within the spatial arrangement of forest. In. Current issues in forest mapping. Technical University Zvolen, S. 125-133.

ŽÍHLAVNÍK, A., 2008: Forest management. Technical University Zvolen, ISBN 80-228-1542-X.

3D-Gebäudeaufnahme mittels UAV-basierter Luftbilder

MARTIN SAUERBIER

Zusammenfassung: Ziel des Projektes war die Evaluierung der Auswertung, insbesondere der Orientierung, von UAV-Bildern zur Gebäudemodellierung mit ERDAS LPS. Im Besonderen ging es darum, die gemeinsame Orientierung von Senkrecht- und Schrägaufnahmen zu untersuchen und in Bezug auf die erreichte Genauigkeit mit Hilfe bekannter Referenzstrecken zu beurteilen. Weiterhin können aus den Erfahrungen, die bei der Auswertung der vorliegenden Bilder gemacht wurden, Rückschlüsse auf mögliche Verbesserungen für zukünftige, ähnlich gelagerte Projekte gezogen werden.

1 Einleitung

Im Rahmen der hier beschriebenen Arbeiten sollen die Bilder einer UAV-Befliegung mit kombinierten Senkrecht- und Schrägaufnahmen des General Herzog Hauses (GHH) in Thun mit ERDAS LPS ausgewertet werden, um festzustellen, ob und mit welcher Genauigkeit sich solche Bilddaten zusammen orientieren lassen. Die Kombination von Senkrecht- und Schrägaufnahmen erlaubt eine bessere stereoskopische Erfassung der Gebäudefassaden im Vergleich zu reinen Senkrechtbefliegungen. Bisher wurde für die Auswertung solcher Bildblöcke zumeist auf Software zurückgegriffen, die für den photogrammetrischen Nahbereich konzipiert ist (PÜSCHEL et al., 2008), daher sollte die Auswertung von Schräg- und Senkrechtaufnahmen mittels der ursprünglich für Senkrechtaufnahmen konzipierten Software ERDAS LPS untersucht werden.

Zum Zweck der Bildorientierung wurden weiterhin ein Satz von Passpunkten, der mittels GPS eingemessen wurde, sowie mit dem Messband aufgenommene Referenzstrecken verwendet.

1.1 Flugplanung und UAV-Befliegung

Bei dem für die Befliegung eingesetzten UAV handelt es sich um das Modell microdrones MD4-1000, welches in der Lage ist, eine maximale Nutzlast von 1,2 kg zu tragen, und daher für den Einsatz mit Spiegelreflex- bzw. Systemkameras geeignet ist. Für die Befliegung des GHH wählten wir die Sony NEX-5N mit einem 16mm-Objektiv mit fester Brennweite aus. Die Spezifikationen der Sony NEX-5N finden sich in Anhang C. Wichtig war hierbei, eine Kamera zu wählen, bei der der Autofokus abschaltbar und sämtliche Einstellungen manuell durchzuführen waren. Auf Basis der gewählten Kamera wurde vorgängig die photogrammetrische Flugplanung mit Hilfe der Software mdCockpit Version 3 durchgeführt. Aufgrund der Nähe zum Flugplatz Thun sollte nicht höher als 80 m geflogen werden, wir planten daher die Aufnahme zweier Bildblöcke mit folgenden Eigenschaften:

- Block 1: dieser Block deckt das GHH ab und wurde mit einer Höhe von 60m über Grund geplant (Abb. 2).
- Block 2: dieser Block deckt den an das GHH angrenzenden Pavillon ab und wurde mit einer Höhe von 40m über Grund beflogen (Abb. 1).

Aufgrund der unterschiedlichen Flughöhe beider Blöcke erreichten wir in etwa die gleiche geometrische Auflösung der Bilder für die jeweiligen Gebäudedächer.

Die Flüge wurden traditionell photogrammetrisch geplant, d.h. das beide Gebäude mäanderförmig in Streifen mit 80% Längs- und Querüberlappung in annähernder West-Ost-Richtung abgeflogen wurden. Die Besonderheit bestand darin, dass die jeweils äusseren zwei Flugstreifen sowie die ersten beiden und die letzen beiden Bilder jedes Streifens als Schrägaufnahmen mit ca. 50° (innen) bzw. ca. 60° (ganz aussen) Neigung in Bezug auf die Horizontale und mit Blickrichtung (Azimuth) senkrecht zum Gebäude aufgenommen wurden. Die Koordinaten und Aufnahmewinkel finden sich im Report zur Aerotriangulation (Anhang A und B). Durch die hohe Längs- und Querüberlappung war ausserdem gegeben, dass jeder der auf den Gebäuden gelegenen Passpunkte in mindestens 4 Bildern sichtbar ist.

Die Befliegung musste aufgrund begrenzter Akku-Kapazität in 3 Einzelflügen absolviert werden.

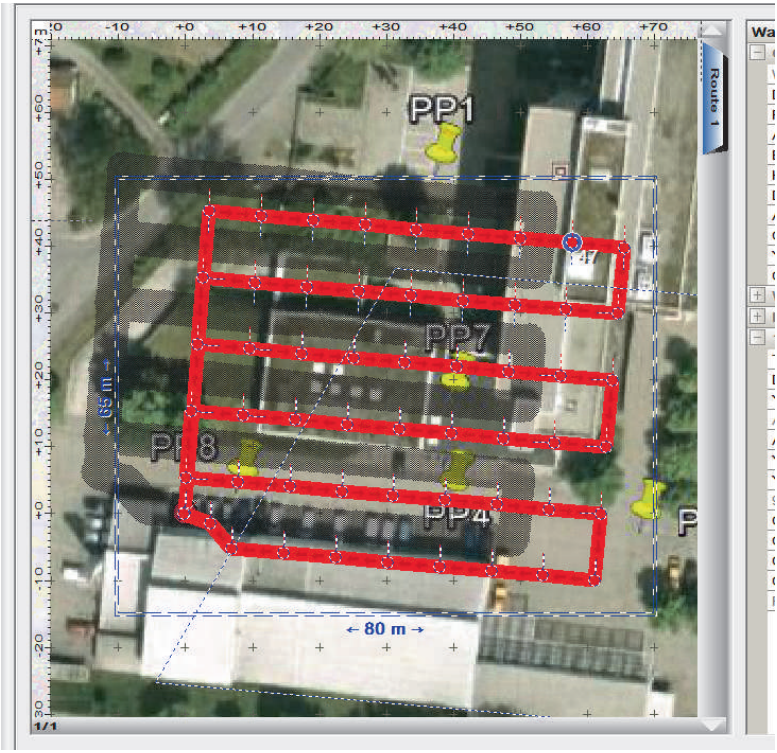


Abbildung 1: Flugplan zur Befliegung des Pavillons in MDCockpit, hier sind exemplarisch einige Passpunkte eingetragen.

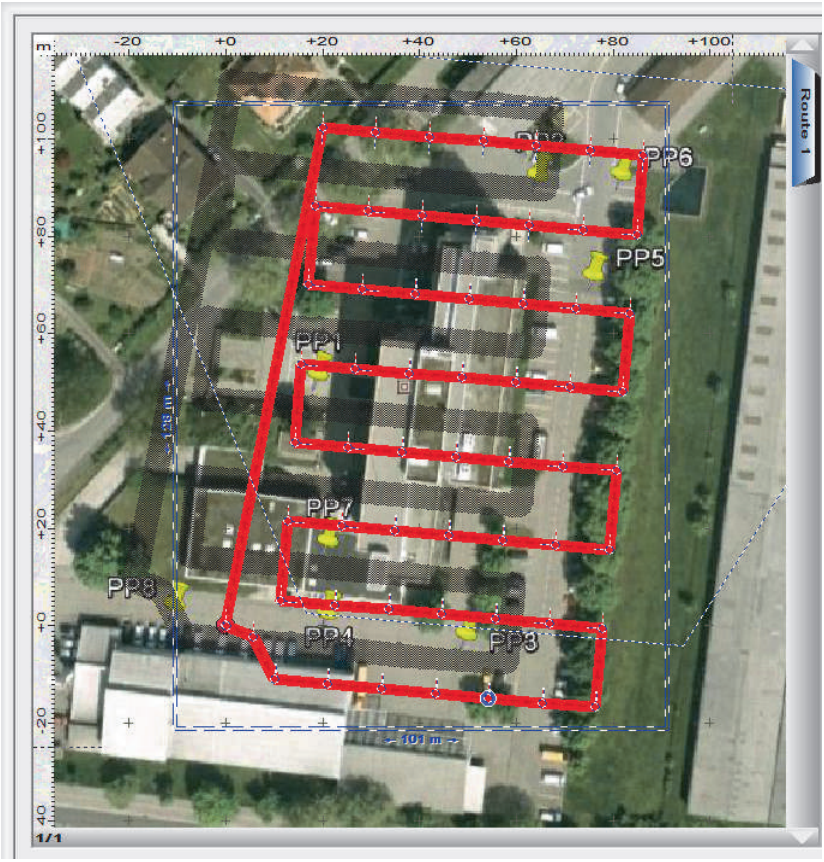


Abbildung 2: Flugplan zur Befliegung des GHH in MDCockpit. Auch hier finden sich exemplarisch einige Passpunkte.

1.2 Kamerakalibrierung

Die Kalibrierung der Kamera wurde zum Einen vor der Befliegung sowie direkt im Anschluss an die Bildflüge mittels Aufnahme eines 3D-Testfelds und Auswertung mit der Software FODIS Measure 3D (PhotoCore, www.photocore.ch) durchgeführt, um eventuelle Änderungen der Kameraparameter aufdecken zu können. Die Kamera wurde zwischen den Testfeldaufnahmen und den Bildflügen weder ausgeschaltet noch wurden Blende oder Fokussierung geändert. Die ermittelten Kameraparameter konnten in LPS importiert und für die Bildauswertung verwendet werden. Tabelle 1 zeigt die in LPS verwendeten Parameter.

Tabelle 1: Kameraparameter nach Kalibrierung.

Parameter	Unit	Fix/Free	Given Value	Given Stdev.	Estimated Value	Estimated Stdev.	Total Correction	Sig. Ratio
c	mm	Free	16	1.00E+30	15.8963	0.00045349	-0.10369	228.65
x0	mm	Free	0	1.00E+30	-0.127863	0.00055391	-0.127863	230.836
y0	mm	Free	0	1.00E+30	0.19188	0.00047026	0.19188	408.027
Sx		Free	0	1.00E+30	-0.00332411	5.80E-06	-0.00332411	573.539
a		Free	0	1.00E+30	-9.31E-05	2.80E-06	-9.31E-05	33.3162
k1	mm^-2	Free	0	1.00E+30	-0.00026088	7.39E-07	-0.00026088	352.791
k2	mm^-4	Free	0	1.00E+30	1.59E-06	9.23E-09	1.59E-06	172.314
k3	mm^-6	Free	0	1.00E+30	-2.19E-10	3.46E-11	-2.19E-10	6.33436
p1	mm^-2	Free	0	1.00E+30	8.90E-05	8.30E-07	8.90E-05	107.238
p2	mm^-2	Free	0	1.00E+30	-1.99E-05	7.04E-07	-1.99E-05	28.205

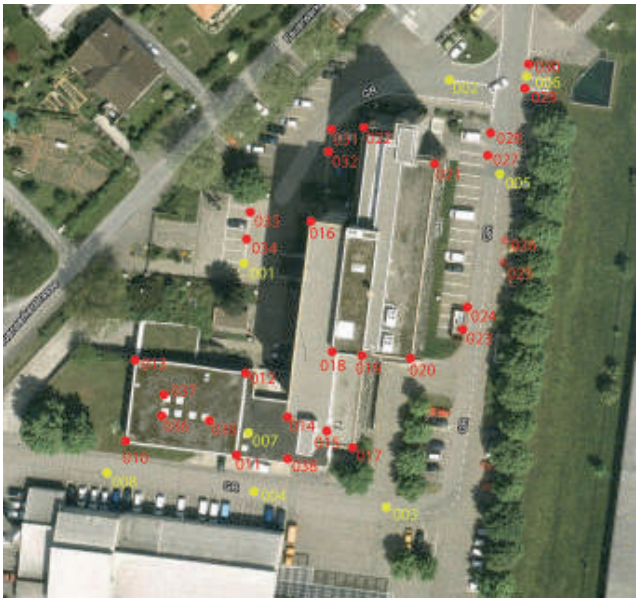


Abbildung 3: Verteilung der Passpunkte im Untersuchungsgebiet.

1.3 Passpunkte

Als Passpunkte dienen überwiegend Gebäudeecken und Fahrbahnmarkierungen, die weitestgehend gleichmässig über das Projektgebiet verteilt sind. Zusätzlich wurden einige Punkte auf der angrenzenden Strasse signalisiert. Abbildung 3 verdeutlicht die Lage der Passpunkte im Untersuchungsgebiet.

2 Bildorientierung mittels Aerotriangulation

Zur Orientierung der UAV-Bilder wurde folgendermassen vorgegangen:

1. Manuelle Messung der Passpunkte in möglichst allen Bildern. Teilweise konnten die Passpunkte nach manueller Messung in einigen Bildern automatisch mittels Bildkorrelation übertragen werden.
2. Zusätzliche Messung von Verknüpfungspunkten (automatisch und manuell) sowie deren Übetragung in alle weiteren Bilder, in denen sie sichtbar waren.
3. Berechnung der Bildorientierungen und der Landeskoordinaten der Verknüpfungspunkte durch Bündelausgleichung.

In beiden Blöcken wurden nicht alle aufgenommenen Bilder orientiert, sondern es erfolgte eine Auswahl, so dass jeweils die Gebäude rundherum von Schrägaufnahmen stereoskopisch abgedeckt waren und dazu wurden jeweils alle Senkrechtaufnahmen orientiert.

Bei der Messung der Verknüpfungspunkte zeigte sich, dass die automatische Übertragung um so besser funktionierte, je ähnlicher die horizontale Ausrichtung der Kamera bei den zu korrelierenden Bildern ist. Automatisch generierte Fehlmessungen wurden durch manuelle Messungen korrigiert. Bei der Berechnung der Bündelausgleichung wurden die Parameter der Kamera und die Koordinaten der Passpunkte fix gesetzt. Auf die erreichten Genauigkeiten wird in Abschnitt 4 eingegangen.

3 Messung von Referenzstrecken

Für Vergleichszwecke wurden einige Strecken in Stereomodellen gemessen, zum Teil zwischen Passpunkten, insbesondere Gebäudekanten, und zum Teil die ausgelegten Paletten. Tabelle 2 zeigt die gemessenen Strecken in dem Bildblock, der den Pavillon abdeckt, Tabelle 3 die gemessenen Strecken im Bereich des GHH-Bildblocks.

Tabelle 2: Soll-Ist-Vergleich der gemessenen Referenzstrecken am Pavillon.

Strecke	Ist [m]	Soll [m]	Δs [m]
PP4 – PP8	32.423	32.468	0.045
PP10 – PP11	24.193	24.342	0.149
PP10 – PP13	18.357	18.36	0.003
PP11 – PP12	18.397	18.373	-0.024
PP12 – PP13	24.254	24.356	0.102

Die in Abbildung 4 eingezeichneten Paletten 1 und 2 wiesen folgende mit dem Messband bestimmte Masse auf: 1.234 m x 0.922 m für Palette 1 und 1.271 m x 0.881m für Palette 2.

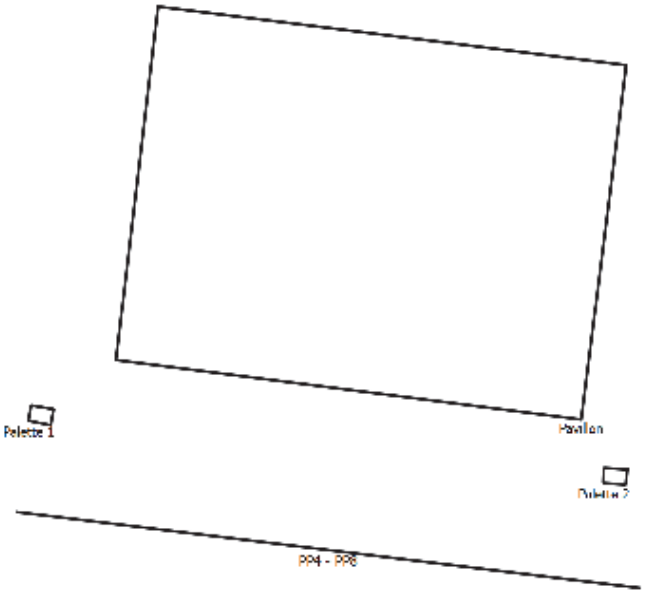


Abbildung 4: Gemessene Referenzstrecken.im Bildblock Pavillon.

Tabelle 3: Soll-Ist-Vergleich der gemessenen Referenzstrecken am GHH.

Strecke	Ist [m]	Soll [m]	Δs [m]
PP14 – PP16	42.235	42.151	-0.084
PP18 – PP19	6.79	6.716	-0.074
PP20 – PP21	41.954	42.148	0.194

4 Fazit

Die erzielten Resultate zeigen, dass die theoretisch zu erwartende Genauigkeit von ca. 1 Pixel im Bildraum, in unserem Fall entsprechend ca. 2 cm im Objektraum, nicht erreicht wurde.

Anhand der Referenzstreckenmessungen lässt sich erkennen, dass insbesondere in Ost-West-

Richtung noch ein Einfluss dieser Unsicherheit aufzutreten scheint (vergleiche Strecken PP10-PP11 und PP12-PP13 in Tabelle 2). Auch die Restfehler in den Passpunkten nach den Bündelausgleichungen treten vor allem in Ost-West (X-) Richtung auf, was ebenfalls auf eine entsprechende Unsicherheit in den GPS-Messungen oder auf Fehler in Punktdefinitionen bzw. Bildmessungen der Passpunkte hindeuten könnte. Die Abweichungen in den Referenzstrecken am GHH deuten ausserdem auf einen Fehler im Massstab hin. Dieser Effekt soll in künftigen Arbeiten noch genauer untersucht werden, was im Rahmen dieses Projektes aus zeitgründen nicht mehr möglich war.

Daher liegen auch die mit LPS ermittelten Genauigkeiten für die Bildorientierungen und die gemessenen Referenzstrecken in dieser Grössenordnung. Wenn nur Senkrechtaufnahmen genutzt werden, wird eine Genauigkeit der Bildorientierung von 2 cm im Objektraum erzielt, was den theoretischen Erwartungen entsprechen würde. Es wird ausserdem deutlich, dass sich durch die Hinzunahme von Schrägaufnahmen die Genauigkeit verringert, in unserem Fall verschlechtert sie sich um einen Faktor von ca. 2 für den Block Pavillon und ca. 3 für den Block GHH. Dieser Effekt dürfte auf die erforderliche Messung von Verknüpfungspunkten auch in von der Kamera weit entfernten Bereichen und der damit einhergehenden Unsicherheit in der Punktdefinition sowie einem ungünstigen Blickwinkel auf die jeweiligen Punkte begründet liegen. Er ist ausserdem umso grösser, je stärker der Aufnahmewinkel von der Vertikalrichtung abweicht und wirkst sich auch in den Stereomodellen mit Schrägaufnahmen in Form einer relativ grossen Restparallaxe in den Randbereichen aus. Abhilfe könnten Passpunkte, die durch kugelförmige Targets signalisiert werden, schaffen, die aus allen Richtungen anzielbar sind und in den weiter entfernten Bereichen (z.B. an der Allmendstrasse) platziert werden.

Da in der Bündelausgleichung alle Bildmessungen in die Bildorientierungen aller Bilder eingehen, beeinflusst der obengenannte Effekt auch die Messung der Referenzstrecken im Bereich der Gebäude.

Grundsätzlich konnte gezeigt werden, dass die Orientierung kombinierter Schräg- und Senkrechtaufnahmen in ERDAS LPS möglich ist. Die vergleichsweise geringe Genauigkeit, die erzielt wurde, kann im Wesentlichen auf Unsicherheiten in der Punktmessung und -definition von Verknüpfungspunkten zurückgeführt werden, was sich auch durch eine vergleichsweise mit FODIS durchgeführte Bildorientierung bestätigte, mit der ähnliche Resultate erzielt wurden.

5 Literaturverzeichnis - Überschrift Ebene 1

PÜSCHEL, H., SAUERBIER, M. & EISENBEISS, H., 2008: A 3D-Model of Castle Landenberg (CH) from combined photogrammetric processing of terrestrial and UAV-based images. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVII, part B6b, pp. 93–98.

Abbildende und nichtabbildende Geländespektrometrie zur Untersuchung von Stressphänomenen an Buchenpflanzen

OKSANA STERN[1]

Im Zuge des Klimawandels wird sich der Einfluss von Hitze, Trockenheit, Frost und Nährstoffmangel auf Waldökosysteme verschärfen. Somit nimmt die Bedeutung quantitativer Monitoringmethoden betreffend biochemischer und biophysikalischer Parameter in der Vegetationsanalyse zu. Für die Fernerkundungssensoren setzt das höhere spektrale und räumliche Auflösungen der Sensoren voraus. Diese Arbeit soll das Potenzial von spektral hochauflösenden, abbildenden und nichtabbildenden optischen Sensoren zur Erfassung von Vegetationsparametern während eines induzierten Trockenstresses untersuchen. Hierzu wurde im Jahr 2011, ein sechswöchiges Austrocknungsexperiment an jungen Buchenpflanzen durchgeführt. Aus der Auswertung von Bilddaten, den spektralen Reflexionssignaturen und daraus abgeleiteten Vegetationsindizes sollen Rückschlüsse auf den allgemeinen pflanzenphysiologischen Zustand sowie auf die Veränderung im Chlorophyll- und Blattwassergehalt der Versuchspflanzen gezogen werden.

1 Einführung

Die Klimaveränderung stellt Risiken für die Wälder und deren nachhaltige multifunktionale Nutzung dar (BOLTE et al., 2009). Das häufigere Auftreten von schweren Hitze- und Dürreperioden als Folge des Klimawandels kann Artenzusammensetzungen und Struktur der Wälder in vielen Regionen verändern (ALLEN et al., 2010). Die fernerkundlich erhobenen Daten können in diesem Kontext eine umfassende Informationsgrundlage liefern. Diese Arbeit soll das bisher wenig untersuchte Potenzial von spektral hochauflösenden, abbildenden optischen Sensoren zur Erfassung von biochemischen und biophysikalischen Vegetationsparametern während eines induzierten Trockenstresses untersuchen. Nach aktueller Stand der Forschung, führt der Wassermangel einerseits zur Dehydratisierung der Zelle, was sich schädigend für Enzyme und Funktionsstrukturen wie z. B. Membranen auswirkt und andererseits behindert der durch Wassermangel bewirkte Stomataverschluss die CO2-Zufuhr, was für den Prozess Photosynthese essenziell ist (SITTE, 2002). In der vorliegenden Arbeit soll diese Veränderung der inneren Blattstruktur, sowie Abnahme der Chlorophyllgehaltes, was mit der photosynthetischen Aktivität der Pflanzen verknüpft ist, mittels eines nichtabbildendes ASD FieldSpec-Spektrometers auf Blattebene und einer Hyperspektralkamera HySpex auf Individuenebene aufgezeigt werden. Aus der Auswertung von Bilddaten, den spektralen Reflexionssignaturen und daraus abgeleiteten Vegetationsindizes sollen Rückschlüsse auf den allgemeinen pflanzenphysiologischen Zustand, sowie auf die Veränderung im Chlorophyll- und Blattwassergehalt der Versuchspflanzen gezogen werden. Im Zentrum der Arbeiten steht eine kritische Analyse von ausgewählten Vegetationsindizes auf ihre Eignung für die Trockenstressanalyse und ihre Sensitivität gegenüber Änderungen von Blattwasser- und Chlorophyllgehalten.

1) Oksana Stern, Umweltfernerkundung und Geoinformatik, Universität Trier, Behringstraße 21, 54286 Trier; shte6b01@uni-trier.de

2 Material und Methoden

Im Jahr 2011, vom 23. August bis 27. September, wurde ein sechswöchiges Austrocknungsexperiment an jungen Buchenpflanzen (Fagus sylvatica L.) durchgeführt, einer der wirtschaftlich bedeutendsten Baumarten Deutschlands. Alle Untersuchungen wurden an den Pflanzen der Trockenstressgruppe (TSG) und der bewässerten Kontrollgruppe (KG) vorgenommen, um eine Vergleichbarkeit der Ergebnisse zu gewährleisten. Die KG setzte sich aus 9 Töpfen mit insgesamt 22 Buchenpflanzen und die TSG aus 8 Töpfen mit 24 Bäumen zusammen (vgl. Abb. 2.1).

In einem wöchentlichen Rhythmus fanden Blattwassergehaltmessungen, Chlorophyllmessungen mittels Minolta Chlorophyll Meter „SPAD-502“, sowie Aufnahmen der Spektralsignaturen mittels des Labor- und Geländespektrometers ASD FieldSpec auf der Blattebene statt. Des Weiteren wurden während des Experimentzeitraums Aufnahmen der beiden Baumgruppen mit dem bildgebenden Spektrometer HySpex VNIR-1600, das im Labor, für Geländemessungen oder auch flugzeuggestützt eingesetzt werden kann, durchgeführt.



Abb. 2.1: Standort und Aufbau des Trockenstressversuchs. Ein Alle Pflanzen stehen unter einem halbtransparenten Kunststoffdach, um sie vor Regen und direkter Sonneneinstrahlung zu schützen.

2.1 Radiometrische Messungen

2.1.1 Spektrometermessung (ASD FieldSpec II)

Die im Rahmen der Arbeit durchgeführten Messungen der spektralen Reflexion wurden mit dem FieldSpec II der Firma Analytical Spectral Devices Inc. durchgeführt, das einen Wellenlängenbereich von 350 bis 2500 nm abdeckt. Im UV/VNIR Bereich von 350 bis 1000 nm wird ein Silizium-Detektor verwendet. Der Detektor setzt sich aus 512 Photodioden zusammen, die so angebracht sind, dass sie das Licht innerhalb einer 1.4 nm schmalen Bande empfangen. Im SWIR1 und SWIR2 (1000 bis 1800 nm bzw. 1800 bis 2500 nm) werden zwei thermoelektrisch gekühlte Indium-Gallium-Arsenid-Detektoren verwendet. Die spektrale Auflösung für den ersten Detektor beträgt 3 nm, für die beiden anderen 10 nm. Für die Messung der Blattspektren wurde der ASD Leafclip eingesetzt. Der Leafclip bietet die Möglichkeit zur Aufzeichnung der Reflexion und Transmission eines einzelnen Blattes ohne es dabei zu beschädigen.

2.1.2 Bildgebende Spektrometrie (HySpex VNIR-1600)

Die HySpex VNIR-1600 ist eine Hyperspektralkamera, die von der norwegischen Firma Norsk Elektro Optikk hergestellt wird. Die HySpex VNIR - 1600 deckt mit 160 Kanälen einen spektralen Bereich zwischen 410 und 990 nm ab. Die spektrale Auflösung der Kanäle beträgt 3.7

nm. Bei der HySpex VNIR-1600 handelt es sich um eine Kamera mit einem zweidimensionalen CCD (charged couple device)-Array Sensor, welcher sich aus 1600 x 1200 Silikon Detektoren zusammensetzt. Mit einem Sichtfeld von 17 Grad quer zum Aufnahmestreifen (FOV across track), nimmt die Kamera eine Zeile von 1600 Pixeln auf.
Die Aufnahmen mit der Hyperspektralkamera erfolgten auf dem Gewächshausgelände der Universität Trier. Der Sensor wurde auf einem Gestell in 3.8 m Höhe über Grund montiert (Abb. 2.2). Neben den drei parallel angeordneten Töpfen befindet sich in dem Aufnahmebereich das rechteckige Referenzpanel aus Spectralon. Die Aufzeichnungen des Referenzspektrums sind notwendig, um die gemessenen Spektraldichten in Reflexionswerte umzurechnen. Die Datenvorprozessierung beinhaltet zuerst eine radiometrische Kalibrierung der Daten. Die Daten werden in Strahldichtewerte (at-sensor-radiance, W/sr nm m²) umgerechnet. Dabei werden auch Streifen im Bild eliminiert, die durch unterschiedliche Empfindlichkeiten der einzelnen Sensorelemente zustande kommen. Anschließend erfolgt eine räumliche Aggregierung, bei der vier Pixel zu einem zusammengefasst werden. Durch Division der Strahldichtewerte durch das mittlere Spektrum des Referenzpanels wird schließlich die relative Reflexion berechnet.



Abb. 2.2: Versuchsaufbau für die Hyperspektralaufnahmen (Die Bewegungsrichtung des Sensors ist mit dem Pfeil angedeutet).

2.2 Vegetationsindizes

Zahlreiche Autoren (PEÑUELAS et al., 1995; ROBERT et al., 2012) haben sich in den letzten Jahren bereits mit der Entwicklung von Indizes, die darauf hinzielen eine hohe Korrelation mit relevanten Vegetationsparametern wie Blattfläche, Gesamtbiomasse, Pflanzenzustand, Pigment- und Wassergehalt herzustellen, beschäftigt. Für diese Arbeit wurden dreizehn Indizes ausgewählt (vgl. Tab.2.1), die sich in zahlreichen Studien bezüglich ihrer Eignung zur Ableitung von biochemischen und biophysikalischen Vegetationsparametern, bewährt haben. Unter den diskutierten Indizes sind unter anderem traditionelle Indizes wie der SR, der NDVI und der MSI, die ursprünglich für Multispektralsensoren entwickelt wurden. Von besonderem Interesse sind jedoch Indizes, die für Hyperspektralsensoren konzipiert wurden und für ihre Berechnung sehr enge Spektralbereiche erfordern.

Tab 2.1: Übersicht über die Vegetationsindizes. Mit (*) sind Indizes markiert, die nur für hyperspektrale Sensoren geeignet sind. ρ gibt den Reflexionsgrad bei der jeweiligen Wellenlänge an.

Vegetationsindex	Indexgleichung	Referenz
Simple Ratio	$SR = \rho_{NIR} / \rho_{NIR}$	Jordan (1969)
Normalized Difference Vegetation Index	$NDVI = (\rho_{NIR} - \rho_{NIR}) / (\rho_{NIR} + \rho_{NIR})$	Rouse et al. (1973)
Green Normalized Difference Vegetation Index	$Green\ NDVI = (\rho_{NIR} - \rho_{GRÜN}) / (\rho_{NIR} + \rho_{GRÜN})$	Gitelson et al. (1996)
Red Edge Normalized Difference Vegetation Index*	$NDVI\ 705 = (\rho_{750} - \rho_{705}) / (\rho_{750} + \rho_{705})$	Gitelson & Merzlyak (1994)
Structure Intensive Pigment Index*	$SIPI = (\rho_{800} - \rho_{450}) / (\rho_{800} + \rho_{680})$	Peñuelas et al. (1995)
Plant Senescence Reflectance Index*	$PSRI = (\rho_{680} - \rho_{500}) / \rho_{750}$	Merzlyak et al. (1999)
Chlorophyll Reflectance Index*	$Chl\text{-}RI = \rho_{860} / (\rho_{708} - \rho_{550})$	Gitelson & Merzlyak (2004)
Red Edge Chlorophyll Reflectance Index*	$Chl\text{-}RI_{red\ edge} = (\rho_{750\text{-}800}) / (\rho_{695\text{-}800}) - 1$	Gitelson & Merzlyak (2004)
Moisture Stress Index	$MSI = \rho_{1600} / \rho_{820}$	Hunt & Rock (1989)
Water Index*	$WI = \rho_{900} / \rho_{970}$	Penuelas et al. (1993)
Normalized Difference Water Index	$NDWI = (\rho_{857} - \rho_{1241}) / (\rho_{857} + \rho_{1241})$	Gao (1996)
Photochemical Reflectance Index*	$PRI = (\rho_{531} - \rho_{570}) / (\rho_{531} + \rho_{570})$	Gamon et al. (1992)
Red Edge Vegetation Stress Index*	$RVSI = ((\rho_{714} - \rho_{752}) / 2) - \rho_{733}$	Merton (1998)

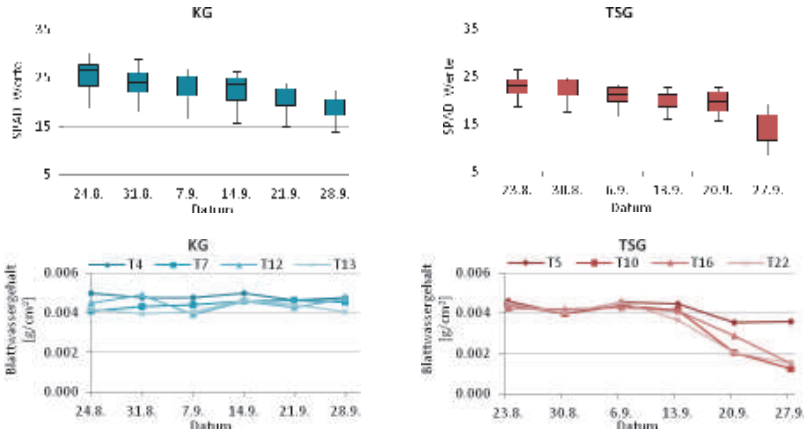


Abb. 2.3: Entwicklung der SPAD-Werte und des Blattwassergehaltes für die Kontrollgruppe (links) und Trockenstressgruppe (rechts).

3 Ergebnisse

3.1 Referenzdaten

Während des Experiments wurde die Austrocknung vieler Baumindividuen anhand der Blattfärbung und des Blattrollens sichtbar. Eine Auskunft über die tatsächliche Veränderung der Bodenfeuchte bzw. den Chlorophyll- und Blattwassergehalt geben experimentelle Messungen (vgl. Abb. 3.1). Diese Daten als Referenzdaten zu Verifizierung der berechneten Vegetationsindizes eingesetzt.

3.2 Nichtabbildende Geländespektrometrie (ASD FieldSpec II)

3.2.1 Spektrale Signaturen

Anhand der spektralen Reflexionssignaturen des Labor- und Geländespektrometers ASD FieldSpec II konnten typische Merkmale der Austrocknung, wie die Abschwächung der Absorption bei den Chlorophyllabsorptionsbanden, die Verschiebung des "Green Peaks" in Richtung des langwelligen Bereiches, der Anstieg der Reflexion im nahen Infrarot und die Abschwächung der Reflexion in den Wasserabsorptionsbanden beobachtet werden. In der Abb. 3.1 werden die Ergebnisse von dem ersten und letzten Messtermin dargestellt.

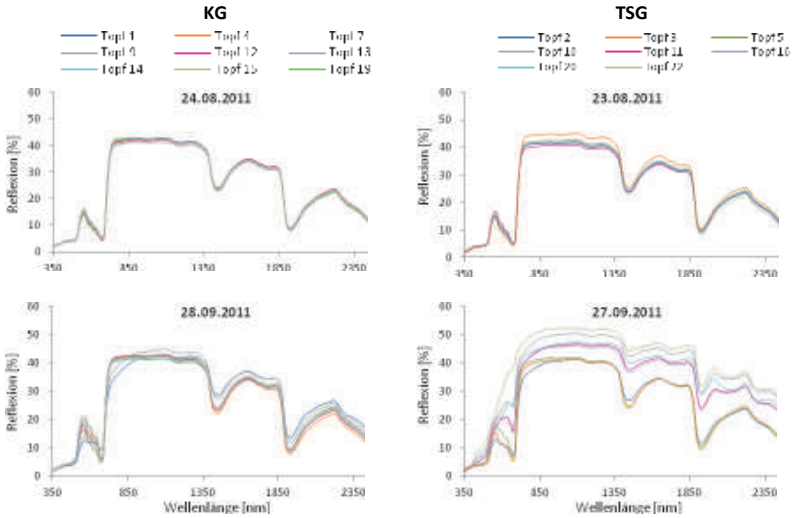


Abb. 3.1: Spektrale Signaturen (ASD FieldSpec) für die Kontrollgruppe (linke Spalte) und die Trockenstressgruppe (rechte Spalte).

3.2.2 Schätzung des Chlorophyllgehaltes und Blattwassergehaltes mithilfe von Vegetationsindizes

Beim Vergleich der Vegetationsindizes mit dem gemessenen Chlorophyllgehalt wurde festgestellt, dass die Indizes NDVI, SIPI und PSRI, die die Reflexion im roten Bereich des elektromagnetischen Spektrums verwenden, nicht sensitiv für ein breiteres Konzentrationsspektrum sind. Der Indexwert sättigt ab einem steigenden Chlorophyllgehalt auf und bleibt unverändert (vgl. Abb.3.3a). Besonders hohe Korrelationen erzielen Indizes, die die Kanäle im "red edge"- Bereich nutzen (vgl. Abb. 3.3b). Dazu gehören der NDVI 705, der Chl-RIred edge und der RVSI. Diese Indizes erwiesen sich als sensitiv unabhängig von der Höhe der Blattchlorophyllkonzentration, sind demzufolge in der Lage die Auswirkungen des Trockenstresses früher wahrnehmen und sie können somit als potentielle Stressindikatoren verwendet werden. Obwohl der PRI nicht zu der Gruppe der Indizes gehört, die sensitiv auf den Chlorophyllgehalt reagieren, zeigt sich trotzdem ein proportionaler Zusammenhang ($R^2 = 0.757$). Des Weiteren wurden in dieser Arbeit wassersensitive Indizes (MSI, WI, NDWI) untersucht und mit dem gravimetrisch ermittelten Blattwassergehalt verifiziert. Alle Indizes zeigen hohe Korrelation, die besten Ergebnisse lieferte der MSI (vgl. Abb. 3.3c).

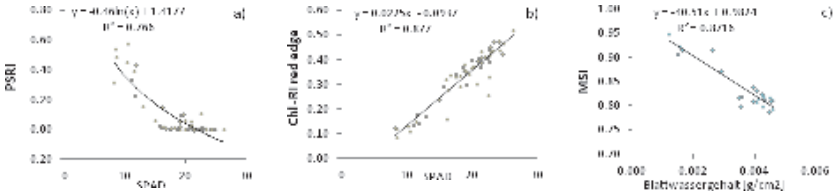


Abb. 3.2: Zusammenhang zwischen den Vegetationsindizes und den SPAD-Werten (a,b) bzw. dem Blattwassergehalt (d). Bei den SPAD-Werten bzw. dem Blattwassergehalt handelt es sich um Topfmittelwerte gemessen vom 23. August bis zum 27. September in der TSG.

3.3 Abbildende Geländespektrometrie (HySpex VNIR-1600)

Innerhalb der Experimentwochen fanden zu fünf Zeitpunkten die spektralen Aufnahmen der Baumgruppen mit der Hyperspektralkamera HySpex statt. Abbildung 3.4 gibt einen Überblick über die erhobenen Daten. Die Signaturen zeigen einen ähnlichen Verlauf wie die FieldSpec-Signaturen in diesem Abschnitt des elektromagnetischen Spektrums. Bei den Spektralsignaturen ist der "Green Peak" zu sehen, der sich mit der Austrocknung in Richtung des langwelligen Bereiches verschiebt. Die Reflexionssignaturen zeigen außerdem eine gut ausgeprägte Wasserabsorptionsbande bei ca. 970 nm. Die Gesamtreflexion ist höher als die Reflexion eines einzelnen Blattes. Die Reflexionsspektren zeigen ein gutes Signal-Rausch-Verhältnis bis auf den Bereich ab ca. 850 nm. Außerdem wird das Spektrum durch die Sauerstoffabsorptionsbande bei 760 nm beeinflusst.

TSG (Töpfe 2, 11, 22) **Topf 11**

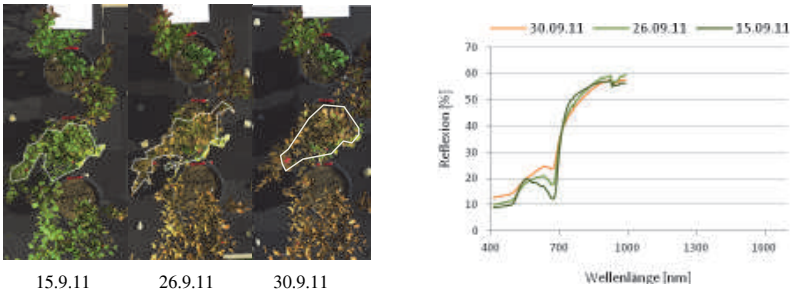
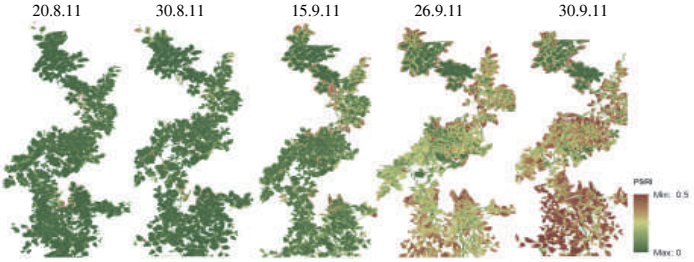


Abb. 3.3: Echtfarbendarstellung von HySpex-Aufnahmen (Kanäle 55, 41, 12) für die Töpfe 2, 11 und 22 (Trockenstressgruppe) (Reihenfolge: von oben nach unten) und die Änderung der spektralen Reflexionssignatur für den Topf und 11.

Die HySpex-Aufnahmen bieten zwei Möglichkeiten um Vegetationsindizes zu berechnen. Die erste ist die Berechnung der Indizes auf Basis von hyperspektralen Bilddaten. Das Ergebnis ist ein hochaufgelöstes Rasterbild der Baumkronen, in dem einzelnen Pixel ein bestimmter Indexwert zugeordnet wird. Eine weitere Möglichkeit ist die Berechnung der Indizes aus den extrahierten Topfmittelwertspektren, die in einem linearen Verlauf dargestellt werden.
Das unterschiedliche Verhalten der Indizes über den Messzeitraum zeigt sich auch bei den Indizes auf Basis der HySpex-Messungen. Abb. 3.4. stellt exemplarisch eine zeitliche Entwicklung des PSRI und PRI der TSG an fünf Messterminen dar. Der PSRI ein Index ist, der zu den früheren Austrocknungsstadien keinen Unterschied in den Reaktionen der verschiedenen Töpfe zeigt. Der PRI zeigt schon an den ersten beiden Terminen ein differenziertes Bild der Baumkronen. Das heißt, der Index ist in der Lage, kleinste physiologische Unterschiede zwischen einzelnen Baumindividuen zu identifizieren. Gleichermaßen wie der PRI erwiesen die Indizes NDVI 705, RVSI und Chl-$RI_{red\ edge}$ als sehr sensitiv gegenüber der durch Trockenstress verursachten Veränderungen.

TSG (Töpfe 2, 11, 22)



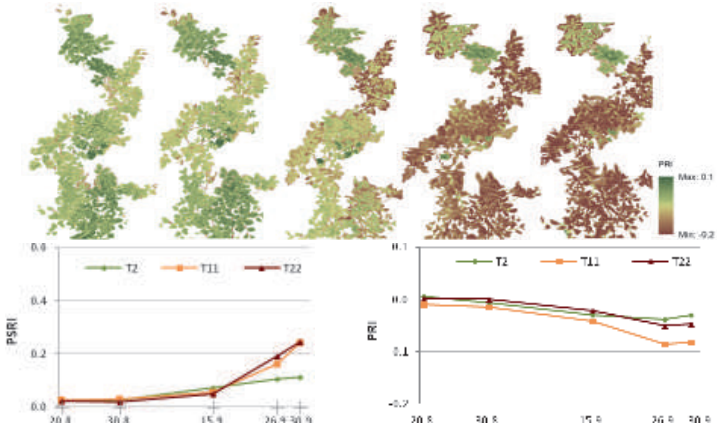


Abb. 3.4: Verlauf des PSRI und PRI für die Trockenstressgruppe (Die Topfnummerierung erfolgt von oben nach unten).

4 Zusammenfassung und Diskussion

Im Rahmen dieser Arbeit wurden Verfahren der abbildenden und nichtabbildenden Spektrometrie getestet, um physiologische Veränderung einzelner Blattorgane und Baumindividuen von Buchenpflanzen (Fagus sylvatica L.) während einer sechswöchigen Trockenstressphase zu untersuchen. Je nach Experimentbedingungen kann der Zeitraum bis zu vollständiger Austrocknung der Blätter schwanken. Es ist zu empfehlen, den Beginn des Trockenstressexperiments in zukünftigen Studien zeitlich früher in die Sommermonate zu setzen, um den Einfluss der natürlichen Seneszenz so gering wie möglich zu halten.

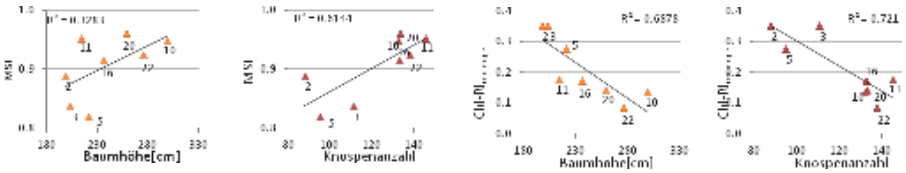


Abb. 4.1: Zusammenhang zwischen Vegetationsindizes (MSI und Chl-$RI_{red\ edge}$) und pro Topf aufsummierten Baummerkmalen (Stammdurchmesser, Baumhöhe, Knospenanzahl).

Festzustellen sind unterschiedliche Reaktionen auf den Trockenstress innerhalb der acht Töpfe in der TSG. Am letzten Termin (27.09.11) kann man dort die Ausbildung dreier Gruppen erkennen, die Ähnlichkeiten im Reflexionsverhalten aufweisen (vgl. Abb. 3.2). Um unterschiedliche Reaktion auf Trockenstress erklären zu können, wurden biometrische Kennwerte wie Baumhöhe

und Knospenanzahl im Zusammenhang mit dem MSI, der als Abschätzung für Blattwassergehalt fungiert und dem Chl-$RI_{red\ edge}$, der in der Beziehung mit dem Chlorophyllgehalt steht, untersucht (Abb. 4.1). Beide Indizes zeigten in dieser Arbeit den besten Zusammenhang mit den experimentell ermittelten Referenzdaten.

Bei den HySpex-Aufnahmen können sich auf Grund der Lage der Sauerstoffabsorptionsbande, mit einem Maximum bei 760 nm Ungenauigkeiten bei der Berechnung des NDVI 705, Chl-RIred edge und RVSI ergeben. Bei Betrachtung der HySpex-Reflexionssignatur fällt auf, dass das Spektrum in dem Bereich zwischen 746 und 764 nm, in dem die Absorptionsbande liegt, verrauscht ist. In unserem Fall konnten zwei Faktoren festgestellt werden, die die Indexwerte bei den HySpex-Bilddaten, unabhängig vom Index verfälschen. In diesem Experiment trugen sie zu einer Erhöhung der Werte bei. Zum einem ist es das Auftreten des Schattens innerhalb der Baumkronen, wodurch der Indexwert an der betroffenen Stelle erhöht wird. Zum anderen verändert sich während der Austrocknung die Aufnahmegeometrie bei den Untersuchungspflanzen.

Der spektrale Bereich, der durch den Hyperspektralsensor HySpex VNIR-1600 abgedeckt wird, reicht nicht aus, um Aussagen über den Blattwassergehalt zu treffen. Die Verwendung des zusätzlichen Moduls HySpex SWIR-320m würde eine Ableitung der wassersensitiven Indizes ermöglichen. Aus der Kombination von beiden Modulen könnten Indizes sowohl aus den pigment- als auch den zellstrukturbeeinflussten Bereichen des elektromagnetischen Spektrums abgeleitet werden. Somit könnten präzisere Aussagen über das Stadium und die Stärke des Trockenstresses getroffen werden.

5 Literaturverzeichnis

ALLEN, C.D., MACALADY, A.K, CHENCHOUNI, H. et al., 2010: A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. Forest Ecology and Management, **259**, S. 660–684.

BOLTE, A., EISENHAUER, D.R., EHRHART, H.P., GROß, J., HANEWINKEL, M., KÖLLING, C., PROFFT, I., ROHDE, M., RÖHE, P. & AMERELLER, K., 2009: Klimawandel und Forstwirtschaft – Übereinstimmungen und Unterschiede bei der Einschätzung der Anpassungsnotwendigkeiten und Anpassungsstrategien der Bundesländer. Landbauforschung - vTI agriculture and forestry research, **59**, S. 269–278.

GAMON, J.A., PEÑUELAS, J. & FIELD, C.B., 1992: A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. Remote Sensing of Environment, **41**, S. 35–44.

GAO, B.C., 1996: NDWI - a normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment, **58**, S. 257–266.

GITELSON, A.A. & MERZLYAK, M.N., 2004: Non-destructive assessment of chlorophyll, carotenoid and anthocyanin content in higher plant leaves: Principles and algorithms. Remote Sensing for Agriculture and the Environment (S. Stamatiadis, J.M., Lynch, J.S. Schepers Eds.). Greece, Ella, S. 78–94.

GITELSON, A.A., KAUFMAN, Y.J. & MERZLYAK, M.N., 1996: Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment, **58**, S. 289–298.

HUNT, E.R. & ROCK, B.N.,1989: Detection of changes in leaf water content using near- and middle-infrared reflectance. Remote Sensing of Environment. **30**, S. 43–54.

JORDAN, C.F., 1969: Derivation of leaf-area index from quality of light on the forest floor. Ecology, **50**, S. 663–666.

MERTON, R.N., 1998:. Monitoring community hysteresis using spectral shift analysis and the red-edge vegetation stress Index. In Proceedings of the Seventh Annual JPL Airborne Earth Science Workshop. Pasadena (CA), USA.

MERZLYAK, J.R., GITELSON, A.A., CHIVKUNOVA, O.B. & RAKITIN, V.Y., 1999: Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. Physiologia Plantarum, **106**, S. 135–141.

PEÑUELAS, J., BARET, F. & FILELLA I., 1995: Semi-empirical indices to assess carotenoids/ chlorophyll-a ratio from leaf spectral reflectance. Photosynthetica, **31**, S. 221–230.

PEÑUELAS, J., FILELLA, I., BIEL, C., SERRANO, L. & SAVE, R., 1993: The reflectance at the 950-970 region as an indicator of plant water status. International Journal of Remote Sensing, **14**, S. 1887–1905.

ROBERT, D.A., ROTH, K.L. & PERROY, R.L. (2012): Hyperspectral vegetation indices. Chapter 14 in Hyperspectral Remote Sensing of Vegetation (Thenkabail, P.S., Lyon, J.G., Huete, A., Eds), Taylor and Francis, S. 141–165.

ROUSE, J.W., HAAS, R.H., SCHELL, J.A. & DEERING, D.W.,1973: Monitoring vegetation systems in the Great Plains with ERTS. In Proceedings of the 3rd ERTS Symposium, NASA SP-351, S. 309–317. Washington (DC), USA.

SITTE, P., WEILER, E.W., KADEREIT, J.W., BRESINSKY A. & KÖRNER, C., 2002: Strasburger. Lehrbuch der Botanik für Hochschulen. Heidelberg, Spektrum.

Möglichkeiten einer stratifizierten Schätzung von forstlichen Kenngrößen unter Verwendung von amtlichen Stereo-Luftbildern, Orthophotos und Laserscannerdaten

CHRISTOPH STRAUB[1], CHRISTOPH STEPPER[1], RUDOLF SEITZ[1] & LARS T. WASER[2]

Zusammenfassung: Ziel einer gemeinsamen Studie der Bayerischen Landesanstalt für Wald und Forstwirtschaft (LWF) und der Eidg. Forschungsanstalt für Wald, Schnee und Landschaft (WSL) ist die Entwicklung praxistauglicher Methoden zur flächendeckenden Schätzung forstlicher Kenngrößen mit amtlichen Fernerkundungsdaten. Im vorliegenden Artikel werden Möglichkeiten einer stratifizierten Schätzung des Holzvorrats und der Grundfläche pro Hektar vorgestellt. Hierfür wurden Punktwolken aus Stereo-Luftbildern und Laserscannerdaten mit Stichprobenpunkten einer Waldinventur in Beziehung gesetzt. Über eine Klassifizierung von Orthophotos wurden die Stichprobenpunkte in laubholz- und nadelholzdominierte Punkte stratifiziert, um jeweils getrennte Modelle zur Vorhersage der forstlichen Kenngrößen zu erstellen. Die Ergebnisse zeigen, dass die Schätzgenauigkeit durch die Stratifizierung leicht verbessert werden kann.

1 Einleitung

Die flugzeuggetragene Laserscannermessung wird häufig als die genauste Fernerkundungstechnik zur Schätzung forstlicher Kenngrößen angesehen (JÄRNSTEDT et al. 2012), da sowohl die Kronenoberfläche als auch das Gelände unterhalb der Vegetation erfasst werden kann (KOCH 2011). Bei entsprechender Punktdichte sind zusätzlich Aussagen zur vertikalen Schichtung möglich. Laserscannerdaten werden in Norwegen bereits operational bei Waldinventuren eingesetzt (NÆSSET 2007).
Alternativ dazu haben einige Studien die Eignung digitaler Stereo-Luftbilder zur Erstellung von hochaufgelösten, digitalen Oberflächenmodellen (DOM) mit Image Matching Techniken in Waldgebieten aufgezeigt (z.B. STRAUB & SEITZ 2011 oder BALTSAVIAS et al. 2008). Steht zusätzlich zu den Stereo-Luftbildern ein Laser-Geländemodell (DGM) zur Verfügung, kann aus der Differenz zwischen Luftbild-DOM und Laser-DGM ein Luftbild-Kronenhöhenmodell (KHM) abgeleitet werden (GINZLER 2012).

1) Dr. Christoph Straub, Bayerische Landesanstalt für Wald und Forstwirtschaft (LWF), Abteilung Informationstechnologie, Hans-Carl-von-Carlowitz-Platz 1, D-85354 Freising; Email: Christoph.Straub@lwf.bayern.de
1) Christoph Stepper, Bayerische Landesanstalt für Wald und Forstwirtschaft (LWF), Abteilung Informationstechnologie, Hans-Carl-von-Carlowitz-Platz 1, D-85354 Freising; Email: Christoph.Stepper@lwf.bayern.de
1) Rudolf Seitz, Bayerische Landesanstalt für Wald und Forstwirtschaft (LWF), Abteilung Informationstechnologie, Hans-Carl-von-Carlowitz-Platz 1, D-85354 Freising; Email: Rudolf.Seitz@lwf.bayern.de
2) Dr. Lars T. Waser, Eidg. Forschungsanstalt für Wald, Schnee und Landschaft (WSL), Gruppe Fernerkundung, Zuercherstrasse 111, CH-8903 Birmensdorf; E-Mail: Lars.Waser@wsl.ch

Aktuelle Untersuchungen zeigen das große Potential von Luftbild-Kronenhöhenmodellen zur Schätzung forstlicher Kenngrößen (JÄRNSTEDT et al. 2012, BOHLIN et al. 2012, STRAUB et al. 2012).
In dieser Studie wurden Möglichkeiten zur Schätzung von zwei wichtigen forstlichen Kenngrößen (Holzvorrat und Grundfläche pro Hektar) untersucht. Als Grundlage dienten amtliche Stereo-Luftbilder, Orthophotos und Laserscannerdaten sowie terrestrische Stichprobenerhebungen einer Waldinventur. In diesem Zusammenhang wurde vor allem geprüft, ob durch eine Stratifizierung der Inventurpunkte in laub- und nadelholzdominierte Stichprobenkreise der Schätzfehler reduziert werden kann.

2 Versuchsfläche und Datensätze

Das Untersuchungsgebiet ist eine Teilfläche des Forstbetriebs Stadtwald Traunstein. Da es sich hierbei um eine langjährige Versuchsfläche der Forstwissenschaftlichen Fakultät der TU München handelt (TUM 2012), stehen umfangreiche Feldaufnahmen zur Verfügung.
Der Stadtwald Traunstein ist durch sehr strukturreiche (oft mehrschichtige) Waldbestände mit einer ausgeprägten Baumartenmischung gekennzeichnet. Die häufigsten Baumarten sind Fichte (*Picea abies*), Buche (*Fagus sylvatica*) und Tanne (*Abies alba*). Im Jahr 2008 wurde vom Lehrstuhl für Waldwachstumskunde der TU München eine Waldinventur mit permanent markierten Stichprobenpunkten durchgeführt. Hierbei wurden die Stichprobenzentren in einem regelmäßigen 100 × 100 m Gitter über die gesamte Waldfläche verteilt. An jedem Stichprobenpunkt wurden einzelne Bäume in konzentrischen Probekreisen in Abhängigkeit ihres Brusthöhendurchmessers (BHD) gemessen. Eine detaillierte Beschreibung der Aufnahmemethodik findet sich in FER (2011). Der größte Probekreis, der als Bezugsfläche für die Auswertungen verwendet wurde, hat einen Radius von 12,62 m, was einer Flächengröße von 0,05 ha entspricht. Auf Grundlage der Einzelbaummessungen konnte an jedem Stichprobenpunkt der gesamte Holzvorrat in Erntefestmeter pro Hektar V [m³ ha^{-1}] sowie die Grundfläche G [m² ha^{-1}] als die Summe aller Stamm-Querschnittsflächen in 1,3 m Höhe modelliert werden (Tabelle 1). Insgesamt lagen 225 Inventurpunkte im Untersuchungsgebiet, die für die Modellierung verwendet wurden.

Tab. 1: Forstliche Kenngrößen Holzvorrat V [m³ ha^{-1}] und Grundfläche G [m² ha^{-1}] im Untersuchungsgebiet Traunstein (berechnet mit 225 Stichprobenpunkten der Waldinventur)

Kenngröße	Mittelwert	Standardabweichung	Min	Max
Holzvorrat V [m³ ha^{-1}]	321,99	225,60	0,00	968,40
Grundfläche G [m² ha^{-1}]	28,72	16,98	0,00	77,27

Für die Versuchsfläche standen insgesamt 30 vierkanalige amtliche Stereo-Luftbilder (Längsüberdeckung: 65%; Querüberdeckung: 30%) zur Verfügung. Die Luftbild-Befliegung erfolgte Ende April 2009 mit einer Vexcel UltraCamX Kamera. Ferner wurden vom Landesamt für Vermessung und Geoinformation (LVG) Bayern Orthophotos mit 20 cm Bodenauflösung berechnet, die in der vorliegenden Studie ebenfalls verwendet wurden.

Die Aufnahme der amtlichen Laserscannerdaten erfolgte Ende März und Anfang April 2010 mit einer Punktdichte von 5-6 Punkten / m². Für die einzelnen Laserpunkte wurde vom LVG eine Höhengenauigkeit von 95% ≤ 0,16 m und eine Lagegenauigkeit von ±0,4 m angegeben. Die Laserpunkte wurden getrennt in zwei Datensätzen (erstes Echo und letztes Echo) geliefert. Zusätzlich stellte das LVG ein gefiltertes und interpoliertes Geländemodell mit 1 m Gitterweite (DGM 1) zur Verfügung.

3 Methodik

3.1 Stratifizierung der Inventurpunkte mittels einer Laub- und Nadelholzklassifizierung

Im ersten Bearbeitungsschritt wurde das Untersuchungsgebiet auf der Grundlage der amtlichen Orthophotos in Laub- und Nadelholzflächen klassifiziert. Die Klassifizierung durchläuft dabei mehrere Prozessierungsschritte, welche in WASER et al. (2011) und in WASER (2012) ausführlich beschrieben werden. Das Verfahren basiert auf dem Prinzip, dass mit Hilfe von logistischen Regressionsmodellen die Beziehung der Zielvariablen (hier Baumarten) von einer Auswahl erklärender Variablen (hier abgeleitete statistische Werte der UltraCamX Reflektanzkanäle) bestimmt wird. In verschiedenen Auswahlverfahren wurden diejenigen erklärenden Variablen bestimmt, welche am meisten zur Unterscheidung der verschiedenen Laub- und Nadelgehölze beitragen. Anschließend wurde die Qualität der Modelle durch eine 10-fache Kreuzvalidierung getestet. Die erzielte Gesamtgenauigkeit der Laub- und Nadelholzunterscheidung beträgt 96%, der entsprechende Kappa-Koeffizient 0.90.
Auf dieser Grundlage wurde für jeden Inventurkreis (0,05 ha) der prozentuale Anteil von Nadelholz (P_N) bzw. Laubholz (P_L) berechnet, um die Stichprobenpunkte entweder dem Stratum „nadelholzdominiert“ oder „laubholzdominiert“ zuzuweisen:

$$\text{Stratum}\begin{cases} \text{nadelholzdominiert falls } P_N > 50\% \\ \text{laubholzdominiert falls } P_L > 50\% \end{cases}$$

Von den 225 Inventurpunkten im Untersuchungsgebiet entfielen somit 179 Stichprobenkreise auf das Stratum „nadelholzdominiert“, die verbleibenden 46 Stichprobenkreise wurden dem Stratum „laubholzdominiert“ zugeteilt. Abbildung 1 zeigt einen Ausschnitt vom Stadtwald Traunstein mit dem Ergebnis der Klassifizierung von Laub- und Nadelholzflächen sowie eine exemplarische Zuordnung von Stichprobenkreisen zu den Straten.

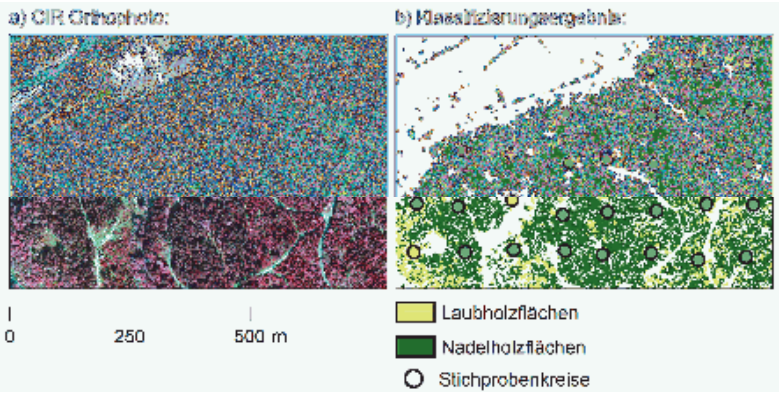


Abb. 1: Ausschnitt vom Stadtwald Traunstein mit a) Farb-Infrarot Orthophoto und b) Klassifizierung von Laub- und Nadelholzflächen sowie Zuordnung der Stichprobenkreise zu den Straten

3.2 Berechnung von Vegetationshöhen mit Stereo-Luftbildern und Laserscannerdaten

Im zweiten Schritt wurde eine dichte photogrammetrische Punktwolke (der Baumkronenoberfläche) aus den Stereo-Luftbildern unter Verwendung eines Image Matching Verfahrens berechnet. Hierfür wurde die Software LPS eATE – Version 11.0.3 verwendet. Für die Ableitung von Vegetationshöhen wurden anschließend die einzelnen photogrammetrischen Höhenmessungen sowie die First-Echo-Laserscannermessungen mit dem DGM 1 normalisiert. Abbildung 2 zeigt beispielhaft einen Querschnitt der Laserpunkte und der photogrammetrischen Messungen.

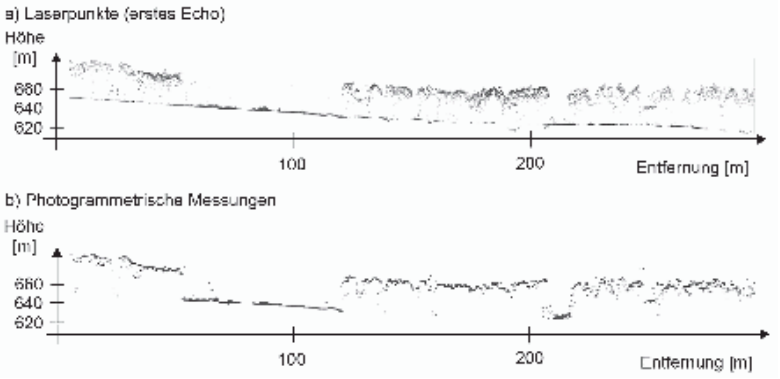


Abb. 2: Querschnitt von a) Laserpunkten (erstes Echo) und b) photogrammetrischen Höhenmessungen

3.3 Schätzung des Holzvorrats und der Grundfläche

Mittels schrittweisen, linearen Regressionen wurden Zusammenhänge zwischen abgeleiteten Variablen aus der photogrammetrischen sowie aus der Laser-Punktwolke mit den terrestrisch erhobenen Holzvorräten und der Grundfläche an den Stichprobenpunkten der Waldinventur untersucht. Dies erfolgte zuerst für alle Stichprobenpunkte (ohne Stratifizierung) und anschließend jeweils getrennt für die Straten „laubholzdominiert" und „nadelholzdominiert". Die folgenden Prädiktorvariablen wurden (sowohl für die photogrammetrische als auch für die Laser-Punktwolke) zur Schätzung verwendet:

1. Mehrere Höhenmetriken bzw. die Quantile h_{10}, h_{20}, ..., h_{90} [m], Maximalwert h_{max} [m] und Mittelwert h_m [m]
2. Der Überschirmungsgrad $\ddot{U}G$ [%] wurde über das Verhältnis von klassifizierten „Vegetationspunkten" zur Gesamtanzahl der Höhenmessungen am jeweiligen Inventurpunkt geschätzt. Die Klassifizierung von Vegetationspunkten erfolgte in Anlehnung an MALTAMO et al. (2006) und NÆSSET (2002) über einen Schwellwert von > 2 m.
3. Die Oberflächenrauhigkeit wurde über den Variationskoeffizient h_{vk} der Höhenmessungen charakterisiert.

Die Beurteilung der Schätzgenauigkeit erfolgte auf Plotebene über den absoluten und relativen Schätzfehler (RMSE), welche über eine Leave-One-Out Kreuzvalidierung ermittelt wurden.

4 Ergebnisse

Tabelle 2 zeigt das Ergebnis der Regressionsanalyse zur Schätzung des Holzvorrats und der Grundfläche auf Grundlage amtlicher Laserdaten. In Tabelle 3 ist die Schätzung bei Verwendung der photogrammetrischen Höhenmessungen aus amtlichen Stereo-Luftbildern dargestellt. Es ist jeweils das Ergebnis ohne Stratifizierung und mit Stratifizierung in nadel- und laubholzdominierte Inventurpunkte dargestellt mit dem zugehörigen Regressionsmodell, Bestimmtheitsmaß R^2 sowie absoluter und relativer RMSE. Für beide Fernerkundungsdatensätze konnte der RMSE durch die Stratifizierung verringert werden. Beim Einsatz von Laserdaten wurde der RMSE für die Schätzung des Holzvorrats von 36,82% auf 34,27% reduziert. Bei Verwendung von Luftbildern konnte eine Reduktion von 41,04% auf 37,68% erzielt werden. Ebenso verbesserte sich die Schätzung der Grundfläche durch die Stratifizierung in einer vergleichbaren Größenordnung.

Tab. 2: Schätzung des Holzvorrats V [m³ ha^{-1}] und der Grundfläche G [m² ha^{-1}] auf der Grundlage von amtlichen Laserdaten (ohne und mit Stratifizierung)

Abhängige Variable	Modell	R^2	RMSE	RMSE [%]
Holzvorrat V [m³ ha^{-1}]	Ohne Stratifizierung:			
	$28{,}849h_m$ -$6{,}063h_{10}$ -$1{,}399\ddot{U}G$ +$2{,}704h_{max}$ -40,788	0,74	118,56	36,82
	Mit Stratifizierung:			
	Nadelholz: $34{,}041h_m$ -$7{,}049h_{10}$ -$1{,}593\ddot{U}G$ +2,454	0,79	110,36	34,27
	Laubholz: $16{,}812h_m$ +17,880	0,61		
Grundfläche G [m² ha^{-1}]	Ohne Stratifizierung:			
	$1{,}641h_m$ -$0{,}454h_{10}$ +$0{,}111\ddot{U}G$ -0,096	0,67	10,00	34,82
	Mit Stratifizierung:			
	Nadelholz: $1{,}798h_m$ $0{,}116\ddot{U}G$ -$0{,}418h_{10}$ -0,836	0,73	9,08	31,62
	Laubholz: $1{,}267h_m$ +3,952	0,60		

Tab. 3: Schätzung des Holzvorrats V [m³ ha^{-1}] und der Grundfläche G [m² ha^{-1}] auf der Grundlage von amtlichen Stereo-Luftbildern (ohne und mit Stratifizierung)

Abhängige Variable	Modell	R^2	RMSE	RMSE [%]
Holzvorrat V [m³ ha^{-1}]	Ohne Stratifizierung:			
	$13{,}227h_{90}$ +$5{,}850h_{30}$ -$0{,}823\ddot{U}G$ +6,175	0,67	132,14	41,04
	Mit Stratifizierung:			
	Nadelholz: $11{,}510h_m$ $9{,}905h_{90}$ -$0{,}819\ddot{U}G$ -5,264	0,74	121,32	37,68
	Laubholz: $10{,}031h_{90}$ +26,071	0,53		
Grundfläche G [m² ha^{-1}]	Ohne Stratifizierung:			
	$0{,}379h_{80}$ +$1{,}633h_m$ -$0{,}730h_{40}$ +5,920	0,58	11,18	38,93
	Mit Stratifizierung:			
	Nadelholz: $0{,}662h_{80}$ +$0{,}659h_m$ +6,986	0,65	10,21	35,55
	Laubholz: $0{,}726h_{90}$ +5,196	0,49		

5 Diskussion und Ausblick

Die Ergebnisse dieser Studie zeigen, dass amtliche Stereo-Luftbilder zur Schätzung des Holzvorrats und der Grundfläche pro Hektar genutzt werden können. Durch eine vorgeschaltete Stratifizierung in laub- und nadelholzdominierte Stichprobenkreise konnte die Schätzgenauigkeit leicht verbessert werden. Hierbei wurden für die nadelholzdominierten Stichprobenkreise (im Vergleich zu den laubholzdominierten Kreisen) etwas höhere Bestimmtheitsmaße erzielt. Die Ergebnisse bestätigen, dass mit Laserdaten etwas geringere Schätzfehler erzielt werden, was sich mit der Untersuchung von JÄRNSTEDT et al. (2012) deckt, in welcher forstliche Kenngrößen in einer Versuchsfläche in Finnland mittels k-NN (k nearest neighbor) geschätzt wurden. Bei Verwendung von Luftbildern berichten JÄRNSTEDT et al. (2012) einen RMSE von 40,39% für die Schätzung des Holzvorrats sowie von 36,23% für die Grundfläche. Für Laserdaten wurde ein RMSE von 31,26% für den Holzvorrat und 27,89% für die Grundfläche angegeben. Obwohl ein direkter Vergleich aufgrund unterschiedlicher Versuchsflächen und Datensätze nicht möglich ist, befinden sich die Schätzfehler der hier vorliegenden Untersuchung in einem ähnlichen Wertebereich.

Zur Ableitung von Vegetationshöhen aus den photogrammetrischen Messungen muss ein Laser-Geländemodell zur Verfügung stehen. In Deutschland sind digitale Geländemodelle aus Laserscannerdaten für mehrere Bundesländer verfügbar, so ist z.B. für Bayern ein hochaufgelöstes DGM aus Laserscannermessung flächendeckend vorhanden (LVG 2012a). Eine

regelmäßige Aktualisierung der amtlichen Laserdaten ist gegenwärtig allerdings nicht vorgesehen. Im Gegensatz dazu werden amtliche Stereo-Luftbilder in Deutschland (und auch in anderen europäischen Ländern) regelmäßig aktualisiert. In Bayern erfolgt die Aktualisierung in einem Drei-Jahres-Turnus, wobei jährlich circa 25000 Quadratkilometer neu beflogen werden (LVG 2012b).

Zusammenfassend zeigt diese Untersuchung, dass wichtige forstliche Kenngrößen unter Verwendung amtlicher Fernerkundungsdaten flächig geschätzt bzw. regionalisiert werden können. Dadurch werden kleinflächige und räumlich detaillierte Abschätzungen möglich, was forstliche Planungsaufgaben und Entscheidungen maßgeblich erleichtern und unterstützen kann.

Eine zukünftige Verbesserung der Schätzung forstlicher Kenngrößen wird durch eine noch feinere Stratifizierung angestrebt, im Idealfall über eine Klassifizierung der häufigsten Baumarten im Untersuchungsgebiet. Ferner ist die Schätzung weiterer Kenngrößen beabsichtigt, wie beispielsweise mittlerer Stammdurchmesser oder das Alter von Bäumen in der Oberschicht. Parallel dazu wird derzeit in einem weiteren Projekt getestet, inwieweit Oberflächenmodelle aus hochaufgelösten Stereo-Satellitenbildern (insbesondere Cartosat-1 und WorldView-2) für die Schätzung forstlicher Kenngrößen geeignet sind.

6 Literaturverzeichnis

BALTSAVIAS, E., GRUEN, A., EISENBEISS, H., ZHANG, L. & WASER, L.T., 2008: High quality image matching and automated generation of 3D tree models. International Journal of Remote Sensing, **29** (5), S. 1243-1259.

BOHLIN, J., WALLERMANN, J. & FRANSSON, J. E. S., 2012: Forest variable estimation using photogrammetric matching of digital aerial images in combination with a high-resolution DEM. Scandinavian Journal of Forest Research, **27** (7), S. 692-699.

FER, 2011: Richtlinie für die mittel- und langfristige Forstbetriebsplanung in den Bayerischen Staatsforsten. Forsteinrichtungsrichtlinie – FER 2011. Bayerische Staatsforsten, Regensburg.

GINZLER, C., 2012: Vom Punkt in die Fläche. LFI info, **14**, S. 1-4.

JÄRNSTEDT, J., PEKKARINEN, A., TUOMINEN, S., GINZLER, C., HOLOPAINEN, M. & VIITALA, R., 2012: Forest variable estimation using a high-resolution digital surface model. ISPRS Journal of Photogrammetry and Remote Sensing, **74**, S. 78-84.

KOCH, B., 2011: Stand und Perspektiven der Nutzung neuer Fernerkundungstechnologien im Waldbereich. Schweizerische Zeitschrift für Forstwesen, **162** (6), S. 156-163.

LVG, 2012a: Faltblatt – Digitale Geländemodelle. Produktinformation vom Landesamt für Vermessung und Geoinformation Bayern. Online: http://vermessung.bayern.de/file/pdf/1614/download_faltblatt-dgm09.pdf (17.12.2012).

LVG, 2012b: Faltblatt – Luftbildprodukte. Produktinformation vom Landesamt für Vermessung und Geoinformation Bayern. Online: http://www.vermessung.bayern.de/file/pdf/1039/download_faltblatt-luftbilder08.pdf (17.12.2012).

MALTAMO, M., MALINEN, J., PACKALÉN, P., SUVANTO, A. & KANGAS, A., 2006: Nonparametric estimation of stem volume using airborne laser scanning, aerial photography and stand-register data. Canadian Journal of Forest Research, **36** (2), S. 426-436.

NÆSSET, E., 2002: Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment, **80** (1), S. 88-99.

NÆSSET, E., 2007: Airborne laser scanning as a method in operational forest inventory: Status of accuracy assessments accomplished in Scandinavia. Scandinavian Journal of Forest Research, **22** (5), S. 433-442.

STRAUB, C. & SEITZ, R., 2011: Möglichkeiten der automatisierten Generierung von Oberflächenmodellen in Waldgebieten aus digitalen Luftbildern. DGPF Tagungsband 20 / 2011, Mainz, S. 153-162.

STRAUB, C., SEITZ, R. & WASER, L. T., 2012: Comparison of digital aerial images and airborne laser scanning data for timber volume estimation on plot level. ForestSAT 2012 conference, 11-14 September, Oregon State University, Corvallis, Oregon, USA.

TUM, 2012: Stadtwald Traunstein, Informationen vom Lehrstuhl für Waldwachstumskunde der TU München. Online: http://www.wwk.forst.wzw.tum.de/info/traunstein/ (17.12.2012).

WASER, L.T., GINZLER, C., KUECHLER, M., BALTSAVIAS, E. & HURNI, L. 2011: Semi-automatic classification of tree species in different forest ecosystems by spectral and geometric variables derived from Airborne Digital Sensor (ADS40) and RC30 data. Remote Sensing of Environment, **115** (1), S. 76-85.

WASER, L.T., 2012: Airborne remote sensing data for semi-automated extraction of tree area and classification of tree species. Dissertation 20464, ETH Zürich. 153 S.

Entwicklung eines robusten Veränderungsdetektionsverfahrens für räumlich hoch auflösende Satellitendaten

FRANK THONFELD[1] & GUNTER MENZ[2]

Zusammenfassung: Veränderungsdetektion kann in unterschiedlichen räumlichen, zeitlichen, spektralen und radiometrischen Dimensionen erfolgen. Neben den unterschiedlichen Sensorcharakteristika haben räumlich (extrem) hoch auflösende Systeme häufig die Fähigkeit, ihre Sensoren zu schwenken. Dadurch entstehen Bildverzerrungen, die die Veränderungsdetektion erschweren, da sie Scheinveränderungen bewirken. Das Ausmaß der Verzerrungen hängt von der Topographie des untersuchten Raums sowie von der Sonne-Sensor-Objekt-Geometrie ab. Zur Optimierung der Veränderungsdetektion wird daher ein robustes Verfahren vorgestellt, das die Pixelnachbarschaften berücksichtigt und die kleinste Differenz zwischen Bildern zweier Zeitpunkte innerhalb eines gleitenden Fensters berechnet. Dieses Prinzip wird auf die Change Vector Analyse (CVA) erweitert, da diese neben einer Unterscheidung von Veränderungen und Nicht-Veränderungen auch die thematische Zuordnung der Veränderungen erlaubt.

1 Einleitung

Im Kontext der Veränderungsdetektion kann zwischen bi-temporalen Verfahren und Verfahren, die mehr als zwei Zeitpunkte betrachten, unterschieden werden. Letztere können in Trajektorenanalyse (z.B. jährliche Aufnahmen eines Gebiets) und Zeitreihenanalyse (d.h. mehr oder weniger kontinuierliche Informationen) untergliedert werden. Für viele Fragestellungen ist jedoch die bi-temporale Veränderungsdetektion ausreichend. Dies gilt insbesondere bei abrupten Veränderungen und bei Konversionen. Konversionen sind definiert als Ersetzung einer Landbedeckung durch eine andere (COPPIN et al. 2004). Um Fehldetektionen zu vermeiden, werden annuelle Aufnahmen zum selben Zeitpunkt im Jahr empfohlen, um phänologische Unterschiede und Effekte durch unterschiedliche Sonnenstände zu minimieren (COPPIN et al. 2004). Empfohlen werden dabei Aufnahmen im Sommer, weil da die Schatten am geringsten sind und die Vegetation phänologisch stabil ist. In Abhängigkeit von der Wolkenbedeckung sind die gewünschten Aufnahmen häufig jedoch nicht verfügbar, insbesondere bei starren Sensoren mit festem Aufnahmeschema.
Im Fall der Notwendigkeit zeitnaher Aufnahmen, beispielsweise zur Katastrophenhilfe, kann mitunter nicht auf Satellitendaten unter optimalen Aufnahmebedingungen gewartet werden. Moderne Satelliten haben häufig schwenkbare Sensoren, um flexibel eingesetzt werden zu können. Dadurch kann kurzfristig von Aufnahmeplänen abgewichen werden, um dringlichere und wichtigere Aufgaben wahrzunehmen bzw. auf wolkenfreie Areale auszuweichen. Durch das

1) Frank Thonfeld, Zentrum für Fernerkundung der Landoberfläche (ZFL), Universität Bonn, Walter-Flex-Str. 3, 53113 Bonn, E-Mail: frank.thonfeld@uni-bonn.de
2) Gunter Menz, Geographisches Institut, Remote Sensing Research Group (RSRG), Universität Bonn, Meckenheimer Allee 166, 53115 Bonn und Zentrum für Fernerkundung der Landoberfläche (ZFL), Universität Bonn, Walter-Flex-Str. 3, 53113 Bonn, E-Mail: g.menz@uni-bonn.de

Schwenken der Sensoren werden Off-Nadir-Aufnahmen erzeugt. Die Einfallswinkel der beiden Aufnahmen sind bei schwenkbaren Sensoren nicht selten unterschiedlich, sodass die damit verbundenen Verzerrungen jeweils individuell ausfallen. An Oberflächen mit ausgeprägtem Relief und dreidimensionalen Strukturen wie Gebäuden oder Waldkanten sind die Verzerrungen und die Auswirkungen unterschiedlicher Sonnenstände am stärksten ausgeprägt. Die Bedeutung der Verzerrungen ist bei räumlich hoch auflösenden Daten wesentlich stärker zu gewichten als bei Sensoren mittlerer oder geringer Auflösung (Landsat-Auflösung 30 m oder geringer). Neben der Sensor-Sonne-Objekt-Geometrie gibt es noch weitere Phänomene, die die Veränderungsdetektion erschweren. Dazu zählt beispielsweise die Tatsache, dass eine Auflösungszelle nicht immer das identische Gebiet an der Erdoberfläche abdeckt (BRUZZONE & COSSU 2003). Selbst bei subpixelgenauer Bildregistrierung verändert sich der Inhalt eines Pixels, ohne dass sich an der Oberfläche etwas ändert. Dies führt zu deutlicher Überschätzung der Veränderung (DAI & KHORRAM 1998). In der Vergangenheit wurde versucht, den Effekt von Fehlregistrierungen zu quantifizieren (DAI & KHORRAM 1998) und zu reduzieren (BRUZZONE & COSSU 2003). Darüber hinaus wurden Veränderungsdetektionsmethoden entwickelt, die die Pixelnachbarschaft berücksichtigen und damit die beschriebenen Effekte minimieren (z. B. CASTILLA et al. 2009; IM & JENSEN 2005). Das Verfahren von CASTILLA et al. (2009) berechnet die kleinste Differenz eines Bildkanals von unterschiedlichen Zeitpunkten innerhalb eines gleitenden Fensters. Es wird dabei davon ausgegangen, dass diese geringste Veränderung der tatsächlichen Veränderung entspricht; die Effekte von Fehlregistrierungen werden minimiert.

2 Methoden

Die von CASTILLA et al. (2009) entwickelte Methode wurde nur für die Verwendung eines Spektralkanals beschrieben. Die Auswahl geeigneter Kanäle ist zwar durchaus sinnvoll, allerdings kann durch die Verwendung weiterer Spektralkanäle die Veränderung umfassender beschrieben und detektiert werden. Ein Verfahren, dass alle Kanäle von Multispektraldaten verwendet, ist die Change Vector Analyse (MALILA 1980). Dabei wird die Veränderungsmagnitude – zu verstehen als Maß für die Veränderungsintensität – als euklidische Distanz im mehrdimensionalen Merkmalsraum berechnet. Die Magnitude *m* berechnet sich aus den Differenzen der Einzelkanäle:

$$x_{diff_i} = x_{2_i} - x_{1_i} \tag{1}$$

$$m = \sqrt{\sum_{i=1}^{n} x_{diff_i}{}^2}, m \in [0, \max(m)] \tag{2}$$

wobei x_{diff} die Differenz in Band *i* ist und *n* die Anzahl der Spektralkanäle. Die Variablen x_1 und x_2 bezeichnen die Multispektraldatensätze zu zwei unterschiedlichen Zeitpunkten. Die Stärke des Verfahrens liegt allerdings darin, dass neben der Veränderungsmagnitude auch eine Information über die Veränderungsrichtung abgeleitet werden kann. Dazu existieren verschiedene Methoden: 1. die Berechnung von Veränderungssektoren (MICHALEK et al. 1993),

2. die Berechnung von Veränderungskosinusrichtungen (CHEN et al. 2003), 3. eine Hauptkomponentenanalyse (PCA) im multitemporalen Merkmalsraum (LAMBIN & STRAHLER 1994) und 4. die Berechnung als mehrdimensionaler Raumwinkel in der polaren Domäne (ALLEN & KUPFER 2000). Die Darstellung der Veränderungsrichtungen in der polaren Domäne wurde ursprünglich für drei Eingangskanäle entwickelt, z.B. Tasseled Cap Brightness, Greenness, Wetness (ALLEN & KUPFER 2000). Dabei wird ein einziger Vektor erzeugt, der auf alle Achsen des Datenraums projiziert wird. Ein systematischer Ansatz zur Präsentation von Veränderungen in der polaren Domäne wurde von BOVOLO & BRUZZONE (2007) vorgestellt. Diese ursprünglich für zwei Spektralkanäle aufgezeigte Darstellungsform wurde später auf beliebig viele Kanäle erweitert (BOVOLO et al. 2010).

2.1 Robuste Change Vector Analyse (RCVA)

In der vorliegenden Arbeit wird das Verfahren von CASTILLA et al. (2009) auf die CVA übertragen, sodass in robuster Weise Veränderungsmagnitude und Veränderungsrichtung abgeleitet werden können. Für jedes Pixel mit der Position *j*,*k* wird in einem gleitenden Fenster der Größe 2*w+1 die minimale Differenz x_{diff} berechnet. Die Pixel des gleitenden Fensters sind mit *p*,*q* gekennzeichnet:

$$x_{diffa_i}(j,k) = \min_{(p\in[j-w,j+w],q\in[k-w,k+w])}\left\{\left(x_{2i}(j,k) - x_{1i}(p,q)\right) > 0\right\} \tag{3}$$

$$x_{diffb_i}(j,k) = \min_{(p\in[j-w,j+w],q\in[k-w,k+w])}\left\{\left(x_{1_i}(j,k) - x_{2_i}(p,q)\right) > 0\right\} \tag{4}$$

Damit das Ergebnis umkehrbar ist, müssen die Differenzen in beide Richtungen berechnet werden (x_{diffa} bzw. x_{diffb}), d.h. die Zeitpunkte werden vertauscht. Die Kombination zu den Differenzbildern für jeden Kanal *i* erfolgt nach:

$$x_{diff_i}(j,k) = \begin{cases} x_{diffa_i} > 0 \Rightarrow x_{diffa_i} \\ x_{diffa_i} \le 0 \Rightarrow 0 - x_{diffb_i} \end{cases} \tag{5}$$

Die Berechnung der Magnitude erfolgt wie in Gleichung 2, jedoch werden die robusten Differenzen eingesetzt. Die Veränderungsrichtungen werden als mehrdimensionaler Raumwinkel in Polarkoordinaten dargestellt (BOVOLO et al. 2010):

$$\alpha = \cos^{-1}\left[\frac{1}{\sqrt{n}}\left(\frac{\sum_{i=1}^{n} x_{diff_i}}{\sqrt{\sum_{i=1}^{n} x_{diff_i}^{\,2}}}\right)\right], \alpha \in [0,\pi] \tag{6}$$

2.2 Veränderungsprozesskette

Veränderungsdetektion sollte nicht nur als Algorithmus zur Unterscheidung von Veränderungen und Nicht-Veränderungen gesehen werden. Vielmehr ist es eine Prozesskette, die je nach Fragestellung, Datengrundlage und Untersuchungsgebiet individuell angepasst sein kann. Eine

grundsätzliche Abfolge ist jedoch jeder Veränderungsanalyse zugrunde liegend (THONFELD et al. 2010). Diese besteht meist aus der an die Fragestellung angepassten Auswahl der Daten, einer entsprechenden Vorprozessierung, dem Algorithmus zur Unterscheidung von Veränderung und Nicht-Veränderung, einem Schwellenwertverfahren, der thematischen Zuordnung der Veränderungen und einer Genauigkeitsanalyse der berechneten Karte. Entsprechend wurde auch in den vorliegenden Beispielen vorgegangen. Eine wichtige Voraussetzung der CVA ist eine möglichst exakte geometrische und radiometrische Vorprozessierung (JOHNSON & KASISCHKE 1998). Die geometrische Übereinstimmung von Bilddaten schwenkbarer Sensoren ist aus den erläuterten Gründen – individuelle Sonne-Sensor-Objekt-Geometrie, unterschiedliche Schattenanteile, nicht-identische Bodenabtastung des Sensors – nur bedingt zu realisieren und daher Gegenstand des hier erläuterten Verfahrens. Die radiometrische Übereinstimmung ist demnach auch nur mit Schwierigkeiten herzustellen, da identische Pixelpaare, anhand derer eine relative radiometrische Anpassung erfolgen kann, ebenfalls von den geometrischen Einflussgrößen betroffen sind. Hier muss die Grundannahme getroffen werden, dass die relative Normierung der Daten mit einem Nachbarpixel des zweiten Datensatzes auch möglich ist, sofern es die gleiche Landbedeckungsklasse repräsentiert, also unverändert ist. Ein automatisches Verfahren zur radiometrischen Normierung ist die Iteratively Re-Weighted Multivariate Alteration Detection (IR-MAD) (CANTY & NIELSEN 2008), die auch hier zur Anwendung kam. Dabei werden iterativ unveränderte Pixel identifiziert, anhand derer eine Regressionsgleichung aufgestellt wird. Diese werden dann genutzt, um die Bilder relativ aneinander anzupassen. Alternativ können absolute Verfahren der Atmosphärenkorrektur angewandt werden. Um konsistente Datensätze zu erhalten, müssten diese jedoch zusätzlich auch relativ aneinander angepasst werden (VICENTE-SERRANO et al. 2008). Zur Ableitung der Veränderungsinformation wurde die RCVA angewandt. Veränderungsmagnitude und -richtung stehen auf einer Ebene, jedoch hat es sich etabliert, zunächst einen Schwellenwert auf die Magnitude anzuwenden und anschließend die Veränderungsrichtung der als verändert identifizierten Pixel zu benennen. Unveränderte Pixel werden für die weiteren Analysen ausmaskiert. Die Unterscheidung von Veränderungen und Nicht-Veränderungen erfolgt durch Anwendung des vergleichsweise einfachen Schwellenwertverfahrens nach ROSIN (2001). Es handelt sich dabei um ein geometrisches Verfahren, bei dem zwischen Maximum und Ende des Histogramms der Veränderungsmagnitude eine Gerade gezogen wird. Der Punkt des Histogramms mit dem maximalen Abstand zur Geraden wird als Schwellenwert angenommen. Dieses Vorgehen entspricht weitgehend dem von CASTILLA et al. (2009). Die zusätzlichen Veränderungsrichtungen stellen jedoch einen deutlichen Informationsgewinn dar bei der Benennung der Veränderungen. Diese erfolgte durch ein Clustering der Veränderungsrichtungen für jene Pixel, deren Veränderungsmagnitude über dem bestimmten Schwellenwert liegt. Als Clustering-Verfahren wurde K-Means angewandt. Die Validierung der Ergebnisse ist eine große Herausforderung, da im Regelfall für keinen der beiden Zeitpunkte tagesaktuelle Landbedeckungsreferenzen vorliegen.

3 Veränderungsdetektion im urbanen Raum – Beispiel Köln

Zur Demonstration der Leistungsfähigkeit der RCVA wurden zwei Beispiele ausgewählt: Beim ersten handelt es sich um zwei RapidEye-Szenen aus dem urbanen Raum, die an zwei

aufeinanderfolgenden Tagen aufgenommen wurden und folglich keine Landbedeckungs- oder Landnutzungsveränderungen aufzeigen sollten. Beim zweiten wurden Daten unterschiedlicher Sensoren (RapidEye und Kompsat-2) und Aufnahmezeitpunkte genutzt, um Veränderungen zu detektieren. Die Aufnahmeparameter sind in Tab. 1 aufgeführt. Mit den zwei unterschiedlichen Testszenarien soll gezeigt werden, dass einerseits mittels RCVA Pseudoveränderungen besser unterdrückt werden können als mit CVA und andererseits reelle Veränderungen robust abgeleitet werden können. Da Verzerrungen insbesondere bei ausgeprägtem Relief auftreten, wurde das Stadtgebiet von Köln als Untersuchungsgebiet gewählt. Die RapidEye-Daten stammen von Ende Mai 2009 und liegen nur einen Tag auseinander. Von einem abweichenden Off-Nadir-Winkel abgesehen, wurden die Daten unter nahezu konstanten Bedingungen aufgenommen. Im Untersuchungsgebiet waren zum späteren Aufnahmezeitpunkt einzelne Wolken vorhanden. Die Daten wurden als Level 1B Format bestellt, mit einem digitalen Geländemodell mit 10 m räumlicher Auflösung orthorektifiziert und mittels bilinearer Interpolation auf eine Pixelgröße von 5 m entzerrt. Zur radiometrischen Anpassung der Daten wurde die IR-MAD angewandt. Als Masterszene wurde für alle Analysen die RapidEye-Szene vom 24.05.2009 genutzt. Zur Angleichung der räumlichen Auflösung von Kompsat-2 wurden die Daten mittels Nearest-Neighbor-Verfahren auf 5 m degradiert. Anschließend wurden sie mittels IR-MAD radiometrisch an die Master-Szene angepasst. Obwohl beide Aufnahmen etwa zur selben Jahreszeit aufgenommen wurden, gibt es zwischen der RapidEye-Szene vom 24.05.2009 und der Kompsat-2-Szene erhebliche Unterschiede in der Aufnahmegeometrie. Die Uhrzeiten des Überflugs liegen mehrere Stunden auseinander mit der Konsequenz, dass in den Kompsat-2-Daten deutlich größere Schattenanteile vorhanden sind. Die Off-Nadir-Winkel unterscheiden sich nur geringfügig. Allerdings wurden die RapidEye-Daten im *descending* Orbit aufgenommen, Kompsat-2 im *ascending* Orbit, sodass der Blickwinkel auf die Erdoberfläche ebenfalls unterschiedlich ist.

Von bewegten Objekten abgesehen (z.B. Autos oder Schiffe) sollte zwischen den beiden RapidEye-Aufnahmen keine Veränderung vorliegen. Unter der Annahme einer korrekten

Tab. 1: Parameter der analysierten Satellitendaten.

	RapidEye	RapidEye	Kompsat-2
Datum	24.05.2009	25.05.2009	04.06.2010
Uhrzeit (UTC)	11:29:16	11:28:29	08:57:35
Off-Nadir-Winkel	8,28 (Ost)	12,99 (Ost)	10,19 (Ost)
Orbit	descending	descending	ascending
Auflösung	6,5 m	6,5 m	4 m (1 m pan)
Blau	440-510 nm	440-510 nm	450-520 nm
Grün	520-590 nm	520-590 nm	520-600 nm
Rot	630-685 nm	630-685 nm	630-690 nm
Red Edge	690-730 nm	690-730 nm	
NIR	760-850 nm	760-850 nm	760-900 nm
Pan			500-900 nm
Grauwerttiefe	12 bit		10 bit

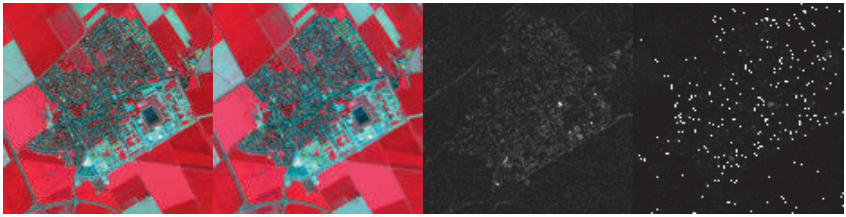


Abb. 1: Ausschnitt zweier RapidEye-Szenen (ca. 2x2 km²) am Stadtrand von Köln, aufgenommen am 24.05.2009 (links) und am 25.05.2009 (Mitte links). Die Magnitude der CVA (Mitte rechts) zeigt Pseudoveränderungen an Objektgrenzen und verstärkt innerhalb der Siedlung. Im Magnitudenbild der RCVA (3x3-Fenster) sind die Pseudoveränderungen unterdrückt.

radiometrischen Anpassung müsste demnach auch die Veränderungsmagnitude der CVA gleich null sein. Dies ist jedoch nicht der Fall. Es handelt sich nicht um eine konstante Differenz zwischen den beiden Aufnahmen. Vielmehr werden Veränderungen an Grenzen unterschiedlicher Landbedeckungsklassen impliziert, besonders stark ausgeprägt in urbanen Räumen. Bei Anwendung der RCVA werden diese Effekte deutlich unterdrückt. Die verbleibenden Unterschiede sind auf bewegte Objekte zurückzuführen und erhöhte Reflektionen an zum Sensor geneigten (Dach-)Flächen (Abb. 1). Die automatische Ableitung eines Schwellenwerts und anschließende Maskierung führt dazu, dass eine Anzahl isolierter Pixel oder kleiner Pixelgruppen als Veränderungsflächen verbleiben. Durch Anwendung einer kleinsten Kartiereinheit (Minimum Mapping Unit MMU) können auch diese Fehlzuweisungen bereinigt werden.
Im zweiten Beispiel sind die Aufnahmeparameter der verwandten Daten deutlich unterschiedlich (Tab. 1). Dementsprechend groß ist der Anteil an veränderten Pixeln im Bildausschnitt (helle

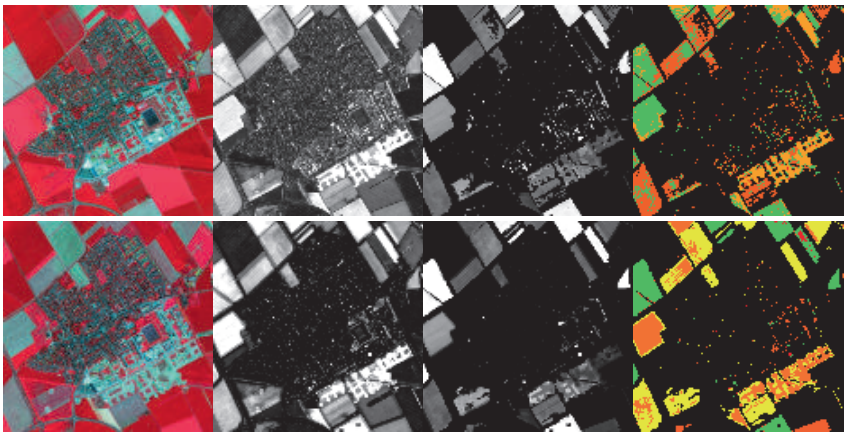


Abb. 2: Obere Zeile v.l.n.r.: RapidEye vom 24.05.2009, CVA Magnitude, CVA Richtung beschränkt auf die Veränderungsgebiete nach Anwendung eines Schwellenwertes, geclusterte Veränderungsrichtungen. Untere Bildzeile wie mittlere, aber für RCVA. Links unten: Kompsat-2 vom 04.06.2010.

Bildbereiche in Abb. 2, obere Bildzeile, Mitte links). Durch Berücksichtigung der Pixelnachbarschaft werden die aufnahmebedingten, durch Verzerrungen bewirkten Pseudoveränderungen unterdrückt. Nach Anwendung der individuellen Schwellenwerte und Clusterung der maskierten Richtungen entsteht bei der RCVA insgesamt ein realistischeres Szenario: die veränderten Flächen werden gut erfasst und es werden keine tatsächlichen Veränderungen verpasst. Die Clusterung der CVA Richtungen unterliegt zudem größeren Fehlern, vermutlich weil diverse Pseudoveränderungen als eigene Klassen erfasst werden und somit die tatsächlichen Veränderungen nicht mehr differenziert werden. Die Cluster-Ergebnisse der RCVA scheinen bei visueller Kontrolle sinnvoll, Felder erscheinen vergleichsweise homogen, Fruchtfolgen sind richtig zugewiesen.

4 Schlussfolgerung

Mit der Robusten Change Vector Analyse (RCVA) wurde ein Verfahren entwickelt, das durch die Berücksichtigung der Pixelnachbarschaft robust gegenüber aufnahmebedingten Verzerrungen, unterschiedlichen Sonne-Sensor-Objekt-Geometrien und Fehlregistrierungen ist. Die aufgeführten Anwendungsbeispiele zeigen, dass der Anteil fehlerhafter Veränderungszuweisungen deutlich reduziert werden kann. Insbesondere bei stark ausgeprägter dritter Oberflächendimension ist es sinnvoll, die Pixelnachbarschaft in die Berechnung der Veränderung einfließen zu lassen. Der Schwerpunkt der hier vorgestellten Methode liegt auf der Optimierung der räumlichen Erfassung von Veränderungen und Nicht-Veränderungen. Die Klassifizierung der Veränderungen wurde in diesem Zusammenhang nicht näher untersucht, kann aber durch andere Methoden weiter verbessert werden.

Danksagung

Die Arbeiten erfolgten innerhalb des Projekts ENVILAND 2, das durch die Raumfahrt-Agentur des Deutschen Zentrums für Luft- und Raumfahrt e.V. (DLR) mit Mitteln des Bundesministeriums für Wirtschaft und Technologie aufgrund eines Beschlusses des Deutschen Bundestages unter den Förderkennzeichen 50EE0844 - 50EE0847 gefördert wurde. Wir danken dem DLR für die Bereitstellung von Produkten aus dem RapidEye Science Archive unter der Projektnummer 272. Unser Dank gilt zudem der Europäischen Raumfahrtbehörde ESA für die Bereitstellung der Kompsat-2-Daten.

5 Literaturverzeichnis

ALLEN, T. R. & KUPFER, J. A., 2000: Application of Spherical Statistics to Change Vector Analysis of Landsat Data: Southern Appalachian Spruce–Fir Forests. Remote Sensing of Environment, **74** (3), S. 482–493.

BOVOLO, F., MARCHESI, S. & BRUZZONE, L., 2010: A nearly lossless 2d representation and characterization of change information in multispectral images. IEEE Geoscience and Remote Sensing Symposium (IGARSS), S. 3074 –3077.

BOVOLO, F. & BRUZZONE, L., 2007: A Theoretical Framework for Unsupervised Change Detection Based on Change Vector Analysis in the Polar Domain. IEEE Transactions on Geoscience and Remote Sensing, **45** (1), S. 218–236.

BRUZZONE, L. & COSSU, R., 2003: An adaptive approach to reducing registration noise effects in unsupervised change detection. IEEE Transactions on Geoscience and Remote Sensing, **41** (11), S. 2455–2465.

CANTY, M. J. & NIELSEN, A. A. 2008: Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. Remote Sensing of Environment, **112** (3), S. 1025–1036.

CASTILLA, G., GUTHRIE, R.H. & HAY, G. J., 2009: The land-cover change mapper (LCM) and its application to timber harvest monitoring in Western Canada. Photogrammetric Engineering & Remote Sensing, **75** (8), S. 941–950.

CHEN, J., GONG, P., CHUNYANG, H., PU, R. & SHI, P., 2003: Land-Use/Land-Cover Change Detection Using Improved Change-Vector Analysis. Photogrammetric Engineering & Remote Sensing, **69** (4), S. 369–379.

COPPIN, P., JONCKHEERE, I. , NACKAERTS, K., MUYS, B. & LAMBIN, E., 2004: Digital change detection methods in ecosystem monitoring: a review. International Journal of Remote Sensing, **25** (9), S. 1565–1596.

DAI, X. & KHORRAM, S., 1998: The effects of image misregistration on the accuracy of remotelysensed change detection. IEEE Transactions on Geoscience and Remote Sensing **36** (5), S. 1566–1577.

IM, J. & JENSEN, J. R., 2005: A change detection model based on neighborhood correlation image analysis and decision tree classification. Remote Sensing of Environment, **99** (3), S. 326–340.

JOHNSON, R. D. & KASISCHKE, E. S., 1998: Change vector analysis: a technique for the multispectral monitoring of land cover and condition. International Journal of Remote Sensing, **19** (3), 411.

LAMBIN, E. F. & STRAHLER, A. H., 1994: Change-vector analysis in multitemporal space: a tool to detect and categorize land-cover change processes using high temporal-resolution satellite data. Remote Sensing of Environment, **48** (2), S. 231–244.

MALILA, W. A., 1980: Change vector analysis: an approach for detecting forest changes with Landsat. Proceedings of the 6th Annual Symposium on Machine Processing of Remotely Sensed Data, Purdue University, Indiana, S. 329–335.

MICHALEK, J. L., WAGNER, T. W., LUCZKOVICH, J. J.& STOFFLE, R. W., 1993: Multispectral change vector analysis for monitoring coastal marine environments. Photogrammetric engineering and remote sensing, **59** (3), S. 381–384.

ROSIN, P. L., 2001. Unimodal thresholding. Pattern Recognition, **34** (11), S. 2083–2096.

THONFELD, F., Hecheltjen, A., Braun, M. & Menz, G., 2010: From algorithms to processing chains: A review of land cover and land use change detection methodologies. Proceedings of ESA Living Planet Symposium, Bergen, CD-ROM.

VICENTE-SERRANO, S. M., PÉREZ-CABELLO, F. & LASANTA, T., 2008: Assessment of radiometric correction techniques in analyzing vegetation variability and change using time series of Landsat images. Remote Sensing of Environment, **112** (10), S. 3916–3934.

VIRAG, L.A. & COLWELL, J.E., 1987: An improved procedure for analysis of change in Thematic Mapper image-pairs. Ann Arbor, Michigan, S. 1101–1110.

Detektion von Fensterverdachungen und Gesims in Fassadenbildern und Punktwolken

WOLFGANG BRANDENBURGER[1], MARTIN DRAUSCHKE[2] WILLIAM NGUATEM[2] & HELMUT MAYER[2]

Zusammenfassung: In dieser Arbeit wird ein Ansatz für die automatische Detektion von Fensterverdachungen und Gesims vorgestellt, der auf der Analyse eines entzerrten Fassadenbildes in Kombination mit einer 3D-Punktwolke erfolgt. Die Punktwolke wird dabei aus einer Bildsequenz abgeleitet, z.B. nach (MAYER et al. 2012). Ausgehend von der Detektion der Fassadenwand als Ebene und der Bestimmung ihrer Begrenzung wird eine Fensterdetektion verwendet, um in einem möglichst kleinen Suchraum die Existenz von Verdachungen zu überprüfen. Neben der geometrischen Analyse werden auch detektierte Kanten ausgewertet. Das in diesem Aufsatz beschriebene Verfahren kann zuverlässig bogenförmige Fensterverdachungen von Dreiecksgiebeln und Gesims unterscheiden. Das wird in Experimenten auf Fassaden aus Freising, München und Schwabach gezeigt.

1 Einleitung

Mittels 3D-Stadtmodelle werden urbane Räume modelliert. Sie werden in diversen Bereichen eingesetzt, z.B. im Tourismus, in der Architektur oder in der Stadt-und Raumplanung. In diesen Bereichen wird eine realitätsnahe Darstellung mit einem hohen Detaillierungsgrad gefordert. Derzeitige 3D-Stadtmodelle weisen aber in aller Regel nur eine geringe Detaillierung auf. Die Modellierung auffälliger Gestaltungsmittel der Architektur wie Fensterverdachungen und Gesimse auf Gebäuden mit historischen Fassaden ist mit einem erheblichen manuellen Arbeitsaufwand verbunden und wird daher (wenn überhaupt) nur für einzelne Sehenswürdigkeiten vorgenommen.

LEBERL et al. (2012) leiten größere detailreiche 3D-Modelle von urbanen Räumen aus diversen Datenquellen ab, wobei Luftbilder für die grobe Modellierung des städtischen Raumes sowie Bilder und LIDAR-Daten von Fahrzeugen verwendet werden. Letztere werden zur Modellierung von Details verwendet. MAYER et al. (2012) kombinieren Bodenaufnahmen und Bilder von Drohnen / Unmanned Aircraft Systems (UAS) zur Generierung von dichten Punktwolken, aus denen detaillierte Oberflächen und grobe Gebäudestrukturen (NGUATEM et al., 2012) konstruiert werden.

Primäre Architekturglieder wie Fenster, Eingänge und schwach ausgeprägte 3D-Strukturen werden häufig unter Verwendung entzerrter Bilder und auf Basis einer baustiltypischen Grammatik detektiert, wie z.B. bei (SIMON et al., 2011). Für die Detektion von Detailstrukturen wie Fensterkreuze oder Ornamente an der Fassade werden zusätzlich noch LIDAR-Punktwolken verwendet, so z.B. bei (BECKER & HAALA, 2008), (HOHMANN et al., 2009) oder (SCHMITTWILKEN & PLÜMER, 2010).

[1] Wolfgang Brandenburger, Amt für Geoinformationswesen der Bundeswehr, Dezernat III 1 (1) Geländeanalyse, Frauenberger Straße 250, 53879 Euskirchen, Email: wolfgangbrandeburger@bundeswehr.org

[2] Martin Drauschke, William Nguatem, Helmut Mayer, Universität der Bundeswehr München, Institut für Angewandte Informatik, Werner-Heisenberg-Weg 39, 85577 Neubiberg, Email: martin.drauschke@unibw.de, william.nguatem@unibw.de, helmut.mayer@unibw.de

In dieser Arbeit wird ein Ansatz für die Detektion auffälliger architektonischer Gestaltungsmittel wie Fensterverdachungen und Gesimse vorgestellt. Zunächst wird die allgemeine Vorgehensweise für die Fassadeninterpretation aufgezeigt (siehe Abb. 1), danach wird das neue Verfahren für die Detektion von Fensterverdachungen und Gesims präsentiert. Abschließend werden Ergebnisse für Fassaden aus dem Großraum München vorgestellt und die Performanz des Verfahrens diskutiert.

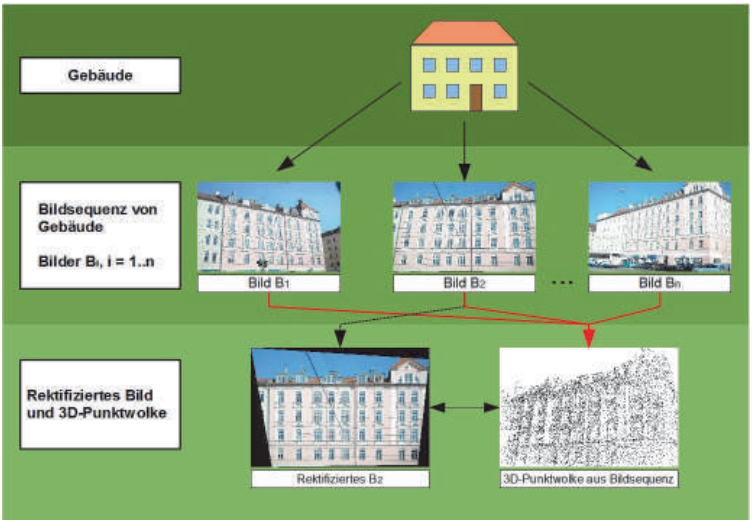


Abb. 1: Konzept für die Fassadeninterpretation mit Bildsequenz und Punktwolke.

2 Konzept für die Fassadeninterpretation

Datengrundlage (siehe Abb. 1) für die Fassadeninterpretation bilden mit handelsüblichen Kameras aufgenommene terrestrische Bildsequenzen. Mit Hilfe eines Rekonstruktionsverfahrens (MAYER et al., 2012) wird aus der Bildsequenz eine mittelmäßig dichte, aber sehr präzise Punktwolke abgeleitet. Neben diesen 3D-Punkten wird auch eine auf die Fassadenfläche entzerrte Aufnahme der Bildsequenz für die Weiterverarbeitung verwendet. Dazu werden zunächst die Lage und die Ausdehnung der Fassadenwand bestimmt. Hierfür eignet sich eine Ebenensuche in der 3D-Punktwolke mittels RANSAC (FISCHLER & BOLLES, 1981). Diese wird durch eine vorhergehende Segmentierung des Bildes ergänzt. Danach wird ein Bild mit möglichst senkrechter Blickrichtung auf die Fassade mittels einer Homographie so entzerrt, dass die Fassadenfläche im Bild rechtwinklig erscheint.
In entzerrten Fassadenbildern können durch Gradientenanalyse nach (LEE & NEVATIA, 2004) oder Implicit Shape Models nach (REZNIK & MAYER, 2008) Fenster erkannt werden. Dadurch kann die Suche nach Fensterverdachungen und Gesims auf kleine Bereiche oberhalb der Fenster eingeschränkt werden. Besonders häufig kommen die flache Gesimse, Dreiecksgiebel sowie

bogenförmige Fensterverdachungen vor (siehe Abb. 2). Deshalb beschränkt sich die Detektion von Verdachungen auf diese drei Erscheinungsformen.



Abb. 2: Verdachungsformen: Fenstergesims (links), Dreiecksgiebel (Mitte) und Bogensegment (rechts).

3 Umsetzung

In dieser Arbeit liegt der Fokus auf der Detektion der Verdachungen in Fassadenbildern und daraus rekonstruierten Punktwolken. Grundlage sind die in Abschnitt 2 beschriebenen Schritte. Diese können vollautomatisch durchgeführt werden. Um das entwickelte Verfahren unabhängig von den vorhergehenden Arbeitsschritten testen zu können, wurden die Vorarbeiten überwacht und teilweise manuell korrigiert.

3.1 Detektion der Fassadenwand sowie der Fenster

Unter Angabe des Maßstabs einer Punktwolke ist es möglich, die rekonstruierten 3D-Punkte in unmittelbarer Nähe der in Abschnitt 2 detektierten Wand von allen anderen zu unterscheiden. Damit ist eine Abgrenzung der Fassade im entzerrten Bild in erster Näherung möglich. Im nächsten Schritt wird der korrespondierende Bildausschnitt in zwei Bereiche segmentiert, wobei die Varianz des Abstands der Punkte des Umfelds von der Ebene und die Intensitätswerte im Bild berücksichtigt werden. Als Resultat erhält man die Bereiche der Fassade, die exakt auf der Fassade liegen (siehe Abb. 3). So lassen sich 3D-Punkte, die auf Fenster oder Ornamenten liegen, für die Verfeinerung der Modellierung der Fassadenebene ausschließen.
Nachdem die Fassadenebene und ihre Abgrenzung bekannt sind, wird nach den Fenstern im entzerrten Fassadenbild gesucht. Unter Annahme einer regelmäßigen Anordnung der Fenster innerhalb einer Fassade wird der Fensterdetektor von LEE & NEVATIA (2004) eingesetzt.

3.2 Detektion der Fensterverdachungen

3.2.1 Objektmodelle für Fensterverdachungen

Fensterverdachungen und Gesimse sind plastisch vor die Außenwand hervortretende Bauglieder einer Fassade, die sich direkt oberhalb von Fenstern befinden. Sie dienen allein der Dekoration und sollen die Fassade durch Schattenbildung beleben (KOEPF & BINDING, 2005). Im Allgemeinen steht die Größe einer Verdachung im direkten Zusammenhang mit den zugehörigen Fenstern: Sie sind etwas breiter. Des Weiteren haben sie eine moderate Höhe. Diese Rahmenbedingungen bzgl. der Lage vor und innerhalb einer Fassade werden beim entwickelten Detektionsverfahren ausgenutzt.



Abb. 3: Segmentierung der Bereiche direkt auf der Fassade. Blau markierte Bereiche zeigen die Fassadenwand, d.h. sie weisen 3D-Punkte mit sehr niedriger Streuung um 3D-Ebene auf.

Die drei gängigsten Formen für Fensterverdachungen und Gesims sind in Abb. 2 dargestellt. Es handelt sich um das Fenstergesims, bestehend aus einer horizontalen Bedachung des Fensters, um den Dreiecksgiebel (gleichschenkliges Dreieck) und um das Bogensegment, das in dieser Arbeit durch einen Kreisbogen approximiert wird.

3.2.2 Suchbereich für Fensterverdachungen

Die detektierten Fenster werden entsprechend ihrer Anordnung mit zwei Indices i und j bezeichnet. Sei das Fenster F_{ij} mit den Parametern (x_{ij}, y_{ij}, B_{ij}, H_{ij}) für seine Lage (Schwerpunkt) und Ausdehnung (Breite und Höhe) gegeben. Dann wird der Suchbereich für eine Fensterverdachung dieses Fenster F_{ij} durch seine beiden horizontalen Nachbarn F_{ij-1} und F_{ij+1} sowie das darüber liegende Fenster $F_{i-1,j}$ definiert. Bei Fenstern ohne Nachbarn wird der Suchbereich bis zum Fassadenrand ausgedehnt. Abb. 4 zeigt einen Suchbereich.

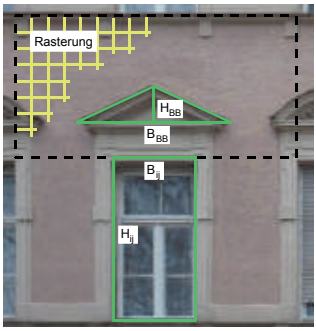


Abb. 4: Fenster und Suchbereich mit Rasterung und Teile der Parameter für Fenster und Dreiecksgiebel.

3.2.3 Erkennung aufgrund der 3D-Geometrie

Es wurde untersucht, ob allein die Analyse der Punkte bzgl. ihrer Lage zur Fassadenebene für die Erkennung von Fensterverdachungen und Gesims ausreicht. Dazu wird der oben definierte Suchbereich diskretisiert, und zwar so, dass die Fensterbreite B_{ij} der Breite für zehn quadratische

Rasterfelder entspricht. In den Rasterfeldern wird die Punktdichte bestimmt und mit der über die gesamte Fassade berechneten Punktdichte statistisch verglichen, um Hinweise auf auffällige Strukturen zu erhalten. Anschließend werden die Ergebnisse mittels eines morphologischen Closings verbessert. Die zusammenhängenden Strukturen mit einer Breite B_{BB} und einer Höhe H_{BB} der zugehörenden Bounding Box werden hinsichtlich der geometrischen Übereinstimmung mit dem zugrunde liegenden Fenster bewertet. Wenn gilt

$$B_{ij} \leq B_{BB} \leq 2B_{ij}, \qquad H_{BB} \leq \frac{1}{2} H_{ij},$$

wird die gefundene Struktur als Fensterverdachung bewertet.

3.2.4 Erkennung aufgrund von Kanten

Für die Erkennung von Fensterverdachungen und Gesims mittels Kanten werden diese im oben definierten Suchbereich mittels Canny-Operator (CANNY, 1986) extrahiert. Aufgrund starker radiometrischer Schwankungen in den Bildern wird ein sehr niedriger Schwellwert zur Konturverfolgung verwendet. Um robuste Ergebnisse zu erhalten werden mittels Tiefpassfilterung feine Strukturen unterdrückt.
In den Kantenbildern werden Geradenstücke und Kreisbögen gesucht (siehe Abb. 5). Mittels RANSAC (FISCHLER & BOLLES, 1981) werden die Parameter über eine zufällige Auswahl von Kantenpunkten geschätzt. Dabei sollen ausschließlich bedeutsame Geradenstücke und Kreisbögen ermittelt werden. Bedeutsame Strukturen liegen vor, wenn sowohl die Menge der ein Modell unterstützenden Punkte als auch die modellierte Instanz eine ausreichende Größe haben. Diese Mindestzahl von Punkten wird in Abhängigkeit von der Fenstergröße (B_{ij}, H_{ij}) festgelegt.



Abb. 5: RANSAC-Algorithmus zur Suche bedeutsamer Strukturen im Kantenbild. Links: Bildausschnitt mit bogenförmiger Verdachung. Mitte: Ergebnis der Kantendetektion mit verschiedenen Schwellwerten. Rechts: Ergebnis der Suche nach Geraden (oben) bzw. nach Kreisbögen (unten).

Es werden alle Geradenstücke entfernt, die einen Winkel betragsmäßig größer 45° zur Horizontalen aufweisen. Dieser Wert rührt daher, dass in der Regel nur stumpfwinklige Dreiecksgiebel vorkommen und vertikale Geraden bei der Unterscheidung der verschiedenen Fensterverdachungen keine Rolle spielen. Ebenso werden alle Kreisbögen entfernt, die keine unterstützenden Punkte im Bereich des oberen Scheitelpunkts haben.
Mittels eines Entscheidungsbaums wird nun klassifiziert, ob eine Fensterverdachung vorliegt und ob es sich um ein Gesims, einen Dreiecksgiebel oder ein Bogensegment handelt. Für ein Bogensegment muss der Mittelpunkt des aus dem Kreissegment abgeleiteten Kreises in einem Toleranzbereich um die vertikale Mittelachse des entsprechenden Fensters liegen. Bei einem Dreiecksgiebel müssen die beiden detektierten Geradenstücke annähernd gleiche Winkel mit

unterschiedlichem Vorzeichen haben und der gemeinsame Schnittpunkt in einem Toleranzbereich um die vertikale Mittelachse des entsprechenden Fensters liegt. Für ein Gesims müssen zwei horizontale Geraden detektiert werden, deren Mitte in einem Toleranzbereich um die vertikale Mittelachse des entsprechenden Fensters liegen. In allen drei Fällen wird der zulässige Toleranzbereich um die vertikale Mittelachse in Abhängigkeit von der Perspektive des Ausgangsbildes gewählt.

3.2.5 Kombination von 3D-Geometrie und Kanten

Die Performanz beider Ansätze zur Detektion von Verdachungen wird stark durch die Wahl der Parameter beeinflusst. Die Parameter können so eingestellt werden, dass keine Falschdetektionen erfolgen, die z.B. durch Schatten oder auf der Fassade gemalten Fensterverdachungen verursacht werden. Dann werden aber auch nicht alle korrekten Hypothesen als Verdachung erkannt. Oder die Parameter werden schwächer gestellt, so dass die Erkennungsrate steigt, aber auch die Zahl der Falschdetektionen. Die Zahl der Falschdetektionen kann dann verringert werden, wenn man beide Verfahren kombiniert. Dieses Zusammenführen wird durch eine logische Und-Operation auf den binären Ergebnisbildern der Einzelverfahren umgesetzt.

4 Experimente

4.1 Daten

Zehn Bildsequenzen von je einer Fassade zeigen meist historische, mittelgroße Fassaden in den Innenstädten Freisings, Münchens und Schwabachs. Sie umfassen zwischen sieben und 14 Bilder und wurden mittels des Verfahrens von (MAYER et al., 2012) orientiert. Die Punktwolken umfassen ca. 9.000 Punkte und die Punktdichte auf den Fassadenwänden liegt häufig bei ca. 20-50 Punkten pro m^2.

4.2 Ergebnisse

In Tab. 1 sind die Resultate der Experimente dargestellt. Die in den Bildsequenzen abgebildeten Fassaden zeigen insgesamt 95 Fenster (ohne Schaufenster und ohne Gaubenfenster). Von diesen haben 16 keine Verdachung, oberhalb von 79 Fenstern sind die Fassaden entweder durch Bogensegmente, Dreiecksgiebel oder Fenstergesimse verziert.

Tab. 1: Ergebnisse der Detektion von Fensterverdachungen und Gesims. Bei Versuch 1 wurde ausschließlich untersucht, ob Fassadenelemente vor der Fassade liegen. Bei Versuch 2 wurde starke Kanten für die Erkennung der Objekte untersucht. Beim dritten Versuch wurden beide Verfahren kombiniert.

	Keine Verdachung	Verdachungen insgesamt	Bogensegment	Dreiecksgiebel	Fenstergesims
Versuch	16	79	39	21	19
1: Geom.	16 (100%)	52 (66%)	29 (74%)	16 (76%)	7 (37%)
2: Kanten	16 (100%)	50 (63%)	32 (82%)	18 (86%)	0 (0%)
3: G + K	16 (100%)	61 (77%)	32 (82%)	18 (86%)	11 (58%)

Von allen drei untersuchten Verfahren wird zuverlässig erkannt, wenn keine Fensterverdachungen existieren. Die Erkennung von vorhandenen Fensterverdachungen ist

dagegen weniger erfolgreich. Die Ergebnisse sind am besten, wenn Informationen aus geometrischer Analyse und Kantendetektion miteinander verknüpft werden. Die Erkennung von Bogensegmenten und Dreiecksgiebel ist wesentlich erfolgreicher als die Erkennung von Fenstergesimsen. Dies liegt vor allem an den charakteristischen Formen der beiden Fensterverdachungen. Dagegen weisen Fenstergesims oft keine starken 3D-Strukturen auf, und horizontale Kanten sind keine zuverlässige Evidenz für Gesimse bilden.
Wenn Verdachungen nicht erkannt wurden, liegt das häufig daran, dass die rekonstruierte Punktwolke in diesen Bereichen zu wenige 3D-Punkte vor der Fassade aufweist. Der Misserfolg bei der Erkennung von Gesimsen mittels Kanten liegt an der Wahl einiger Parameter.

5 Diskussion und Ausblick

Es wurde ein Verfahren zur Detektion von Fensterverdachungen und Gesims in Fassadenbildern vorgestellt. Die Fassadenbilder werden zusammen mit präzisen Punktwolken aus Bildsequenzen abgeleitet. Die Fassadenbilder sind so entzerrt, dass die Strukturen der Fassade wieder eine rechteckige Form zeigen. Aufbauend auf einer präzisen Extraktion der Fassadenebene sowie einer zuverlässigen Fensterdetektion, durch die der Suchbereich eingeschränkt wird, werden Fensterverdachungen und Gesimse erfolgreich erkannt. In den Experimenten wurde gezeigt, dass die alleinige Ausnutzung der Information, dass zahlreiche 3D-Punkte im Suchbereich vor der Fassadenebene liegen, für eine zuverlässige Erkennung nicht ausreicht. Ähnlich ungeeignet ist die alleinige Betrachtung detektierter Kanten. Dagegen liefert die Kombination beider Informationen zufriedenstellende Ergebnisse. Abb. 6 zeigt die Ergebnisse bzgl. der Detektion mittels 3D-Geometrie und Kanten für zwei repräsentative Fassaden aus Freising. Die eingezeichneten Modelle verdecken dabei häufig die auffälligsten Bildstrukturen der Fensterverdachungen.



Abb. 6: Ergebnisse der Detektion von Verdachungen bei zwei Bildern von Fassaden aus Freising. Beide Bildausschnitte wurden entsprechend entzerrt und ausgeschnitten. Neben den Verdachungsformen für Bogensegmente, Dreiecksgiebel und Gesims werden auch die vor der Fassadenwand liegenden Punkte angezeigt. Diese liegen vor allem im Bereich der Verdachungen.

Die Erkennung der Fensterverdachungen und Gesimse kann verbessert werden, indem eine extrem dichte Punktwolke verwendet wird, wie sie z.B. durch das Semiglobale Zuordnungsverfahren von HIRSCHMÜLLER (2008) erzeugt wird. Für eine Automatisierung der

Detektion der Fassadenebene, der Bestimmung ihrer Begrenzung, die Auswahl und Entzerrung eines Bildes für die Detailanalyse sowie die Fensterdetektion im entzerrten Bild kann diese Arbeit in die Prozesskette von NGUATEM et al. (2012) integriert werden.

6 Literaturverzeichnis

BECKER, S. & HAALA, N., 2008: Integrated LIDAR and Image Processing for the Modelling of Building Facades. Photogrammetrie – Fernerkundung – Geoinformation, **2008** (2), S. 65-81.

CANNY, J., 1986: A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, **8** (6), S. 679-698.

FISCHLER, M.A. & BOLLES, R.C., 1981: Random Sample Consensus: A Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography. ACM, **24** (3), S. 381-395.

HISCHMÜLLER, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. IEEE Transactions on Pattern Analysis and Machine Intelligence, **30** (2), S. 328-341.

HOHMANN, B., KRISPEL, U., HAVEMANN, S. & FELLNER, D., 2009: Cityfit: High-Quality Urban Reconstruction by Fitting Shape Grammars to Images and Derived Textured Point Clouds. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **38** (5/W1).

KOEPF, H. & BINDING, G., 2005: Bildwörterbuch der Architektur. Vierte, überarbeitete Auflage, Alfred Kröner Verlag, Stuttgart. 678 Seiten.

LEBERL, F., MEIXNER, P., WENDEL, A. & IRSCHARA, A., 2012: Automated Photogrammetry for Three-Dimensional Models of Urban Spaces. Optical Engineering, **51** (2), S. 1-12.

LEE, S.C. & NEVATIA, R., 2004: Extraction and Integration of Window in a 3D Building Model from Ground View Images. IEEE Conference on Computer Vision and Pattern Recognition, S. II: 113-120.

MAYER. H., BARTELSEN, J. HIRSCHMÜLLER, H. & KUHN, A., 2012: Dense 3D Reconstruction from Wide Baseline Image Sets. Outdoor and Large-Scale Real-World Scene Analysis – 15th International Workshop on Theoretical Foundations of Computer Vision, Lecture Notes in Computer Science 7474, S. 285-304.

NGUATEM, W., DRAUSCHKE, M. & MAYER, H., 2012: Finding Cuboid-Based Building Models in Point Clouds. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **39** (B3), S. 149-154.

REZNIK, S. & MAYER, H., 2008: Implicit Shape Models, Self-Diagnosis, and Model Selection for 3D Facade Interpretation. Photogrammetrie – Fernerkundung – Geoinformation, **2008** (3), S. 187-196.

SCHMITTWILKEN, J. & PLÜMER, L., 2010: Modelbased Reconstruction and Classification of Facade Parts in 3D Point Clouds. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **38** (3A), S. 269-274.

SIMON, L., TEBOUL, O., KOUTSOURAKIS, P. & PARAGIOS, N., 2011: Random Exploration of the Procedural Space for Single-View 3D Modeling of Buildings. International Journal of Computer Vision, **93** (2), S. 253-271.

Traditionelle Luftbilder in 3D

UTA GRIWODZ[1] & MARKUS BELLINGHOFF[2]

Zusammenfassung: Traditionelle Luftbilder liegen heute zahlreich als zweidimensionale Rasterdaten vor - sowohl aktuelle als auch historische Luftbilder mit interessanten Informationen für die Gegenwart (historische Bach- und Flussläufe, Gebietsentwicklungen, etc.). Gleichzeitig werden heutzutage immer mehr Sensordaten erfasst, z.B. umfangreiche Punktwolken. Die sehr großen Punktdatenmengen enthalten zwar viele Informationen, welche aber vom normalen Betrachter und in Standardviewern und -formaten kaum betrachtet, erkannt und genutzt werden können. So werden Sensordaten zumeist nur von Experten genutzt. Luftbildern wiederum fehlt die 3D-Komponente, die für viele Betrachtungen und Analysen interessant und hilfreich wäre. In diesem Beitrag wird eine einfache Möglichkeit aufgezeigt, beide Datenquellen mit Hilfe von Standard Spatial ETL-Werkzeugen zu kombinieren. Es wird gezeigt, wie Punktwolken auf der Basis von Luftbildern mit realen Farben versehen und wie Luftbilder mit Höhenwerten angereichert werden können, so dass der Eindruck entsteht, man sähe Luftbilder in 3D. Es werden Anwendungsszenarien und Zieldatenformate erörtert. So können diese Daten z.B. ins 3D-PDF-Format geschrieben und so für alle Interessierte zugänglich gemacht werden, ohne dass eine Spezialsoftware benötigt würde. Der Beitrag beschäftigt sich neben der Erzeugung ebenfalls mit den Mehrwerten aber auch Grenzen, die solche realistisch aussehenden Punktwolken bzw. Luftbilder mit 3D-Eindruck bieten können.

1 Einleitung / Ausgangssituation

Luftbilder liegen heute zahlreich vor. Die meisten Städte und Gemeinden lassen ihre Gebiete in unterschiedlichem Turnus regelmäßig befliegen. Es gibt aber auch Angebote über flächendeckende Luftbilder für ganz Deutschland und darüber hinaus. Durch die regelmäßigen Befliegungen können durch den Vergleich der historischen zu den aktuellen Luftbildern auch Entwicklungen abgelesen werden, wie z.B. historische Bach- und Flussläufe, Gebietsentwicklungen, etc.. In vielen Fällen werden diese zweidimensionalen Raster aber nur als Hintergrundbilder genutzt. Vielen Menschen fällt das Lesen und Interpretieren von gewöhnlichen Karten schwer. Manche Menschen können sich auf Grund der realistischeren Ansicht besser anhand von Luftbildern orientieren.
Immer gewohnter wird und ist die Nutzung von 3D-Informationen: in Computerspielen und Apps navigiert der Nutzer durch 3D-Welten. Die Navigation und Orientierung in der 3D-Welt ist inzwischen etwas Selbstverständliches, unsere Kinder wachsen damit auf.
Diese 3D-Komponente fehlt den traditionellen Luftbildern aber.

1) Uta Griwodz, con terra GmbH, Martin-Luther-King-Weg 24, 48455 Münster, E-Mail: u.griwodz@conterra.de
2) Markus Bellinghoff, con terra GmbH, Martin-Luther-King-Weg 24, 48455 Münster, E-Mail: m.bellinghoff@conterra.de

Parallel dazu werden heutzutage immer mehr Sensordaten erfasst, die großen Punktwolken erzeugen. Diese Daten beinhalten die 3D-Komponente. Aufgrund der großen Punktmengen sind sie aber mit den herkömmlichen Werkzeugen nur schwer zu bearbeiten und zu betrachten. Daher werden die Sensordaten meist nur von Experten genutzt.
Eine Kombination von Luftbildern und Sensordaten sowie die Bereitstellung für einen breiten Anwenderkreis ist daher wünschenswert.

2 Sensordaten, Laserscandaten, LIDAR

Es gibt eine Vielzahl von Sensordaten und -technologien, die in der Fernerkundung zum Einsatz kommen. Hierzu gehören unter anderem Photographische analoge Aufnahmesysteme, Panchromatische Filme (inkl. NIR), Farbfilme (inkl. NIR), Digitale Photosysteme (CCD) und Abtastsysteme (Scanner), Satellitenscanner, Hyperspektrale Spektrometer und Radarsysteme.

Eine dieser Möglichkeiten ist das Laserscanning. Bei diesem Verfahren werden Oberflächen oder Körper mittels Lasertechnologie abgetastet. Dabei existieren verschiedene Vorgehensweisen und Anwendungsgebiete.
Ein System, das die Laserlichtquelle nutzt, ist LIDAR. Der Begriff "LIDAR" ist eine Abkürzung für "Light Detection and Ranging".
Die Grundlagen für LIDAR Mapping sind
a) die genaue Lokalisierung der Position des Sensors, und
b) die Zeitmessung dafür, wie lange es dauert, bis der Lichtimpuls auf ein Objekt trifft, dort reflektiert wird und zurück zum Sensor kommt. Der reflektierte Lichtimpuls wird vom Sensor erfasst. Wenn der Laserimpuls dabei z.B. auf ein "festes" Objekt trifft (wie ein Gebäude, Fahrzeug, Pflaster, einen Felsen oder nackten Boden) wird die auftreffende Energie komplett reflektiert und nur ein Reflexionswert ("Return") vom Sensor aufgezeichnet. Dies kann allerdings auch für sehr dichte Vegetation gelten. Wenn die Vegetation aber nicht so dicht ist, kann das Laserlicht weiterreisen, z.B. zwischen einzelnen Blättern und Ästen hindurch, teilweise sogar bis zum Erdboden. In einem solchen Fall gibt es dann mehrere Reflexions („Return“)-Werte für einen Ort.

Diese LIDAR-Daten (vor einer Klassifizierung) sind im Wesentlichen eine Wolke von 3D-Punkten von "single returns" oder "multiple returns" von Laserimpulsen. Der Effekt der Reflexion wird in folgender Abbildung veranschaulicht.

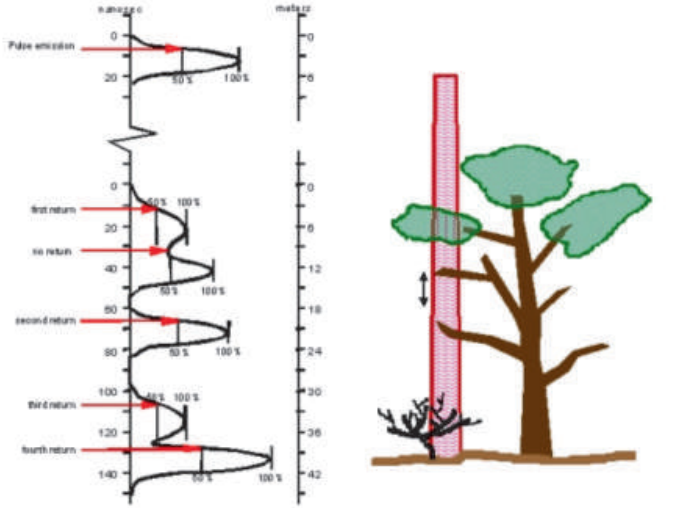


Figure 1.05: Multiple lidar returns can be generated from a single emitted pulse. Depending on the detection capabilities of the lidar sensor, two or more of these returns may be recorded as data points.

SOURCE: ASPRS

Abb. 1: Reflexionswerte bei LIDAR-Daten

In Abhängigkeit vom verwendeten System werden unterschiedlich viele Reflexionswerte aufgezeichnet. Je mehr Reflexionswerte aufgezeichnet werden, desto größer wird der resultierende Datensatz. LIDAR-Datensätze sind allein schon aufgrund der hohen emittierten Impulsfrequenz sehr umfangreich. Die ersten Reflexionswerte können genutzt werden, um digitale Geländemodelle zu erstellen, die die Objekte oberhalb der Erdoberfläche enthalten, wie Gebäude, Brücken, Baumkronen, etc. (Bedeckungsprofil). Mittlere Reflexionswerte werden genutzt, um die Vegetation von festen oberirdischen Objekten zu unterscheiden. Die letzten Reflexionswerte liefern die erste Näherung für die Erdoberfläche (Bodenprofil).

Zusätzlich zur Zeit eines Rücklaufimpuls zeichnen die meisten LIDAR-Systeme auch die Intensität, d.h. die Stärke des Rücklaufimpuls auf. D.h. sie messen nicht nur, dass der Impuls zurückkehrt, sondern auch mit welcher Stärke (Intensität) er dies tut. Generell haben Objekte mit hohem Reflexionsvermögen, wie Schnee oder ein Metalldach, eine höhere Intensität des Reflexionswertes als dunkle Objekte, wie z.B. eine asphaltierte Straße.

LIDAR-Daten werden sowohl in Form von einfachen ASCII-Textdateien abgegeben als auch in Binärformaten. Bei den ASCII-Textdateien enthält jede Zeile der Datei die x-, y- und z-Koordinaten eines LIDAR-Punktes. Die hohe Punktzahl macht dieses Format für Datenlieferungen aber sehr unhandlich. Außerdem gibt es neben den x-, y- und z-Werten eine

Reihe von weiteren Informationen, die man für jeden Punkt speichern möchte, wie die Rücksendenummer, die Klassifizierung, den Intensitätswert, den Abtastwinkel, usw.
Ein heute sehr verbreiteter Standard für ein Binärformat ist das LAS-Format. Das LAS-Dateiformat ist ein veröffentlichtes Dateiformat für den Austausch von 3-dimensionalen Punktwolkendaten. Die Spezifikation ist von der
American Society for Photogrammetry and Remote Sensing (ASPRS) (http://asprs.org/a/society/committees/standards/LAS_1_4_r12.pdf) veröffentlicht und freigegeben.

3 Kombination von Luftbildern mit Sensordaten

In diesem Kapitel soll nun beispielhaft gezeigt werden, wie mit Hilfe von Spatial ETL-Technologie Luftbilder mit Sensordaten kombiniert werden können.
Ein Spatial ETL-Werkzeug, mit dem solche Daten kombiniert werden können, ist FME.

3.1 Spatial ETL-Technologie

Die FME (Feature Manipulation Engine) ist ein leistungsfähiges, universelles Spatial ETL-Werkzeug des kanadischen Unternehmens Safe Software Inc.. Spatial ETL Technologie dient der formatunabhängigen Geodatenprozessierung und Konvertierung räumlicher Informationen. Die beispielhaft verwendete FME-Technologie ermöglicht den Zugriff und die Verarbeitung von über 275 Datenquellen, d. h. CAD- und GIS-Formate, Datenbanken und Dienste.
Dabei unterstützt die Technologie eine Vielzahl von 3D-Formaten, darunter u.a. Point Cloud XYZ (ASCII-Format), ASPRS LIDAR Data Exchange Format (LAS) und 3D-pdf.

Die Software bietet – neben dem Lesen und Schreiben der zahlreichen Formate – verschiedene Möglichkeiten zur Geodatenprozessierung. Es stehen u.a. diverse Werkzeuge zur Bearbeitung von Punktwolken zur Verfügung, z. B. zum Zerlegen und Kombinieren der Farbwerte und zum Ausdünnen und Skalieren der Punktwolke. Es können die Eigenschaften der Punktwolke einfach ermittelt werden. Mit FME können TINs und Höhenlinien generiert, Sachdaten und Geometriedaten aus verschiedenen Quellen zusammengeführt, Rasterdaten mit Vektordaten und Punktwolken verschnitten werden, um nur einige wenige Beispiele zu nennen.

Zur Konfiguration der Geodatenprozessierung und Konvertierung gibt es eine komfortable grafische Benutzeroberfläche („FME Workbench“), die die Konfiguration anschaulich und einfach handhabbar macht. Programmierkenntnisse sind dazu nicht erforderlich. Für die Prozessierung steht eine Vielzahl von Werkzeugen (derzeit über 400), so genannte 'Transformer', zur Auswahl.

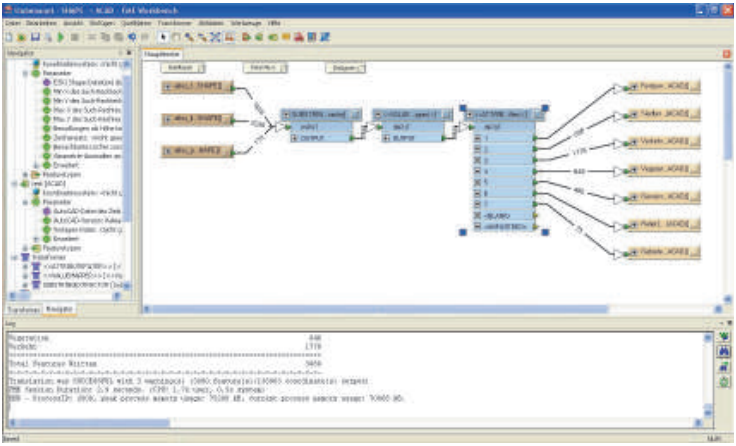


Abb. 2: FME Workbench

Die Viewer-Komponente FME Data Inspector ermöglicht die Visualisierung aller von FME unterstützten Formate inklusive der 3D-Formate. Es können sowohl die Geometriedaten in ihrer spezifischen Projektion als auch die zugehörigen Sachinformationen betrachtet werden. Neben der Viewerfunktion ist der Data Inspector auch ein ideales Werkzeug für die visuelle Qualitätskontrolle, u.a. des in der FME Workbench konfigurierten Prozesses.

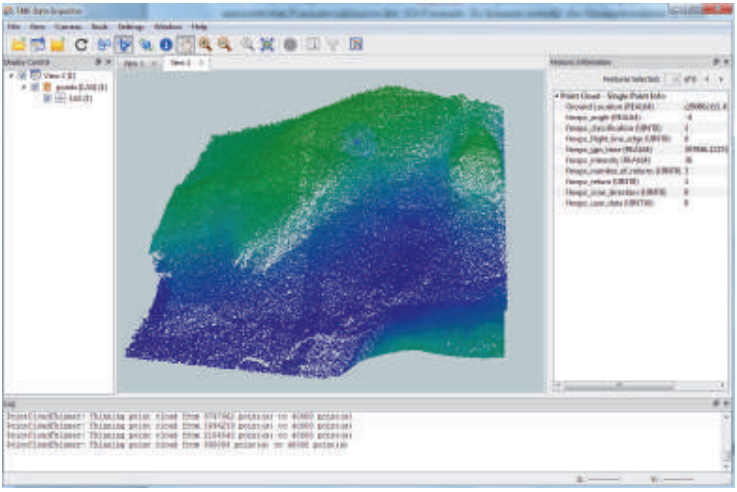


Abb. 3: FME Data Inspector

FME-Technologie kann auch im Intranet und Internet genutzt werden, Dazu wird der FME Server verwendet. Über den FME Server können z.B. Auswertungen zu LIDAR-Daten zentral im Intranet des Unternehmens oder der Organisation zur Verfügung gestellt werden.

3.2 Punktwolken auf der Basis von Luftbildern mit realen Farben versehen

Ein Beispiel für die Kombination von Punktwolken mit Luftbildern ist das Versehen von Punktwolken mit realen Farben. Liegt z.B. eine Punktwolke im LIDAR-Format vor (Abbildung links) und ein Orthophoto vom selben Gebiet (Abbildung rechts), so können beide verknüpft werden.

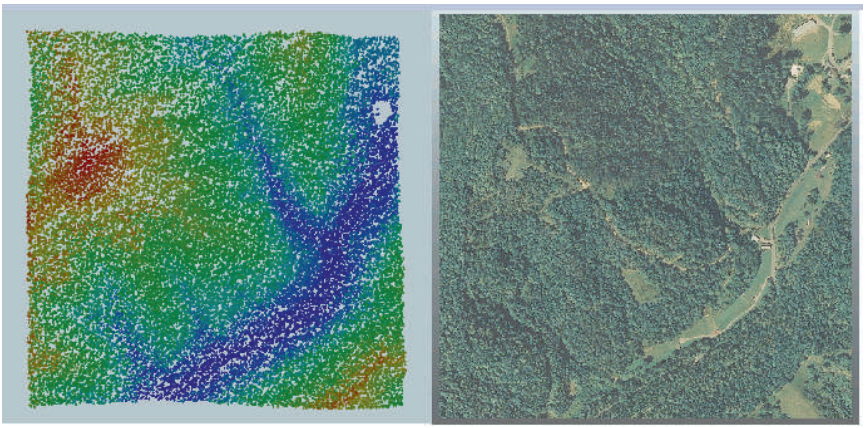


Abb. 4: Links Punktwolke im LIDAR-Format, rechts Orthophoto

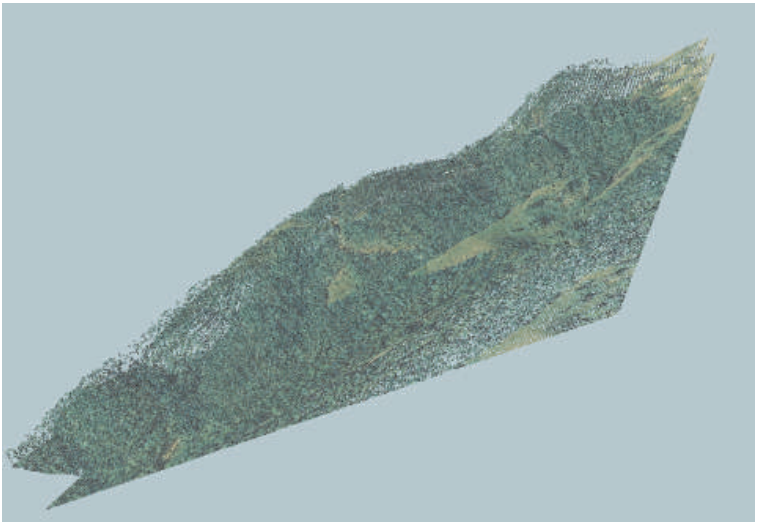


Abb. 5: Mit realen Farbwerten angereicherte Punktwolke

Dabei wird in dem Prozess aus der Punktwolke ein numerisches Raster erzeugt, das Orthophoto auf das Gebiet der Punktwolke geometrisch geklippt, die Raster zusammengerechnet und aus dem Ergebnisraster wieder ein Punktwolke erzeugt. Das Ergebnis kann dann wieder in ein LIDAR-Format geschrieben werden und auch mit dem FME Data Inspector direkt betrachtet werden.

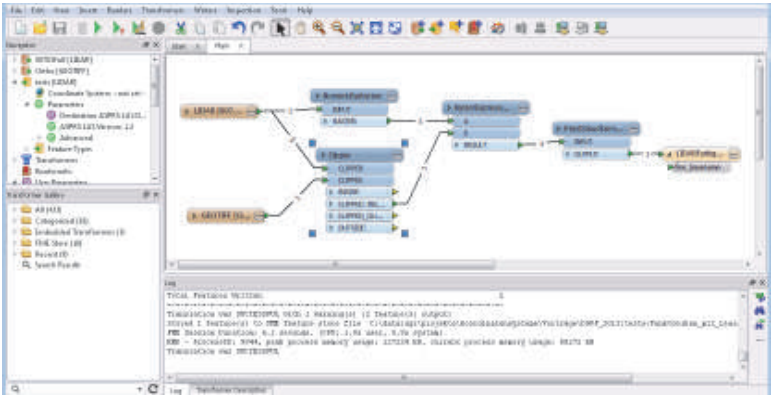


Abb. 6: FME-Prozess zur Anreicherung von Punktwolken mit realen Farben

3.3 Luftbilder mit Höhenwerten angereichert

Ein weiteres Beispiel für die Kombination von Sensordaten mit Luftbildern ist das Anreichern von Luftbildern mit Höhenwerten. Liegen die Höhenwerte z.B. in einem Digitalen Geländemodell (DGM) im XYZ-Format vor (Abbildung rechts) und ein Orthophoto vom selben Gebiet (Abbildung links), so können beide verknüpft werden und als ein 3-D-Luftbild ausgegeben werden. Die Beispieldaten wurden uns freundlicherweise vom Vermessungs-und Liegenschaftsamt der Stadt Düsseldorf für diesen Beitrag zur Verfügung gestellt.

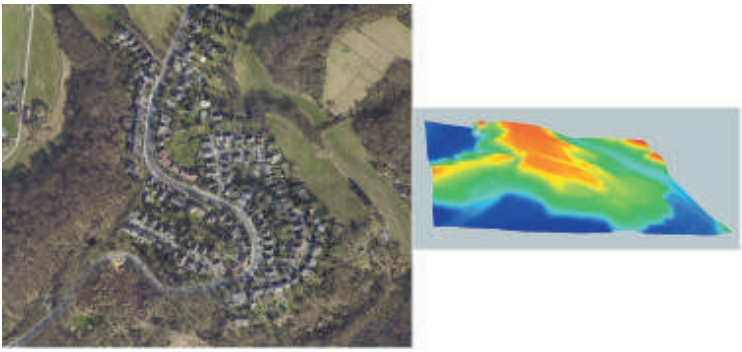


Abb. 7: Luftbild und Punktwolke im XYZ-ASCII-Format

Das Ergebnis kann z.B. nach 3D-pdf geschrieben und im Adobe Reader betrachtet werden.

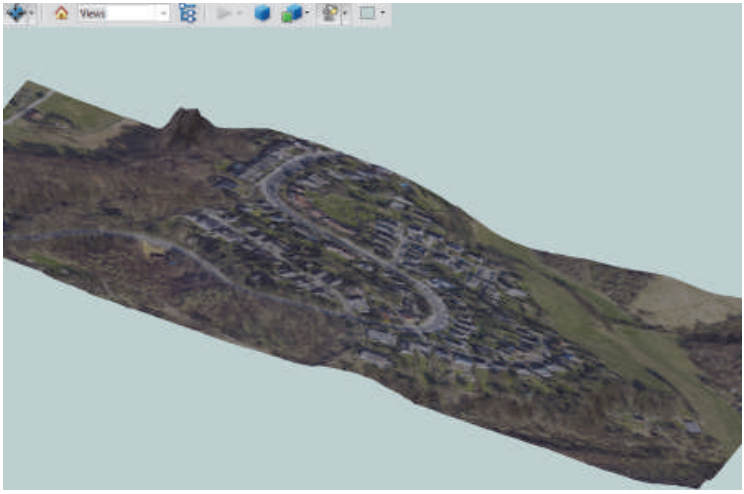


Abb. 8: 3D-Luftbild im 3Dpdf-Format

Dieses anschauliche Bild kann mit einer sehr einfachen Konfiguration der FME Workbench erzeugt werden. Dazu wird aus den Höheninformationen des Digitalen Geländemodells (XYZ-Format) eine TIN-Fläche generiert. Im nächsten Schritt wird die TIN-Oberfläche mit dem Luftbild überlagert und das Ergebnis der Überlagerung ins 3D-pdf-Format geschrieben.

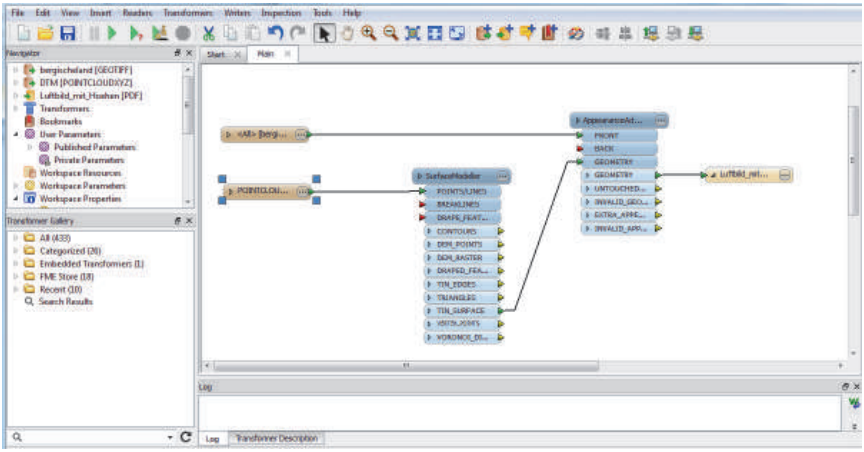


Abb. 9: FME-Prozess zur Erstellung von 3D-Luftbildern

4 Zielformate und Anwendungsszenarien

Neben der Möglichkeit, die Daten wieder als Punktwolken oder Rasterdateien zu speichern, z.B. im LAS-Format, ERDAD IMAGINE oder im XYZ-ASCII-Format, geotif, jpeg2000, etc. können auch andere 3D-Formate geschrieben werden, wie z.B. 3D-pdf, 3D-dxf, VRML und X3D. Auf diese Weise können die Daten einem breiten Anwenderkreis bereit gestellt werden. Vor allem 3d-pdf, das über den Adobe Reader auf nahezu jedem Rechner nutzbar ist, bietet hier gute Möglichkeiten, ein breites Publikum zu erreichen.

Aus Punktwolken generierte Geländemodelle kombiniert mit digitalen Orthophotos sind z.B. ideale Datenquellen für 3-D-Stadtmodelle, für Senderstandortplanung im Mobilfunk, für Hochwassersimulation, für Waldgebietskartierung und Virtual Reality-Szenen.
3D-pdfs können z.B. Planer nutzen, um sich einen Eindruck vom Gelände zu verschaffen. Sie können bei Präsentationen in Bürger-Informationsveranstaltungen gezeigt und online bereit gestellt werden.
Große Mehrwerte bestehen darin, dass die Daten in dieser Form nicht nur den Experten zugänglich bleiben, sondern auch von gewöhnlichen Anwendern genutzt werden können. Sie können so aufbereitet werden, dass sie auch in allgemein gebräuchlichen Formaten zur Verfügung stehen.

Grenzen sind dort erreicht, wo es nicht um den Eindruck, sondern um hohe Genauigkeiten geht. Es kann zu fehlerhaften Interpretationen und Klassifizierungen der Sensordaten kommen. Bei der Verknüpfung von 3d-Punktdaten mit Luftbildern z.B. muss es möglich sein, durch die Klassifizierung das Bodenprofil vom Bedeckungsprofil mit guter Qualität zu unterscheiden. Ansonsten werden dem Luftbild falsche Höhenwerte zugeordnet und es entsteht ein falsches Bild. Die Unterscheidung von Vegetation und Erdboden ist, je nach Vegetation, nicht immer einfach aus der Klassifizierung zu ermitteln. Die Unterscheidung der Vegetation zu festen Objekten und vor allem die Differenzierung der Vegetation ist anhand der Klassifizierung nicht immer einfach.

5 Fazit

Standard Spatial ETL-Werkzeuge bieten heutzutage hervorragende Möglichkeiten, traditionelle Luftbilder und moderne Sensordaten zusammenzuführen und gemeinsam zu nutzen. Es besteht die Möglichkeit, die kombinierten Daten in Formate auszugeben, die von jedem nutzbar sind, wie z.B. 3D-pdf. So können komplexe Daten und Sachverhalte einfach und anschaulich in 3D dargestellt und betrachtet werden. Diese können einem breiten Publikum zur Verfügung gestellt werden.

6 Literaturverzeichnis

ASPRS, 2012: LASer (LAS) File Format Exchange Activities, http://asprs.org/Committee-General/LASer-LAS-File-Format-Exchange-Activities.html

CON TERRA GMBH, 2012: FME Desktop, http://www.conterra.de/de/software/fme/desktop/desktop.shtm

PRINZ, T. & HANDT, S.,2007: Digitale Fernerkundungsmethodik in den Geowissenschaften, Vorlesung Geofernerkundung, http://ivvgeo.uni-muenster.de/Vorlesung/FE_Script/2_2.html

SAFE SOFTWARE, 2012: Complete list of Formats Supported by FME, http://www.safe.com/fme/format-search/

SAFE SOFTWARE, 2012: Colouring Point Clouds Using Orthophotos, http://fmepedia.safe.com/articles/Samples_and_Demos/Colouring-Point-Clouds-Using-Orthophotos

SCHUCKMAN, K. & RENSLOW, M., 2012: Lidar Sensor Design, https://www.e-education.psu.edu/lidar/l1_p3.html

UNIVERSITÄT ROSTOCK, 2008: Geoinformatik-Lexikon, Airborne Laserscanning (ALS), http://www.geoinformatik.uni-rostock.de/einzel.asp?ID=-1616705597, dort benannte Quelle: Bill, R. Grundlagen der Geo-Informationssysteme Band 1. Hardware, Software und Daten

U.S. GEOLOGICAL SURVEY , USDA-FSA-APFO AERIAL PHOTOGRAPHY FIELD OFFICE, 2011: National Agriculture Imagery Program (NAIP) Orthoimagery for Zone 11 California State Quarter Quadrangle HALF DOME

Zuordnung von 3D Gebäudemodellen und IR-Videosequenzen mit Linienverfolgung

DOROTA IWASZCZUK [1], LUDWIG HOEGNER [1] & UWE STILLA [1]

Zusammenfassung: Die Position und Orientierung eines Luftfahrzeuges im Raum wird durch GPS, oft mit einer IMU, bestimmt. Oft ist die mit GPS/IMU aufgenommene Position und Orientierung nicht genau genug um die in den aufgenommen Daten detektierten Objekte besser zu georeferenzieren und um die Daten aus verschiedenen Quellen zu fusionieren. In diesem Beitrag wird eine Methode zur Verbesserung der Georeferenzierung eines mit einer Infrarotkamera aufgenommenen Videostromes mittels Verfolgung markanter Geradenstücke durch die Bildsequenz vorgestellt. Es wird zwischen normalen Frames und Key-Frames unterschieden. In einem Key-Frame werden für jede Modelllinie innerhalb eines Puffers passende Bildlinien gesucht und als Kandidaten für Korrespondenzen weiter bearbeitet. Die richtigen Korrespondenzen werden mit einem RANSAC-Verfahren selektiert. Dann werden die Position und Orientierung der Kamera neu berechnet. Es werden dabei Unsicherheiten sowohl der äußeren Orientierung wie auch des Gebäudemodells berücksichtigt. Einmal im Bild zugeordnete Modellkanten werden mittels Korrelation im Bildraum durch die Bildsequenz verfolgt. Die Korrektheit der Linienverfolgung wird überprüft, indem in einem regelmäßigen Abstand „Key-Frames" eingeführt werden.

1 Einleitung

Die Wärmetexturen, die der Detektion und Dokumentation von Wärmeleckagen in der Gebäudehülle dienen, werden immer häufiger aus Infrarotdaten (IR) abgeleitet und den Gebäudemodellen zugeordnet. Die Außenfassaden können von einem Fahrzeug aufgenommen werden (HOEGNER et al, 2007). Von dem Fahrzeug sind jedoch die Dächer nicht sichtbar und die Innenhoffassaden sind oft für die Messfahrzeuge nicht zugänglich. Um diese Flächen abbilden zu können, werde fliegende Plattformen, wie z.B. Hubschrauber (IWASZCZUK, et al., 2012), Flugzeuge FRUEH et al. (2004) oder UAVs, eingesetzt. Um die Bilddaten mit den 3D-Gebäudemodellen zu kombinieren, werden die Parameter der äußeren und inneren Orientierung benötigt. Die Parameter der äußeren Orientierung des Luftfahrzeuges werden durch ein GPS, oft mit einer IMU (Eng. inertial measurement unit) kombiniert, aufgezeichnet. Die innere Orientierung des Kamerasystems wird durch den Selbstkalibrierungsprozess oder im Labor bestimmt. Oft ist die mit GPS und IMU aufgenommene Position und Orientierung nicht genau genug um die in den Daten detektierten Objekte mit hoher Genauigkeit zu georeferenzieren und um die Daten aus verschiedenen Quellen zu fusionieren. Um die äußere Orientierung der Kamera zu verbessern wird oft Model-to-Image Matching durchgeführt.

1) Photogrammetrie & Fernerkundung, Technische Universität München (TUM), Arcisstraße 21, 80333 München, http://www.pf.bv.tum.de

FRUEH et al. (2004) projizierte das 3D-Gebäudemodell mit verschiedenen, zufälligen Parametern ins Bild und korrelierte die Modelllinien mit den extrahierten Liniensegmenten. Aus mehreren Versuchen wurde dazu die Kameraposition mit der besten Anpassung ausgewählt. In jedem Bild der Sequenz verlief der Vorgang gleich. Diese Methode nutzt keine Zusammenhänge zwischen den Frames. Hu et al. (2006) verbinden die Frames mit markanten Punkten, die jedoch keinen Zusammenhang mit den Gebäudestrukturen aufweisen. Unter Annahme sog. „Manhattan“ oder „Legoland scenes“ mit dominanten horizontalen und vertikalen Richtungen kann die Ausrichtung der Kamera mithilfe der Fluchtpunkte berechnet werden (DING & ZAKHOR, 2008; CHAM et al., 2010; FOERSTNER, 2010). Dieses Verfahren hilft bei schlechten Navigationsparametern, die Fehler der Drehwinkel zu reduzieren. Um die Texturen präzise zuordnen zu können, wird eine Methode benötigt, die für das Matching bei Gebäuden charakteristische Bildmerkmale verwendet und diese von Frame zu Frame verfolgt.

2 Methode

Die in diesem Beitrag vorgestellte Methode dient der Verbesserung der Georeferenzierung eines mit der Infrarotkamera aufgenommenen Videostromes mittels Verfolgung markanter Liniensegmente durch die Bildsequenz. Es wird dabei angenommen, dass sowohl die Lage der extrahierten Bildkanten wie auch die Lage der 3D-Gebäudemodelle unsicher sind. Die Fortpflanzung der Unsicherheit wird schematisch in Abb. 1 dargestellt. Es wird ebenso angenommen, dass die Größe der Unsicherheit der 3D-Modellpunkte bekannt ist, und dadurch die Unsicherheit der 3D-Kante berechnet werden kann. Diese Unsicherheit wird während der Projektion in 2D fortgepflanzt. Im 2D-Bild werden die Kanten extrahiert und die Unsicherheiten der Kanten bestimmt. Dann werden die projizierten Kanten des 3D-Gebäudemodells mit den Bildkanten gematcht. Dabei werden die fortgepflanzten Unsicherheiten des Gebäudemodells und die Unsicherheiten der Bildkanten berücksichtigt.
Für die Projektion werden die mit dem GPS bzw. der IMU ermittelten Parameter der äußeren Orientierung sowie die durch Kalibrierung bekannten Parameter der inneren Orientierung verwendet. Da die mit GPS/IMU aufgezeichneten Parameter der äußeren Orientierung nur als Näherungswerte betrachtet werden können, sollen diese im Matching-Prozess so verbessert werden, dass das projizierte Modell und die Bildstrukturen möglichst gut aufeinander angepasst werden. Dazu werden zwei Typen von Frames definiert: Einerseits die sog. Key-Frames, in denen das vollständige Matching durchgeführt wird (Kap. 2.1), anderseits die normalen Frames in denen das vereinfachte Matching mittels Linienverfolgung stattfindet (Kap. 2.2). Die verfolgten Linien und deren Zuordnung zu den Gebäudekannten wird in einer Chronik gespeichert. Die hier präsentierte Methode setzt voraus, dass die Bewegungen zwischen den Frames gering sind.

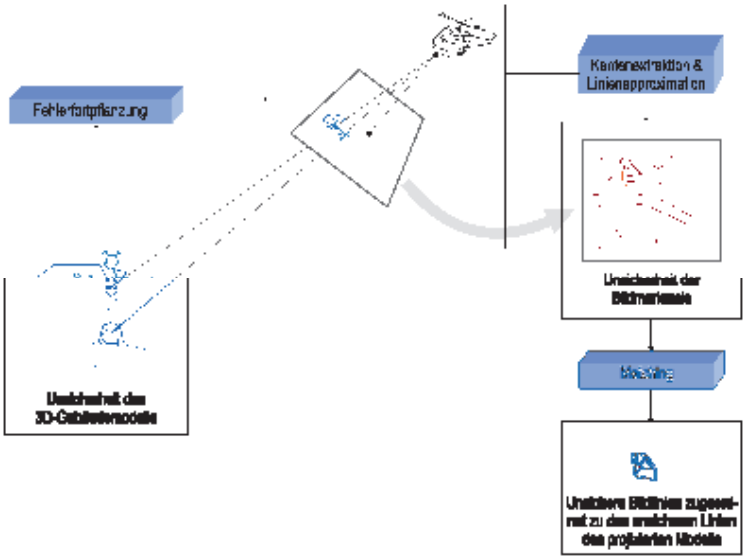


Abb. 1: Schematische Darstellung der Fehlerfortpflanzung und des Matchings

2.1 Key-Frame - Einzelbildfall

Um den Ablauf zu starten wird der erste Frame immer als ein Key-Frame definiert. Weitere Key-Frames werden in regelmäßigen Abständen eingesetzt. Ein Key-Frame wird als Einzelbild betrachtet. Zunächst werden die Kanten des 3D-Gebäudemodells ins Bild projiziert und für jede Kante wird in der lokalen Umgebung nach passenden Liniensegmenten im Bild gesucht. Die Größe des Suchbereiches wird durch die Fehlerfortpflanzung der Unsicherheit der Kameraposition und –Orientierung und der Unsicherheit der Modellkante berechnet. Die richtigen Korrespondenzen werden mit einem RANSAC-Verfahren selektiert. Dann werden die Position und Orientierung der Kamera neu berechnet. Um die Unsicherheiten sowohl der äußeren Orientierung wie auch des Gebäudemodells als Konsequenz der ungenauen Erfassung und der Generalisierung zu berücksichtigen, wird ein stochastisches Modell in der Ausgleichung formuliert (IWASZCZUK, et al., 2012). Die neu berechnete Position und Orientierung des Sensors werden auf die weiteren Frames übertragen, da die Verschiebung und Rotation des Sensors zwischen den Frames relativ genau mit der IMU bestimmt werden kann. Daher werden nach jedem Key Frame als Ergebnis die Korrekturen der Position und Orientierung für die nächste Aufnahme weitergegeben. Außerdem werden die Korrespondenzen gespeichert.

2.2 Linienverfolgung

Jede Linie, die in einem Key-Frame einer Modellinie zugeordnet wird, wird im nächsten Frame gesucht. Da in einem Videostrom mehrere Frames pro Sekunde aufgenommen werden, sind die Unterschiede in der Erscheinung der Objekte in benachbarten Frames sehr gering. Deswegen eignet sich die Kreuzkorrelation für die Linienverfolgung, da die geometrischen Verzerrungen von Frame zu Frame sehr gering sind und die Ergebnisse der Kreuzkorrelation kaum beeinflussen. Die neue Position der Linie im nächsten Frame kann aus dem Pfad des Luftfahrzeuges geschätzt werden. Dafür wird um jede zugeordnete Linie ein Bildbereich gewählt und mit dem Bildbereich, wo dieselbe Linie erwartet ist, korreliert. Dadurch ergibt sich für jede Linie ein $\mathbf{offset}_i=[\Delta row_i,\Delta col_i]$, aus dem eine Verschiebung zwischen den Frames als $median(\mathbf{offset}_i)$ berechnet wird. Dann wird in der Umgebung der verfolgten Linie überprüft, ob sich dort eine passende Linie befindet, die in dem neuen Frame als dieselbe Linie erkannt werden kann. Dafür werden statistische Tests nach HEUEL (2002) verwendet. HEUEL (2002) definiert die algebraische Bedingung für die Identität zweier Linien als

$$S(\mathbf{l})\mathbf{m} = -S(\mathbf{m})\mathbf{l} = \mathbf{0}, \qquad (1)$$

wo

$$S(\mathbf{a}) = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}, \qquad (2)$$

$\mathbf{m}=[m_1,m_2,m_3]^T$, $\mathbf{l}=[l_1,l_2,l_3]^T$ sind Geraden in projektiven Koordinaten.
So kann eine Hypothese H_0: $\mathbf{d}=S(\mathbf{l})\mathbf{m} = -S(\mathbf{m})\mathbf{l} = \mathbf{0}$ aufgestellt werden. Die Kovarianzmatrix für den Distanzvektor $\mathbf{d}$ wird als

$$\Sigma_{\mathbf{dd}} = S(\mathbf{l})\Sigma_{\mathbf{mm}}S(\mathbf{l})^T + (-S(\mathbf{m}))\Sigma_{\mathbf{ll}}(-S(\mathbf{m})^T) \qquad (3)$$

berechnet. H_0 kann mit dem Signifikanzniveau α abgelehnt werden, wenn

$$T = \underline{\mathbf{d}}^T\Sigma_{\mathbf{dd}}^{-1}\underline{\mathbf{d}} > \varepsilon_H = \chi^2_{1-\alpha;r}, \qquad (4)$$

wo $r = rank(\Sigma_{\mathbf{dd}})$.

Die Linienverfolgung wird nur für die Linien durchgeführt, die im nächsten Frame sichtbar sind. Es wird kontinuierlich geprüft, ob durch die Bewegung des Sensors zusätzliche Kanten des Gebäudemodells im Bild sichtbar werden oder verschwinden. Diese Information wird aus der bekannten Geometrie des 3D Gebäudemodells mithilfe des Hidden-Line-Algorithmus berechnet. Dafür werden Ereignisse und Zustände für jede Modelllinie definiert. Prinzipiell kann eine Linie folgende Zustände aufweisen: *lebend vollkommen sichtbar, lebend verletzt (teilsichtbar), verdeckt, tot*. *Lebend vollkommen sichtbar* ist eine Linie, die nach der Projektion vollkommen im Bild erscheint und von keinen Objekten verdeckt wird. *Lebend verletzt* ist dagegen eine Linie,

die nur Teilweise im Bild erscheint oder zum Teil von einem Objekt verdeckt wird. Im *verdeckten* Zustand befindet sich eine Linie die zwar ins Bild projiziert wird, aber von einem anderen Objekt vollkommen verdeckt wird. Fällt eine Linie aus dem Bildbereich, so wird sie als *tot* gekennzeichnet. Der Zustand der Modelllinie definiert, wie die korrespondierenden Bildlinien verfolgt werden. Die *lebenden* Linien werden mit den extrahierten Bildlinien abgeglichen, während die Verfolgung der *verdeckten* Linien unterbrochen wird, bis diese wieder erscheinen. Die toten Linien werde nicht mehr verfolgt. Die Tab. 1 definiert sechs Ereignisse (Events), die den Zustand der Linie verändern. Es werden mögliche Zustände vor und nach jedem Ereignis aufgeführt.

Ereignis	Mögliche Zustände vor dem Ereignis	Mögliche Zustände nach dem Ereignis
Geburt	∅	Lebend vollkommen sichtbar Lebend verletzt
Erscheinung	Verdeckt	Lebend vollkommen sichtbar Lebend verletzt
Verletzung	Lebend vollkommen sichtbar	Lebend verletzt
Heilung	Lebend verletzt	Lebend vollkommen sichtbar
Verschwinden	Lebend vollkommen sichtbar Lebend verletzt	Verdeckt
Tod	Lebend vollkommen sichtbar Lebend verletzt Verdeckt	Tot

Tab. 1: Ereignisse und Zustände, die für eine Linie auftreten können

3 Testdaten

Die IR-Daten wurden während einer Flugkampagne über einem dicht bebauten Bereich mit einer im Hubschrauber montierten IR-Kamera aufgenommen. Die Frequenz des Videostroms lag bei 25 Bildern pro Sekunde. Die Flughöhe betrug ca. 400 m über Grund und die Bilder wurden schräg nach vorne unter einem Pitch-Winkel von ca. 45º aufgenommen.
Der Hubschrauber verfügte über ein GPS/INS Navigationssystem. Das Projektionszentrum der Kamera war im Verhältnis zur Antenne des Navigationsgerätes versetzt. Die innere Orientierung der Kamera war nur ungefähr bekannt. Daher wurde eine erweiterte Systemkalibrierung mit der Berechnung von boresight & leverarm Parametern (KOLECKI et al., 2010) durchgeführt.
Die 3D-Gebäudemodelle wurden im Detailierungsgrad LOD2 aus Luftbildern semiautomatisch erstellt und im *CityGML*-Format gespeichert.

4 Ergebnisse

Die ersten Tests wurden mit den oben beschriebenen Daten durchgeführt. Es wurden fünf nacheinander folgende Frames gewählt. Der erste Frame wurde als Key-Frame definiert. Es

wurden automatisch mit RANSAC zum Modell korrespondierende Kanten gewählt und die Parameter der äußere Orientierung neu berechnet. Die Abb. 2 zeigt das mit den neu berechneten Parametern projizierte Modell (links) und die mit RANSAC gewählten Liniensegmente im Bild. Die dem Modell zugeordneten Linien wurden in den folgenden Frames verfolgt. Die Ergebnisse der Verfolgung wurden in Abb.3 dargestellt. In den Frames 2-4 wurde das in den Frame 1 projizierte Modell um den entsprechenden mittleren Offset verschoben. Die Genauigkeit der in dem ersten Frame neu bestimmten Parameter der äußere Orientierung lag bei 3-4 [m] und 0.3-0.5[˚].

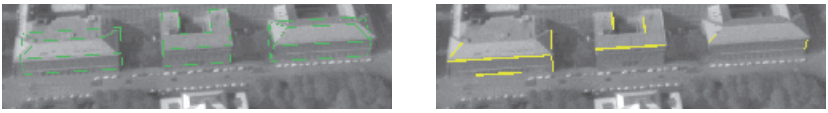


Abb. 2: Ausschnitt aus dem Key-Frame: links - das mit neu berechneten Parametern projizierte 3D-Gebäudemodell, rechts - die mit RANSAC gewählten Liniensegmente, die in den nächsten Frames verfolgt werden

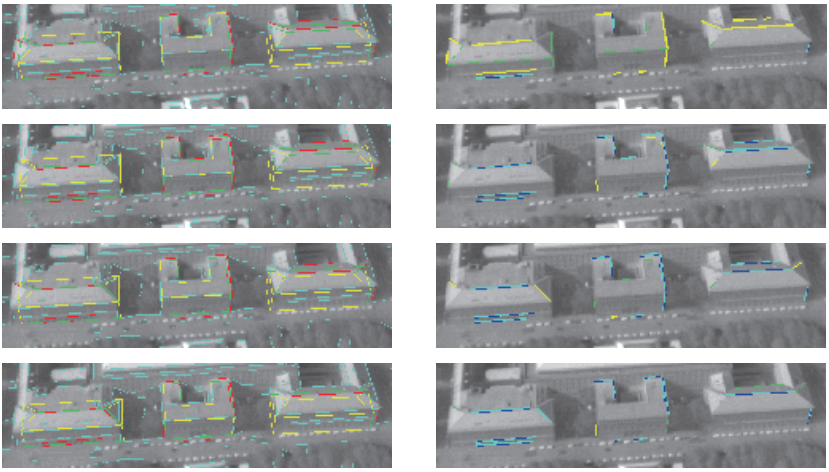


Abb. 3: Ausschnitte aus vier nacheinander folgenden Frames. Links: extrahierte Bildkanten, die dem verfolgten Modell zugeordnet worden sind (Rot), nicht zugeordnete Bildkanten (Cyan), Modellkanten mit zugeordneten Bildkanten (Grün), Modellkanten ohne Zuordnung (Gelb). Rechts: Aus dem vorherigen Frame verfolgte Kanten, die wiedergefunden wurden (dunkel Blau) und Linien die keinen Kanten im aktuellen Bild (Cyan) zugeordnet worden sind (Grün); in Gelb werden die neugefundenen Bildkannten dargestellt, die im vorherigen Bild nicht vorkamen.

5 Diskussion

Die Verfolgung der Geradenstücke schränkt den Suchraum für Linienkorrespondenzen ein und beschleunigt dadurch den Prozess. Die Genauigkeit der neuberechneten Parameter der äußeren Orientierung ist jedoch zu schlecht um ab dem zweiten Frame die Linien im Bild ohne RANSAC verfolgen zu können. Deswegen sollen in Zukunft Experimente mit mehreren Key-Frames am Anfang der Sequenz durchgeführt werden. Die Key-Frames sollen auch nicht am Anfang festgelegt werden, sondern es soll dynamisch entschieden werden, ob der nächste Frame als Key-Frame festgelegt wird. Dafür werden Qualitätsmaße und Kriterien benötigt, die diese Entscheidung automatisch ermöglichen. Es soll untersucht werden, welche Größen (z.B. Genauigkeit des Matchings, Anzahl der gefundenen Korrespondenzen) dafür geeignet sind. Es sollen auch Zuverlässigkeitsmaße für die verfolgten Linien definiert werden. Die Linien, die häufiger gefunden werden, werden als zuverlässiger eingestuft und länger verfolgt, während die nur gelegentlich gefundenen Linien als weniger zuverlässig gesehen werden und die Verfolgung schneller verworfen wird.
Es soll außerdem die Korrespondenzsuche verbessert werden. Insbesondere bei ungenauen GPS/INS Daten können mit RANSAC nicht immer die richtigen Korrespondenzen gefunden werden (IWASZCZUK, et al., 2012). RANSAC lässt auch mehrere Bildlinien die einer Modelllinie zugeordnet sind zu. Deswegen soll in der Zukunft eine Methode eingesetzt werden, die nur 1-zu-1 Korrespondenzen zulässt. Alternativ kann ein Akkumulator-Verfahren angewandt werden, das noch robuster als RANSAC die richtigen Korrespondenzen bestimmt.
Die Analyse der Ereignisse und Zustände hilft die Zuverlässigkeit der Verfolgung zu erhöhen und kann in der Zukunft auch bei der Flugplanung genutzt werden, um die optimale Flugroute des Luftfahrzeuges zu bestimmen.

6 Danksagung

Die Autoren danken Fraunhofer IOSB für die Bereitstellung der IR-Daten.

7 Literaturverzeichnis

CHAM T.J.; CIPTADI A.; TAN W.C.; PHAM M.T. & CHIA L.T., 2010: Estimating Camera Pose from a Single Urban Ground-View Omnidirectional Image and a 2D Building Outline Map. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, USA

DING, M. & ZAKHOR, A., 2008: Automatic registration of aerial imagery with untextured 3D LiDAR models. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)

FOERSTNER, W., 2010: Optimal vanishing point detection and rotation estimation of single image from a Legoland scene. The International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences, 38(3B): 157-162

FRUEH C.; SAMMON R. & ZAKHOR A., 2004: Automated Texture Mapping of 3D City Models With Oblique Aerial Imagery. Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'04)

HEUEL, S., 2002: Statistical Reasoning in Uncertain Projective Geometry for Polyhedral Object Reconstruction. Photogrammetry Department, Institute for Geodesy and Geoinformation, Bonn University, Germany, PhD Thesis

HOEGNER, L.; KUMKE, H.; MENG, L. & STILLA, U., 2007: Automatic extraction of textures from infrared image sequences and database integration for 3D building models. Photogrammetrie Fernerkundung Geoinformation (PFG), 2007(6): 459-468

HU, J.; YOU, S. & NEUMANN, U., 2006: Automatic Pose Recovery for High-Quality Textures Generation. Proceedings of 18th International Conference in Pattern Recognition: 561-565

IWASZCZUK, D.; HOEGNER, L.; SCHMITT, M. & STILLA, U., 2012: Line based Matching of Uncertain 3D Building Models with IR Image Sequences for Precise Texture Extraction. Photogrammetrie, Fernerkundung, Geoinformation (PFG), 2012(5): 511-521

KOLECKI, J., IWASZCZUK, D. & STILLA, U., 2010: Calibration of an IR camera system for automatic texturing of 3D building models by direct geo-referenced images. Proceedings of Eurocow 2010, Castelldefels, Spain

Geometrie- und farbbasierte Punktwolkenklassifizierung von komplexen urbanen Laserscanningszenen mittels überwachter SOM-Klassifikation

ERIC MATTI[1] & STEPHAN NEBIKER[1]

Zusammenfassung: In diesem Bericht wird eine robuste Methode zur direkten Klassifizierung texturierter 3D-Punkte komplexer urbaner Laserscanningszenen vorgestellt. Die Methode basiert auf den geometrischen und farblichen Charakteristika der Punkte, indem für jeden Punkt sein jeweiliges Fast Point Feature Histogramm, sein Zenitwinkel zur Flächennormalen sowie sein HSV-Farbwert berechnet wird. Daraus resultiert ein 37-dimensionaler Featurevektor pro Punkt als Input für die überwachte Punktklassifikation mittels einer trainierten Self-Organizing Map (SOM). Erste Untersuchungen des vorgeschlagenen Verfahrens an einer realen komplexen urbanen Laserscanningszene zeigen auf, dass die Methode in der Lage ist, Punkte nach den Objektklassen Strasse, Baumstamm/-ast, Gebäudefassaden und Vegetation zu klassifizieren.

1 Einleitung

Der urbane Raum lässt sich heute mittels dichter, texturierter 3D-Punktwolken aus terrestrischen oder mobilen Laserscans sehr effizient abbilden. Diese 3D-Puntwolken bilden in aller Regel die Basis, um explizite urbane 3D-Modelle, bestehend aus Punkten, Linien, Flächen und Volumen, in einem interaktiven oder bestenfalls semi-automatischen Prozess abzuleiten. Obwohl die dichten 3D-Punktwolken einen hohen geometrischen und radiometrischen Informationsgehalt aufweisen, ist die automatische Ableitung verschiedenster semantischer oder geometrischer Informationen und Modelle aus Punktwolken keine einfache Fragestellung. Nebst der Komplexität der realen Welt führen auch die Eigenschaften von 3D-Punktwolken, wie z.B. Ausreisser in den Messungen, unterschiedliche Punktdichten oder fehlende Aufnahmegebiete, dazu, dass verschiedene Methoden verwendet werden, um die gewünschten Informationen aus Punktwolkenszenen abzuleiten.

In dieser Arbeit wird eine neue robuste direkte Methode zur Klassifizierung von 3D-Punkten vorgestellt. Mithilfe dieser Methode können texturierte Punkte von komplexen urbanen Laserscanningszenen nach einer finiten Anzahl von gewünschten disjunkten Objektklassen klassifiziert werden. Diese Methode basiert auf den geometrischen und farblichen Charakteristika der Punkte, indem für jeden Punkt sein jeweiliges Fast Point Feature Histogramm (FPFH), sein Zenitwinkel zur Flächennormalen sowie sein HSV-Farbwert berechnet wird. Anhand des Point Feature Histogramm Wertes lässt sich bestimmen, ob sich der Punkt auf einer Ebene, Kugel, Zylinder, Kante oder Ecke befindet. Unter Verwendung des Zenitwinkels lässt sich in einem weiteren Schritt die absolute Neigung der unterliegenden Geometrie definieren. Nebst diesen Berechnungen von geometrischen Punktcharakteristika werden zudem die RGB-Farbwerte pro Punkt in den HSV-Farbraum transformiert, um dekorrelierte und somit robustere Farbwerte zu erhalten. Daraus resultiert ein 37-dimensionaler Featurevektor pro Punkt als Input für die überwachte Punktklassifikation mittels einer trainierten Self-Organizing Map (SOM).

1 Eric Matti, Stephan Nebiker, Fachhochschule Nordwestschweiz, Institut Vermessung und Geoinformation, 4132 Muttenz, Schweiz; E-Mail: eric.matti@fhnw.ch, stephan.nebiker@fhnw.ch

2 Problemstellung

Ziel der Objekterkennung in einer Punktwolke ist, einzelne abgegrenzte Teilmengen von Punkten einer Objektklasse der realen Welt zuzuordnen. Dabei können die Objektklassen in abstrahierter Form mittels einer Menge geometrisch oder nicht-geometrisch beschreibender Charakteristika, so genannten Features oder Deskriptoren, definiert werden. Werden nun für jeden Punkt seine Charakteristika berechnet, so kann mittels einer Klassifikationsmethode, die jeden Punkt abhängig von seinen Charakteristika einer Objektklasse zuweist, die abgegrenzte Menge einer Objektklasse gebildet werden. Die dabei verwendeten Punktcharakteristika sollten nach BIOSCA & LERMA (2008) das umliegende Gebiet des Punktes möglichst gut repräsentieren. Nebst dieser Bedingung hängt die Wahl der Charakteristika auch stark davon ab, welche Objekte voneinander unterschieden werden sollen, denn um eine möglichst geringe Fehlklassifikation erster und zweiter Art zu erreichen, sollten die Menge der Charakteristika unterschiedlicher Objektklassen möglichst disjunkt sein.

3 Verwandte Arbeiten

Obwohl es eine Vielzahl unterschiedlicher Punktklassifikationsmethoden gibt, können diese in die zwei Gruppen der indirekten und direkten Punktklassifikationsmethoden aufgeteilt werden. Die direkten Methoden können wiederum in geometrie- und punktcharakteristika-basierte Punktklassifikationsmethoden unterteilt werden.

Das Prinzip der geometrie-basierten direkten Punktklassifikationsmethoden besteht darin, sämtliche Punkte, die einer gesuchten Geometrie am besten entsprechen, zu finden. Dazu können nach VOSSELMAN (2009) zum Beispiel eine 3D-Houghtransformation oder eine RANSAC-Ebenendetektion verwendet werden. Eine weitere Methode, um Punkte nach geometrischen Kriterien direkt einer geometrischen Objektklasse zuzuweisen, ist die Verwendung von Grammars. Beispiele von grammar-basierten, Modellierungsansätzen finden sich in den Arbeiten von WAN & SHARF (2012) und von HOHMANN et al. (2009).

Punktcharakteristika-basierte direkte Punktklassifikationsmethoden basieren darauf, für jeden Punkt der Punktwolke seine Charakteristika zu berechnen. Mittels einer Klassifikationsmethode werden dann in einem zweiten Schritt die einzelnen Punkte anhand ihres Eigenschaftsvektors einer Klasse zugewiesen. Die Ansätze von LIM & SUTER (2007) sowie RUSU et al. (2009) verwenden dabei Conditional Random Fields (CRF), um Gebäude, Gelände, Vegetation und Baumstamm/-ast bzw. planare, konvexe und konkave Flächen sowie Kanten zu erkennen.

4 Lösungskonzept zur geometrie- und farbbasierten Punktwolkenklassifizierung

Unser Lösungsansatz zur geometrie- und farbbasierten Punktwolkenklassifizierung setzt sich aus den folgenden Schritten zusammen:

1. **Erstellen der Trainingsdaten:** In einem Punktwolkeneditor werden mehrere Punkte manuell den zu detektierenden Objektklassen zugewiesen.
2. **Punktcharakteristikaberechnung:** Berechnung der HSV-Farbwerte, der Flächennormalen n_q, des Zenitwinkels z_q zur Flächennormalen und der 33 Features

des Fast Point Feature Histogramms $fpfh_{33_q}$ für jeden Punkt p_q der Trainingsdaten und der zu klassifizierenden Punktwolke.

3. **Überwachte SOM-Klassifikation:** Training des SOM-Klassifikators und Klassifikation der zu klassifizierenden Punktwolke unter Verwendung des Featurevektors $[h, s, v, z, fpfh_{33}]$.

Die Schritte zwei und drei werden in den nachfolgenden Kapiteln kurz vorgestellt.

4.1 Transformation der RGB-Farbwerte nach HSV-Farbwerten

Aufgrund der hohen Korrelation zwischen den einzelnen Farbwerten im RGB-Farbmodell, werden die RGB-Farbwerte ins robustere HSV-Farbmodell transformiert. Dabei wird eine Farbe im HSV-Farbmodell primär durch ihren Farbton (H) charakterisiert, während unterschiedliche Farbsättigung- (S) und Helligkeitswertwerte (V) Variationen des gleichen Farbtons entsprechen. Als Resultat erhält man somit eine robustere, auf dem Farbton basierende farbliche Charakterisierung des 3D-Punktes.

4.2 Berechnung der Flächennormalen

Als Basis zur Berechnung der Features des Zenitwinkels zur Flächennormalen und des Fast Point Feature Histogramms, wird die Flächennormale jedes Punktes berechnet. Die Flächennormale n_q eines Punktes p_q wird nach der Methode von BERKMANN & CAELLI (1994) berechnet, indem die Kovarianzmatrix C_q nach der Formel (1) unter Verwendung von k Nachbarpunkten berechnet wird und dessen Eigenvektor n_0 mit dem geringsten Eigenwert λ_0 bestimmt wird.

$$C_q = \frac{1}{k}\sum_{j=1}^{k}(p_j - \bar{p})(p_j - \bar{p})^T \qquad (1)$$

Die berechneten Flächennormalen sind jedoch hinsichtlich ihrer Orientierung nach innen oder aussen gegenüber der angenommenen Fläche nicht eindeutig und werden somit, unter der Bedingung des Vorliegens einer Scanaufnahme, in Richtung des Instrumentenstandorts orientiert. Dabei werden sämtliche Flächennormalen n_q in Richtung des Instrumentenstandorts v orientiert, indem dass Vorzeichen so gewählt wird, dass die Gleichung

$$n_q \cdot (v - p_q) > 0 \qquad (2)$$

erfüllt wird.

4.3 Berechnung des Zenitwinkels zur Flächennormalen

Um ein absolutes geometrisches Punktcharakteristikum zu erhalten, das Rückschlüsse auf die absolute Neigung der umliegenden Punktoberflächen erlaubt, wird ausgehend von den berechneten und orientierten Flächennormalen der Zenitwinkel z_q zur Flächennormale für jeden Punkt berechnet (MATTI, 2012).

4.4 Berechnung des Fast Point Feature Histogramms (FPFH)

Als relatives Punktcharakteristikum wird das Fast Point Feature Histogramm (FPFH) von RUSU (2009) verwendet. Ziel dieses lokalen geometrischen Punktdeskriptors ist, die geometrische Eigenschaft der Nachbarpunkte durch die generalisierte mittlere Krümmung, um den Punkt in einem multi-dimensionalen Histogramm zu beschreiben. Dabei ist er invariant gegenüber der Orientierung und Lage der beschreibenden Oberfläche und robust

gegenüber unterschiedlichen Scanauflösungen und Rauschen in den Nachbarpunkten. Die Berechnung des FPFH für den Punkt p_q und seine Nachbarpunkte p_k ist im Algorithmus 1 und in der Abbildung 1 ersichtlich.

Berechnung des FPFH im Punkt p_q:

1. Berechnung des Simplified Point Feature Histogramms SPFH im Punkt p_q:

1.1 Bestimmung sämtlicher Nachbarpunkte p_k (siehe Abbildung 1a)).

1.2 Berechnung der k-Winkel (α, ϕ, θ) zwischen den Punkten p_q und p_k:

Indem ausgehend von den entsprechenden Flächennormalen n_q und n_k ein Darboux-Koordinatensystem in einem der Punkten p_q und p_k nach,

$$u = n_s, v = u \times \frac{(p_t - p_s)}{\|p_t - p_s\|_2}, w = u \times v$$

mit

falls $\arccos(n_q p_{kq}) \leq \arccos(n_k p_{qk}), p_{kq} = p_k - p_q, p_{qk} = p_q - p_k$

$p_s = p_q,\ n_s = n_q$ *und* $p_t = p_k,\ n_t = n_k$

sonst

$p_s = p_k,\ n_s = n_k$ *und* $p_t = p_q,\ n_t = n_q$

festgelegt wird (siehe Abbildung 1b)) und nach,

$$\alpha = v \cdot n_t, \phi = u \cdot \frac{(p_t - p_s)}{\|p_t - p_s\|_2}, \theta = \arctan(w \cdot n_t, u \cdot n_t)$$

die Winkel berechnet werden.

1.3 Erstellung des Histogramms für die k-Winkel (α, ϕ, θ).

2. Berechnung des Simplified Point Feature Histogramms SPFH pro Nachbarpunkt p_k:

Siehe Berechnungsschritt 1.

3. Berechnung des FPFH im Punkt p_q nach:

$$FPFH(p_q) = SPFH(p_q) + \frac{1}{k}\sum_{i=1}^{k}\frac{1}{w_k} \cdot SPFH(p_k)$$

mit

w_k = Distanz zwischen Punkt p_q und p_k.

Alg. 1 - Berechnung des FPFH (nach RUSU(2009)).

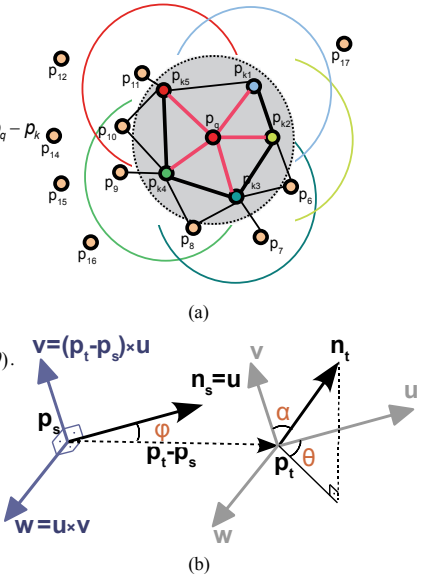


Abb. 1 - Berechnung des FPFH (RUSU, 2009).

4.5 Überwachte Self-Organizing Maps Klassifikation

Zur robusten und schnellen überwachten Klassifikation der 3D-Punkte wird das von Kohonen ab 1981 entwickelte Verfahren der Self-Organizing Maps (SOM) verwendet. Dieses ist ein wettbewerbsorientiertes selbstlernendes künstliches neuronales Netzwerk. Der Algorithmus hat nach KOHONEN (2001) im Wesentlichen zum Ziel, die Daten eines hochdimensionalen Featureraums auf ein üblicherweise zweidimensionales reguläres Raster von Knoten, sogenannten Neuronen, abzubilden. Dabei werden die originalen Informationen komprimiert, während die wichtigsten topologischen und metrischen Beziehungen der primären Datenstrukturen erhalten bleiben. Des Weiteren werden Objekte, die im n-dimensionalen Featureraum nahe beieinander liegen, im SOM räumlich nahe abgebildet (KOHONEN, 2001). Die SOM können somit zum einen für das Clustering verwendet werden und zum anderen

kann durch eine geringfügige Anpassung am Algorithmus auch eine überwachte Klassifikation durchgeführt werden.

4.6 Implementiertes Klassifikationssystem

Das vorgeschlagene Verfahren zur benutzergestützten Klassifikation einzelner 3D-Punkte, wurde unter Verwendung von Python zur Workflowsteuerung, der Point Cloud Library (POINTCLOUDS, 2012) zur Berechnung der Punktcharakteristika und der SOM Toolbox für Matlab von ALHONIEMI et al. (2012) als Prototyp-Klassifikationssystem implementiert.

5 Erste Untersuchungen

5.1 Verwendete urbane Laserscanningszene

Als Testdatensatz für erste Untersuchungen der Unterscheidbarkeit der berechneten Punktcharakteristika und der Klassifikationsgenauigkeit wurde eine mit der Leica ScanStation 2 erfasste terrestrische Laserscanningaufnahme des FHNW-Campus, Muttenz (CH) verwendet. Um die vorgeschlagene Punktklassifikationsmethode optimal untersuchen zu können, wurde eine Subszene mit ca. 1 Mio. Punkten extrahiert, welche die zu erkennenden Klassen „Strasse“, „Vegetation“, „Baumstamm/-ast“ und „Gebäudefassade“ in möglichst getrennter Form beinhaltet. Die Subszene mit einer Dimension von 50 x 50 x 15 Meter ist in Abbildung 2a) abgebildet. Nebst einer Vielzahl verschiedener Objektklassen, die in der Szene vorhanden sind, enthält die Punktwolke zudem unterschiedliche Punktauflösungen, fehlende Aufnahmegebiete und weist eine geringe Texturierungsqualität auf.



Abb. 2 – Verwendete urbane Laserscanningszene: Originale (a) und die klassifizierte (b) Punktwolke.

5.2 Untersuchung der Aussagekraft der berechneten Punktcharakteristika

Vor der Untersuchung der erreichbaren Klassifikationsgenauigkeit der vorgeschlagenen Klassifikationsmethode wurde unter Verwendung verschiedener graphischer Methoden der explorativen Statistik untersucht, ob die verwendeten Punktfeatures $[h, s, v, z, fpfh_{33}]$ eine Unterscheidung der gewählten Objektklassen überhaupt erlauben und welche Punktcharakteristika dazu am meisten beitragen.

Als erste Visualisierung wurden die vier Punktfeatures [h, s, v, z] des Trainingsdatensatzes in einer Scatterplotmatrix paarweise gegenübergestellt. Abbildung 3 zeigt auf, dass es die HSV-Farbwerte erlauben, die Objektklassen „Gebäudefassade“ sowie „Strasse“ von den Klassen „Baumstamm/-ast“ und „Vegetation“ einigermassen zu trennen, dass der Zenitwinkel der Flächennormalen jedoch mitberücksichtig werden muss, um die im HSV-Farbraum überlappenden Objektklassen „Baumstamm/-ast“ und „Vegetation“ voneinander trennen zu können. Des Weiteren ist ersichtlich, dass die Zenitwinkel der Vegetationsklasse stark streuen und somit dennoch einige Punktbeobachtungen innerhalb der Klasse Baumstamm/-ast zu liegen kommen. Der Grund der Streuung ist vor allem darauf zurückzuführen, dass viele Flächennormalen als Folge der abnehmenden Punktauflösung bei zunehmender Aufnahmedistanz falsch berechnet werden.

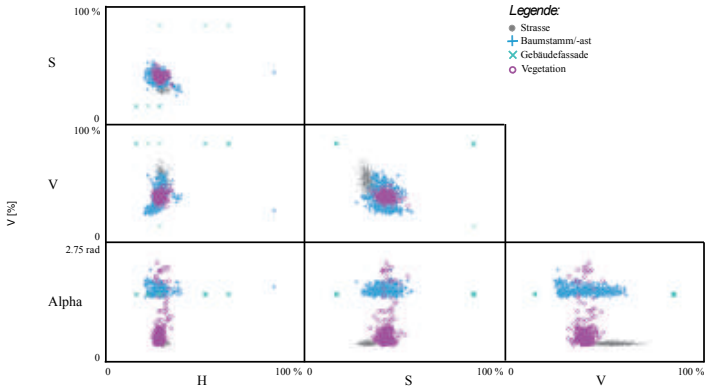


Abb. 3 - Scatterplotmatrix der HSV-Farbwerte und der Zenitwinkel zur Flächennormalen (mit k = 32 für die Flächennormalenberechnung).

Zur weiteren Untersuchung der Aussagekraft der berechneten Punktcharakteristika wurden die 37-dimensionalen Featurevektoren [$h, s, v, z, fpfh_{33}$] der Trainingsdaten mittels einer Grand Tour-Darstellung (ASIMOV, 1985; BUJA & ASIMOV, 1986) in einer interaktiven und dynamischen Grafik visualisiert. Diese Darstellungsmethode, bildet im Wesentlichen sämtliche Features eines Punktes durch eine Linearkombination auf einer Hyperebene niederer Ordnung ab. Sie dient dazu, in einer niedrigdimensionalen Projektion der Daten die multivariante Verteilung zu untersuchen und somit visuell Clusters erkennen zu können. Um eine optimale Clusteranalyse durchführen zu können, wurden die Projektionsparameter mittels einer Linear Discriminant Analysis (LDA) bestimmt (COOK et al., 1995; COOK & SWAYNE, 2007) und die Werte der Featurevektoren ins orthogonale Koordinatensystem ihrer Hauptkomponenten überführt. Abbildung 4 zeigt auf, dass unter Verwendung der HSV-, Zenitwinkel- und FPFH-Werte sämtliche Punkte einer Objektklasse der Trainingsdaten saubere, voneinander getrennte, Clusters bilden und somit die verwendeten Punktcharakteristika geeignet sind, die zu detektierenden Objektklassen voneinander zu unterscheiden.

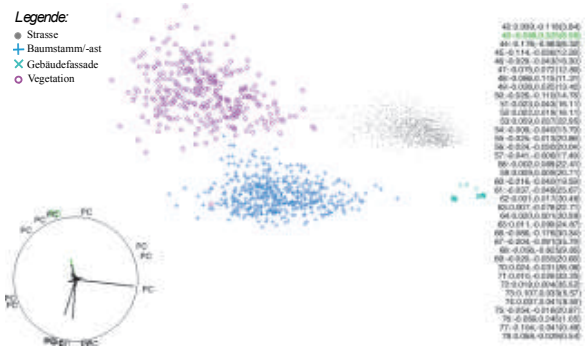


Abb. 4 - Grand Tour-Darstellung der HSV-, Zenitwinkel- zur Flächennormalen und FPFH-Werte pro Punkt (mit k = 32 für die Flächennormalen- und FPFH-Berechnung).

5.3 Untersuchung der Klassifikationsgenauigkeit

Um die Klassifikationsgenauigkeit numerisch ausweisen zu können, wurde manuell ein Referenzdatensatz erstellt. Der Datensatz diente dazu, um basierend auf der tatsächlichen Objektklassenzugehörigkeit jedes klassifizierten Punktes die in Tabelle 1 abgebildete Konfusionsmatrix erstellen sowie die Klassifikationsgenauigkeiten berechnen zu können. Tabelle 1 zeigt auf, dass die Herstellergenauigkeit für alle Klassen, ausser für die Klasse der Baumstämme/ -äste, über 90 Prozent liegt. Demgegenüber entspricht die Benutzergenauigkeit für alle Klassen, ausser für die Klassen der Baumstämme/-äste und der Vegetation, über 98 Prozent. Die hohe Anzahl an fälschlicherweise als Vegetation und Baumstamm/-ast klassifizierten Strassenpunkten lässt sich dabei auf die Tatsachen zurückführen, dass viele Strassenpunkte durch die ungenügende Texturierung der Punktwolke eine grüne Farbe aufweisen und dass als Folge der unterschiedlichen Punktauflösung die Zenitwinkel zur Flächennormalen sowie die FPFH falsch berechnet werden. Des Weiteren zeigt die Analyse des Klassifikationsresultats der Objektklasse Baumstamm/-ast, dass die Mehrheit der Falschklassifikationen auf die Schwierigkeit bei der Unterscheidung zwischen Blättern und Ästen zurückzuführen ist.

		Referenzdaten					
		Strasse	Baumstamm/ -ast	Gebäude- fassade	Vegetation	Total	Benutzer- genauigkeit [%]
Klassifikation	Strasse	32405	13	59	319	32796	99
	Baumstamm/-ast	200	421	36	31	688	61
	Gebäudefassade	13	17	4237	35	4302	98
	Vegetation	2320	748	373	7331	10772	68
Total		34938	1199	4705	7716	48558	
Herstellergenauigkeit [%]		93	35	90	95		
Klassengenauigkeit [%]		96	48	94	82		

Gesamtgenauigkeit [%]	91

Tab. 1 – Konfusionsmatrix und Genauigkeiten der Klassifikation (mit k = 32 für die Flächennormalen- und FPFH-Berechnung).

6 Fazit und Ausblick

Mit dieser Arbeit wurde eine Methode vorgestellt, mit welcher texturierte Punkte einer komplexen urbanen Laserscanningszene nach gewünschten Objektklassen klassifiziert werden können. Dabei konnte unter Verwendung der Punktcharakteristika des Zenitwinkels zur Flächennormalen, des Fast Point Feature Histogramms und des HSV-Farbwertes pro Punkt aufgezeigt werden, dass sich Punkte einer komplexen urbanen Laserscanningszene nach den Objektklassen „Strasse", „Baumstamm/-ast", „Gebäudefassade" und „Vegetation" klassifizieren lassen. Die zentrale Erkenntnis dieser Arbeit ist, dass die Punktcharakteristika so gewählt und korrekt berechnet werden müssen, dass eine Objektklassenunterscheidung gewährleistet werden kann und somit die Klassifikationsmethode eine sekundäre Rolle beim Erreichen optimaler Resultate einnimmt. Wie in den Resultaten ersichtlich, führt die zu geringe Berücksichtigung von Nachbarpunkten bei der Flächennormalenberechnung in Gebieten mit geringer Punktdichte dazu, dass viele Punkte falsch klassifiziert werden. Um diese Problematik der Flächennormalenberechnung beheben zu können, ist in künftigen Arbeiten vorgesehen, die verwendete Flächennormalenberechnungsmethode durch eine robustere Methode zu ersetzen. Eine dabei prüfenswerte Methode wäre, in Anlehnung an BRODU & LAGUE (2012), die verwendeten Punktcharakteristika über mehrere Stufen der Nachbarschaftsberücksichtigung zu berechnen, um so die tendenzielle Flächenneigung und -form über mehrere Auflösungsstufen zu erhalten.

7 Literaturverzeichnis

ALHONIEMI, E., HIMBERG, J., PARHANKANGAS, J. & VESANTO, J., 2012. SOM Toolbox. Verfügbar unter: http://www.cis.hut.fi/somtoolbox/ [Zugegriffen 24. April 2012].

ASIMOV, D., 1985. The grand tour: A Tool for Viewing Multidimensional Data. *SIAM J. Sci. Stat. Comput.,* **6** (1), S.128–143.

BERKMANN, J. & CAELLI, T., 1994. Computation of Surface Geometry and Segmentation using Covariance Techniques. *IEEE transactions on pattern analysis and machine intelligence*, **16** (11), S.1114 –1116.

BIOSCA, J.M. & LERMA, J.L., 2008. Unsupervised Robust Planar Segmentation of Terrestrial Laser Scanner Point Clouds Based on Fuzzy Clustering Methods. *ISPRS Journal of Photogrammetry and Remote Sensing*, **63** (1), S.84–98.

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**, S.121–134.

BUJA, A. & ASIMOV, D., 1986. Grand Tour Methods: An Outline. *Computing Science and Statistics*, **17**, S.63-67.

COOK, D., BUJA, A., CABRERA, J. & HURLEY, C., 1995. Grand Tour and Projection Pursuit. *Journal of Computational and Graphical Statistics*, **4**, S.155–172.

COOK, D. & SWAYNE, D.F., 2007. *Interactive and Dynamic Graphics for Data Analysis*. *With R and GGobi.* New York: Springer.

HOHMANN, B., KRISPEL, U., HAVEMANN, S. & FELLNER, D., 2009. Cityfit: High-quality Urban Reconstructions by Fitting Shape Grammars to Images and Derived Textured Point Clouds. In: ISPRS, *3d-Arch 2009: „3D Virtual Reconstruction and Visualization of Complex Architectures"*. Trento, Italien 25-28. Februar 2009.

KOHONEN, T., 2001. *Self-Organizing Maps*. 3. Aufl.. Heidelberg: Springer.

LIM, E.H. & SUTER, D., 2007. Conditional Random Field for 3D Point Clouds with Adaptive Data Reduction. In: IEEE, *2007 International Conference on Cyberworlds*. Hannover, Deutschland 24-26. Oktober 2007.

MATTI, E., 2012. *Geometrische und farbbasierte Punktwolkenklassifizierung von komplexen urbanen Laserscanningszenen mittels überwachter SOM-Klassifikation*. Projektarbeit Master Studium. Muttenz: Institut für Vermessung und Geoinformation, FHNW.

POINTCLOUDS, 2012. PCL. Verfügbar unter: http://pointclouds.org/ [Zugegriffen 24. April 2012].

RUSU, R.B., 2009. *Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments*. Ph.D. Computer Science Department, Technische Universität München, Germany.

RUSU, R.B., HOLZBACH, A., BLODOW, N. & BEETZ, M., 2009. Fast Geometric Point Labeling using Conditional Random Fields. In: IEEE/RSJ, *International Conference on Intelligent Robots and Systems*. St Louis, Missouri 11-15. Oktober 2009.

VOSSELMAN, G., 2009. Advanced Point Cloud Processing. In: *Photogrammetrische Woche 2009*. Stuttgart, Deutschland 7-11. September 2009.

WAN, G. & SHARF, A., 2012. Grammar-based 3D Facade Segmentation and Reconstruction. *Computers & Graphics*, **36** (4), S.216 – 223.

3D-Bildanalyse mit Scene Flow-Algorithmen

SILVIO TRISTRAM[1] & RALF REULKE[2]

Im Bereich der Bewegungsanalyse werden Verfahren angewandt, die neben dem Optical Flow auch den Scene Flow ermitteln können. Dieser setzt sich aus einer dreidimensionalen Beschreibung der Bewegung in einer Szene zusammen. Die dabei eingesetzten Verfahren zur Erzeugung der einzelnen Bewegungsvektoren basieren zumeist auf der Nutzung einer Stereokamera und sind somit nur eingeschränkt einsetzbar.
Im Rahmen dieses Beitrages soll ein Ansatz präsentiert werden, der eine Abwandlung des ICP-Algorithmus darstellt. Die Schätzung des Scene Flows wird somit auf Grundlage von 3D-Punktwolken durchgeführt und ist daher weitgehend unabhängig vom Aufnahmesystem.
Der ICP-Algorithmus wird normalerweise zum Stitching mehrerer Punktwolken verwendet. Deshalb wurde er speziell dahingehend modifiziert, dass er auch die Bewegung von nicht-starren Objekten ermöglicht. Ein besonderer Aspekt liegt auf der Erkennung der Körperbewegungen von Menschen. Bei der Umsetzung wurde speziell die Performanz durch die Nutzung von Multithreading und eines eigens implementierten OcTrees berücksichtigt.

1 Einleitung

Die Analyse von Bewegungen auf Grundlage zweidimensionaler Bilddaten ermöglicht die Erfassung der Handlungen in einer Szene. Der resultierende optische Fluss kann in unterschiedlichen Anwendungsfällen, wie der autonomen Navigation computergestützter Systeme oder Motion Capturing-Verfahren, genutzt werden.
Der optische Fluss hat jedoch den Nachteil, dass er lediglich eine zweidimensionale Beschreibung der Bewegung aus Sicht der Kamera wiederspiegelt und der perspektivischen Verzerrung unterliegt. Aus diesem Grund wurde dieser Ansatz dahingehend erweitert, dass die geschätzten Bewegungsvektoren dreidimensional sind und somit auch eine Bewegung in der Tiefe beinhalten. Da dadurch eine Beschreibung der gesamten Szene folgt, wird das Ergebnis als Scene Flow bezeichnet.
Häufig wird dabei der Einsatz einer Stereokamera vorausgesetzt, da viele Ansätze auf dieser Technik basieren. Jedoch existieren weitere Möglichkeiten wie Lichtschnittverfahren und Laserscanner, die eine alternative Akquirierung dreidimensionaler Bilddaten ermöglichen.
Im Folgenden wird daher eine Abwandlung des ICP-Algorithmus vorgestellt, der direkt auf Punktwolken arbeitet. Somit ist der vorgestellte Ansatz unabhängig von der eingesetzten Aufnahmetechnik.

2 Problemstellung

Das Ziel des hier vorgestellten Ansatzes soll die Schätzung der Bewegung einzelner Punkte einer Punktwolke *PC* sein. Es soll somit eine Aussage darüber getroffen werden, wohin sich ein Punkt vom Zeitpunkt t zu Zeitpunkt t+1 bewegt hat. Dabei wird speziell darauf Wert gelegt, dass auch die Bewegung von Menschen, also nicht-starren Körpern, ermöglicht wird. Im Folgenden

1) Silvio Tristram, Humboldt Universität zu Berlin, silvio.tristram@informatik.hu-berlin.de
2) Ralf Reulke, Humboldt Universität zu Berlin, reulke@informatik.hu-berlin.de

werden die zugrunde liegenden Algorithmen auszugsweise vorgestellt und auf Ihre Probleme eingegangen.

2.1 ICP-Algorithmus

Der Iterative Closest Point-Algorithmus (BESL & MCKAY) kann verwendet werden, um eine Bewegungsschätzung durchzuführen. Das Verfahren setzt voraus, dass es sich um einen starren Körper handelt und dieser zu beiden Zeitpunkten vollständig erfasst wurde.
Während der Ausführung findet eine iterative stetige Annäherung statt, in der die Bewegung durch einen Translationsvektor $t = (t_x, t_y, t_z)^T$ und einer Rotationsmatrix R (gebildet aus den Rotationen α_x, α_y, α_z) beschrieben wird. In einer Schleife werden dabei nacheinander mehrere Schritte ausgeführt.
Zunächst wird bei einem Matching - auf Grundlage der aktuell geschätzten Bewegung - eine Zuordnung der Punkte aus den unterschiedlichen Zeitpunkten auf Basis ihrer Distanz getroffen. Als nächstes wird diese genutzt, um eine Schätzung der Bewegung unter dem Aspekt der Minimierung des Abstandes der gefundenen Matches zu berechnen. Diese wird anschließend verwendet, um die Schätzung der Gesamtbewegung zu verbessern.

Pseudocode:

$t_x = t_y = t_z = \alpha_x = \alpha_y = \alpha_z = 0$
while *no break condition full filled* ***do***
  $M = Matching(PC_t, PC_{t+1}, t_x, t_y, t_z, \alpha_x, \alpha_y, \alpha_z)$
  $compute\ t'_x, t'_y, t'_z, \alpha'_x, \alpha'_y, \alpha'_z\ by\ M\ w.r.t.\ minimize$ point cloud distance
  $update\ t_x, t_y, t_z, \alpha_x, \alpha_y, \alpha_z\ by\ t'_x, t'_y, t'_z, \alpha'_x, \alpha'_y, \alpha'_z$
end while
return $t_x, t_y, t_z, \alpha_x, \alpha_y, \alpha_z$

Somit resultiert ein einziger Parametersatz zur Beschreibung der Bewegung für die gesamte Punktwolke.

2.2 TrICP-Algorithmus

Wie bereits erwähnt, setzt der ICP-Algorithmus voraus, dass die Objekte, deren Bewegung geschätzt werden soll, vollständig sein müssen. Die Punktewolken müssen also deckungsgleich sein. Dies ist in der Praxis jedoch häufig nicht gegeben, weshalb in solchen Fällen der Trimmed Iterative Closest Point-Algorithmus (CHETVERIKOV, SVIRKO, STEPANOV & KRSEK) eingesetzt werden kann.
Neben der Nutzung einer effizienten Datenstruktur unterscheidet sich dieser Ansatz vor allem dahingehend, dass bei der Minimierung des Abstandes nicht alle Matchings verwendet werden, sondern nur solche, die eine geringe Distanz aufweisen.

Pseudocode:

$t_x = t_y = t_z = \alpha_x = \alpha_y = \alpha_z = 0$
while *no break condition full filled* ***do***
$M = Matching(PC_t, PC_{t+1}, t_x, t_y, t_z, \alpha_x, \alpha_y, \alpha_z)$
sort M by distance
discard Matchings with high distance
compute $t'_x, t'_y, t'_z, \alpha'_x, \alpha'_y, \alpha'_z$ *by M w.r.t. minimize point cloud distance*
update $t_x, t_y, t_z, \alpha_x, \alpha_y, \alpha_z$ *by* $t'_x, t'_y, t'_z, \alpha'_x, \alpha'_y, \alpha'_z$
end while
return $t_x, t_y, t_z, \alpha_x, \alpha_y, \alpha_z$

Durch diese Veränderung ist es möglich die Bewegung auch dann zu schätzen, wenn die erfassten Punktwolken aufgrund von Sichtschatten nicht deckungsgleich sind. Jedoch bleibt nach wie vor das Problem, dass nur eine einzige Bewegung für die gesamte Punktwolke berechnet wird.

3 Flexible ICP

Zur Realisierung einer Bewegungsanalyse nicht-starrer Körper ist es erforderlich, eine separierte Schätzung durchzuführen. Die im Folgenden geschilderten Modifikationen sollen es ermöglichen, dass der ICP- bzw. TrICP-Algorithmus eine Bewegungsanalyse nicht-starrer Körper durchführt. Aufgrund dieser Unabhängigkeit wird die Bezeichnung Flexible ICP verwendet.

Im Folgenden werden zur besseren Visualisierung anstelle von Punktwolken zweidimensionale Kurven verwendet. Dies dient lediglich der besseren Veranschaulichung und stellt keine Einschränkung dar. Als Grundannahme wird zunächst angenommen, dass die Bewegung eines starren Körpers vorliegt.

In Abbildung 1 ist anhand einer Kurve exemplarisch die Bewegung und Verformung dargestellt. Die schwarze Kurve stellt dabei den Zustand vom Zeitpunkt t dar und die rote Kurve den Zustand vom Zeitpunkt $t + 1$. Nach einigen Iterationen resultiert eine grobe Schätzung der Bewegung, die je nach Wahl des Algorithmus ein abweichendes Ergebnis liefert (vgl. Abb. 2 und Abb. 3). Diese geschätzten Positionen werden sich nun nicht mehr gravierend ändern.

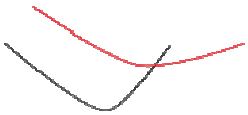


Abb. 1: Ausgangszustand

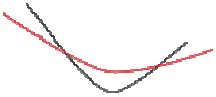


Abb. 2: Resultat beim ICP

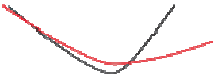


Abb. 3: Resultat beim TrICP

Zur Verbesserung dieser Schätzung soll nun eine Unterteilung durchgeführt werden. Zur Umsetzung wird für jeden Punkt aus der Punktwolke PC_t in jeder Iteration gespeichert, mit welchem Punkt aus Punktwolke PC_{t+1} er gematcht wurde.
Nach einer zuvor festgelegten Anzahl *n* an Iterationen wird in jedem Schleifendurchlauf nach dem Matching-Schritt geprüft, ob die aktuelle Zuordnung in den vergangenen *m* Iterationen mindestens *r*-mal aufgetreten ist. Sollte dies der Fall sein und der Abstand der zugeordneten Punkte gering ausfallen, so wird für diese Punkte (Fixpunkte) die aktuelle Schätzung der Bewegung fest zugewiesen und sie fallen aus der weiteren Analyse aus. Diese Kriterien treten in Bereichen von Überschneidungen auf (vgl. Abb. 4 und Abb. 5).

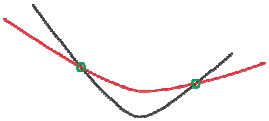


Abb. 4: Gefundene Fixpunkte beim ICP



Abb. 5: Gefundene Fixpunkte beim TrICP

Da die Fixpunkte wegfallen, entstehen an diesen Stellen Lücken. Nach deren Bestimmung wird eine Segmentierung der Punktwolke durchgeführt, wodurch die Punktwolke an diesen Stellen aufgeteilt wird. Die Bewegungsanalyse wird anschließend auf den Bruchstücken getrennt voneinander fortgesetzt, wodurch eine separate Schätzung folgt (vgl. Abb. 6 und Abb. 7).

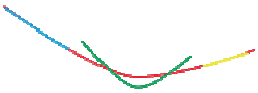


Abb. 6: Segmentierung und Annäherung (ICP-Kern)

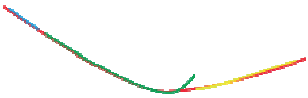


Abb. 7: Segmentierung und Annäherung (TrICP-Kern)

Durch wiederholte Durchführung dieses Vorgehens wird die Punktwolke zu immer kleineren Bruchstücken zerlegt, wodurch eine stetige Abstandsreduzierung der Punkte folgt.

Pseudocode:

$t_x = t_y = t_z = \alpha_x = \alpha_y = \alpha_z = 0$
while *no break condition full filled* ***do***
  $M = Matching(PC_t, PC_{t+1}, t_x, t_y, t_z, \alpha_x, \alpha_y, \alpha_z)$
  if (iteration >= n) {
    check for fixpoints
    segmentation

}
[$sort\ M\ by\ distance$]
[$discard\ Matchings\ with\ high\ distance$]
$compute\ t'_x, t'_y, t'_z, \alpha'_x, \alpha'_y, \alpha'_z\ by\ M\ w.r.t.\ minimize\ point\ cloud\ distance$
$update\ t_x, t_y, t_z, \alpha_x, \alpha_y, \alpha_z\ by\ t'_x, t'_y, t'_z, \alpha'_x, \alpha'_y, \alpha'_z$
end while
return $t_x, t_y, t_z, \alpha_x, \alpha_y, \alpha_z$

4 Beschleunigung des Verfahren

In der praktischen Umsetzung wurde Wert darauf gelegt, dass die Abarbeitung des Algorithmus möglichst schnell erfolgt. Das Aufwändigste ist die Ermittlung des Matches und die Durchführung der Segmentierung. Dabei wird zum einen der nächste Nachbar ermittelt, zum anderen werden Bereichsabfragen durchgeführt.
Beide Aufgaben können mit Hilfe einer räumlichen Datenstruktur effizient durchgeführt werden. Zum Einsatz kam dabei eine eigens implementierte OcTree-Datenstruktur. Es hat sich dabei gezeigt, dass eine vollständige Aufspaltung des Raumes nicht der optimalen Lösung entspricht. Dies ist darauf zurückzuführen, dass die Navigation durch den Baum ebenfalls zeitliche Kosten erzeugt.
Stattdessen wurden in den Blättern mehrere Punkte in einer unsortierten Liste verwaltet. Zu einer Aufteilung der Blattknoten kam es, wenn eine festgelegte Größe überschritten wird. Da auf beiden Punktwolken unterschiedliche Operationen durchgeführt wurden (in PC_t werden Bereichsabfragen durchgeführt, während in PC_{t+1} nächste Nachbarn ermittelt werden), war das optimale Limit abweichend (PC_t: 60 Punkte, PC_{t+1}: 100 Punkte).
Eine weitere Beschleunigung wurde durch eine Parallelisierung der Matching-Operation erreicht. Bei der Umsetzung wurde diese Teilaufgabe durch Multithreading auf die zur Verfügung stehenden CPU-Kerne verteilt, um diese bestmöglich auszulasten.

5 Resultate

Zur Überprüfung der Funktionalität des Flexible ICP-Ansatzes wurden verschiedene Punktwolken verwendet. Dabei wurden zur Akquirierung der Bilddaten das Kinect System und Stereokameras (Die Tiefenkarte wurde mit dem Semi Global Matcher ([HIRSCHMULLER]) ermittelt) genutzt.
In Abbildung 8 und 9 sind exemplarisch dazu zwei Punktwolken vom Zeitpunkt t und t + 1 dargestellt.



Abb. 8: Punktwolke vom Zeitpunkt t



Abb. 9: Punktwolke vom Zeitpunkt t + 1

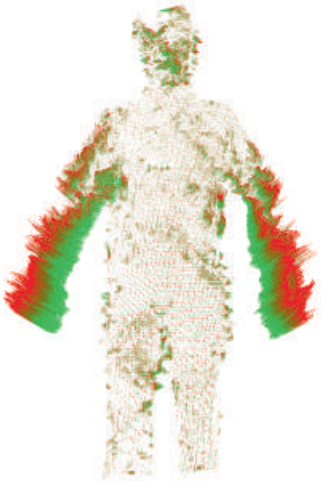


Abb. 10: Dichtes Vektorfeld

Bei der Ausführung des Algorithmus hat sich gezeigt, dass die geschätzten Zielpositionen nahe bei einander liegender Punkte aus PC_t aufgrund der separierten Betrachtung leicht schwanken. Jedoch wird davon ausgegangen, dass sie tatsächlich einen ähnlichen Verlauf aufweisen. Daher wurde zusätzlich eine Angleichung benachbarter Bewegungsvektoren durch eine Mittelung durchgeführt. Das Resultat des Flexible ICP-Algorithmus mit TrICP-Kern ist in Abbildung 10 dargestellt (Parameterwahl: n = 6, m = 6, r = 3).

In der Darstellung werden die Bewegungsvektoren durch Linien dargestellt. Die grünen Enden kennzeichnen dabei den Ausganspunkt und die roten Enden die geschätzten Zielpositionen. Sowohl die Bewegung der Arme als auch die Drehung des Kopfes wurden erfasst.

6 Future Work

Trotz der Modifikationen unterliegt auch diese Variante des ICP-Algorithmus dem Problem der lokalen Tiefpunkte. So können zu große Bewegungen zu Fehlern bei der Schätzung der Bewegung führen.
Zur Reduzierung der Fehler empfiehlt sich der Einsatz von Skelettmodellen. Dadurch können aufgenommene Punktwolken vor dem Start der Bewegungsanalyse bereits in verschiedene Partien untergliedert und dann passend gematcht werden. Ein weiterer Vorteil dieses Vorgehens ist eine Reduzierung des Rechenaufwandes.

7 Literaturverzeichnis

BESL, P.J. & MCKAY, N.D.: A method for registration of 3-D shapes. In: IEEE Transactions on pattern analysis and machine intelligence (ISSN 0162-8828) 14 (1992), Nr. 2, S. 239–256

CHETVERIKOV, D.; SVIRKO, D.; STEPANOV, D. & KRSEK, P.: The trimmed iterative closest point algorithm. In: Pattern Recognition, 2002. Proceedings. 16th International Conference on (ISSN 1051-4651) Bd. 3 IEEE, 2002, S. 545–548

HIRSCHMULLER, H.: Stereo vision in structured environments by consistent semi-global matching. In: Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on (ISBN 0769525970) Bd. 2, 2006, S. 2386–2393

Automated Multi-temporal Landslide Identification in Kyrgyzstan Using Satellite Remote Sensing Data

ROBERT BEHLING, SIGRID ROESSNER, HANS-ULRICH WETZEL & HERMANN KAUFMANN

In Southern Kyrgyzstan large landslides represent a wide-spread and regularly occurring natural hazard Due to the large extent of the affected area, automated analysis of multi-temporal satellite data represents the only way for efficient landslide detection. A RapidEye image database could be established for an area of about 12,000 km² containing multiple acquisition dates between 2009 and 2012. The developed automated approach for landslide detection is based on the analysis of temporal trajectories of NDVI values derived from the entire multi-temporal data stack. Specific temporal footprints of vegetation changes enable identification of landslide events. Applying this approach to the whole study area, about 250 landslides could be automatically detected between 2009 and 2012. Subsequent field checks have revealed that the vast majority of events have been correctly identified. Thus, the developed approach is capable of automatically detecting different kinds of mass movements under diverse natural conditions.

1 Introduction

In Southern Kyrgyzstan ongoing tectonic processes and high topographic relief cause natural hazards related to active mountain building affecting the Pamir-Tienshan orogenic system. Besides earthquakes, floods, and avalanches, landslides are one of the major natural hazards in this area and are especially concentrated along the Eastern rim of the Fergana Basin (Figure 1). In this densely populated mountainous region almost every year large landslides endanger human lives and infrastructure. Therefore local authorities responsible for disaster management and risk reduction have a big need for objective and spatially differentiated hazard assessment as a main prerequisite for sensible decision making related to precaution and mitigation. In this context, multi-temporal landslide inventories enabling the analysis of past and recent landslide activity as well as fostering an improved process understanding are of special importance (FELL et al. 2008; VAN WESTEN et al. 2008). In Kyrgyzstan, landslide inventories have been carried out since the 1950'ies, whereas approximately 3000 landslides have been reported by the local authorities. However, regular inventories have been limited to the time period between 1968 and 1992 and focused on areas in the vicinity of populated places. The resulting reports (e.g., IBATULIN 2011) contain verbal descriptions of major events of the last 50 years. However, in most of the cases precise geographic locations of the events are missing. Thus, the existing knowledge on landslide events is incomplete in space and time and leaves the need for a systematic multi-temporal landslide inventory as a basis for objective and comprehensive hazard assessment.

Landslide inventories were first accomplished by field based mapping carried out by geological experts. With the wider availability of aerial photographs, remote sensing methods were introduced in the mapping procedure starting out with visual interpretation. This labor-intensive method was also applied to satellite remote sensing data whereas most of the time their suitability had been limited to the assessment of large events due to the rather coarse spatial resolution of these data. Only with the recent availability of high-resolution satellite remote sensing data, they have been widely replacing aerial photographs (VAN WESTEN et al. 2008). Due to the commonly used time consuming method of visual analysis of remote sensing data, often only mono-temporal inventories have been established, especially if large areas are affected by landslides. However, because of ongoing process activity such

inventories usually are outdated in the moment of their completion. Therefore, there has been a big need for developing methods for automated landslide detection based on multi-temporal satellite remote sensing data. In the past decade the increasing availability of optical sensors with high spatial resolution has resulted in first approaches identifying landslides in an (semi-) automated way based on visible signs of surface cover changes caused by slope failures (GUZETTI et al. 2012). But up to now there is a lack of approaches suitable of deriving multi-temporal inventories for large areas mainly due to the limited spatial and temporal availability of optical remote sensing data.

In this context, the Rapid Eye satellite remote sensing data characterized by high temporal and high spatial resolution of 6.5 m open up new opportunities for automated landslide detection with the goal of establishing multi-temporal landslide inventories for large areas whereas the possible high temporal frequency of data acquisition allows a systematic assessment of landslide events independent from the knowledge about specific triggering events. The presented work is based on the RESA (RapidEye Science Archive) program (grant number 424) enabling customized data acquisition since 2010 for the study area at the Eastern rim of the Fergana Basin described in section 2. The resulting multi-temporal database (see section 2) forms the basis for the development of an approach for automated landslide detection based on the analysis of surface cover changes related to landslide activity (see section 3). The results obtained for the study area at the Eastern rim of the Fergana Basin are described in section 4 and further discussed in section 5.

2 Study Area and Database

At the Eastern rim of the Fergana Basin landslides are mainly occurring within the foothills of its surrounding mountain ranges in an elevation range between 700 and 2000 m (Figure 1). The majority of these landslides occur in form of rotational and translational slides in weakly consolidated Quaternary and Tertiary sediments. The lithology of the landslides represents the main factor in determining their type of movement. One type is related to the presence of massive Quaternary loess units of up to 50 m thickness that moves rapidly down the slope with several meters per second. These avalanche-like movements are especially dangerous because of their great destructive power and their sudden occurrence. The other main type is represented by slowly moving (up to several meters per day) landslides occurring in Meso- and Cenozoic sediments with intercalated clays. Independent from their specific type all of these landslides result in degradation and decrease of scarce human living space limited by the high mountain ranges of Kyrgyzstan. The destructive power of landslides affects transportation and energy infrastructure as well as settlements and often results in evacuation and relocation of people (ROESSNER et al. 2005).

The RapidEye data which are acquired with high temporal and spatial resolution in a swath of 70 km enable detailed and efficient analysis of landslide activity for large areas. For this purpose, orthorectified Level-3A data products have been selected in order to minimize pre-processing efforts and to foster an approach which has the potential for global and operational application. The studied region shown in Figure 1 comprises an area of approx. 12.000 km² which is covered by 21 Level-3A RapidEye tiles. The RESA data grant has allowed acquiring data in pre-defined time periods of high process activity with a high temporal repetition rate of up to one acquisition per month since 2010. Together with archive data which have been available for selected parts of the study area since 2009, up to 25 acquisition dates are available for the whole time period from 2009 to 2012. The resulting multi-temporal database consists of approx. 500 Level-3A datasets representing about 20 acquisition dates during the main period of landslide activity between April and September (Figure 1). In areas of highest

process activity indicated in Figure 1 by the reported landslides and the landslide prone slopes the temporal coverage is even higher. This high temporal and at the same time high spatial resolution of the RapidEye data are of key importance for a detailed analysis of the current state of landslide activity and the comprehensive assessment of seasonal surface cover changes. Additionally, the high temporal repetition rate enables a rather precise temporal determination of landslides in the frame of a multi-temporal landslide inventory. The high spatial resolution allows detecting smaller events which may be an indicator for subsequent larger and more destructive slope failures. This way, the RapidEye data support identification of possible precursors for upcoming hazardous events which can support early warning efforts.

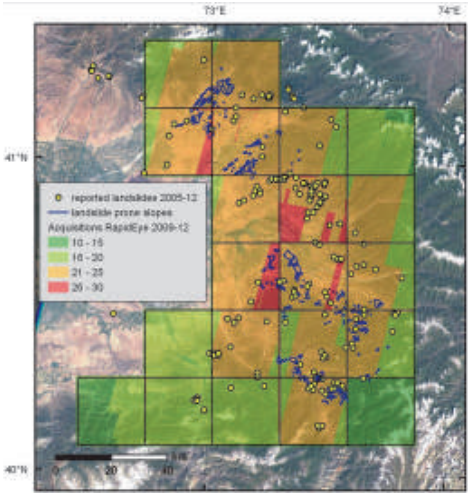


Figure 1: RapidEye coverage of the study area

3 Automated Landslide Detection

3.1 State of the Art

Existing approaches for automated landslide detection in optical remote sensing data mainly rely on the assumption that slope failures result in disrupted or absent vegetation cover. Their specific realizations represent varying degrees of automation and according to GUZETTI et al. 2012 they can be classified by methodological aspects (e.g. pixel-based, object-based) and the number of dates included in the analysis (post-event image analysis, bi-temporal change analysis of pre- and post-event images). The majority of these approaches have been developed for event-based inventories providing information about landslides related to a specific triggering event, such as earthquakes or hydrometeorological extreme events (e.g., tropical rainstorms). Most of the published studies have been targeted at areas smaller than 100 square kilometers. So far, there have been no attempts made to develop an automated approach for a larger area and for a systematic inventory of landslide events occurring independent from specific and well-known triggering events. Such an approach is needed for the study area at the Eastern rim of the Fergana Basin, since in this region landslide activity is caused by complex interactions between tectonic, geological, geomorphological and hydrometeorological factors leading to regularly reoccurring landslide events (ROESSNER et

al. 2005). Additionally, the limited existing knowledge about past landslide activity in this region requires an automated approach which can extended to a longer-term time series of multi-temporal image data originating from different sensors.

3.2 Overall Approach

The multi-temporal RapidEye database in combination with additional spatial information, such as digital elevation models and geological information forms the basis for the development of an automated approach for landslide identification in order to derive a GIS-based multi-temporal landslide inventory containing the individual landslide events in an object-based form including their qualitative and quantitative characterization. The main part of the approach represents the multi-temporal change detection allowing a separation between changes caused by landslide activity from other land cover changes (e.g. agriculture) and from artifact changes caused by geometric mismatches and radiometric differences between image data with different acquisition properties. In order to minimize such artifacts, automated change detection requires adequate pre-processing of the multi-temporal image database. Taking into account the high number of approx, 500 datasets, pre-processing has to be carried out in an automated and robust form. Section 3.3 describes the pre-processing methods which have been developed specifically in the frame of the change detection approach which is described in more detail in section 3.4.

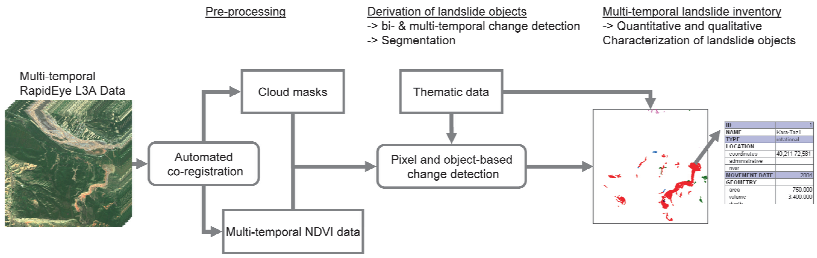


Figure 2: Automated approach for deriving a multi-temporal landslide inventory

3.3 Pre-processing

Before automated change detection can be applied to a multi-temporal image database, pre-processing of the remote sensing images has to be performed. Pre-processing usually includes the three steps of geometric co-registration, radiometric normalization and masking of irrelevant changes in order to generate a comparable dataset and minimize artifact changes (COPPIN et al. 2004; JIANYA et al. 2008; LU et al. 2010).

Although in the presented study methodological developments are based on the orthorectified Level-3A RapidEye data products, quality checks have shown that some geometric shifts of up to 50 m exist. In order to compensate for these mismatches an image-to-image co-registration method was applied. It performs an area-based correlation between two images in order to derive optimized shifting parameters for each image datasets without requiring another resampling of the already orthorectified datasets (BEHLING et al. 2012). In the result, a geometrically adjusted multi-temporal database was created which is characterized by a co-registration accuracy of about one pixel amounting to 5 m. This co-registered image database allows change detection with minimized artifact changes caused by geometric mismatches.

Besides geometric adjustment, the multi-temporal datasets also have to be corrected for radiometric differences caused by different acquisition parameters (e.g. atmospheric conditions, viewing angles, sun angles). Spectral indices, such as the normalized difference vegetation index (NDVI) are commonly used to reduce the effects of such radiometric

differences. However, without comprehensive atmospheric corrections or relative radiometric normalization even these indices show a slight variability of their values caused by the above mentioned radiometric differences. In case of the developed NDVI–based change detection approach (see section 3.4), the change of NDVI values caused by the removal or disturbance of vegetation related to slope failures is much higher than the variability of the NDVI values caused by radiometric differences. However, in order to obtain robust indicators for landslide-related surface changes, the developed change detection approach compensates for this remaining variability by implementing a multi-threshold change detection approach (see section 3.4 for more detail).

The presence of clouds and snow is another problem for automated change detection especially in humid areas of pronounced topographic relief, such as the Eastern rim of the Fergana Basin. They lead to detection of changes which are not related to land cover changes. In order to eliminate them, in a first step the cloud mask product supplied by RapidEye had been used. Analysis of these products has shown that dense clouds and snow cover are well represented. However, in part these products do not contain small and/or thin clouds and they also overestimate the size of the clouds because of a buffer which is applied to the identified cloud pixels. Therefore, a new approach for cloud detection has been developed using the original level 3A data. This approach consists of a double threshold technique based on the radiance values of the blue band and results in three classes of likelihood that a pixel is covered by a cloud. Based on this likelihood dataset, image segmentation is performed. The resulting segments are classified as clouds according to the areal coverage of the likelihood classes within the segments and additional size parameters. In the result this approach enables the detection of dense clouds and snow cover as well as thinner cloud layers and small amounts of snow which is not achieved by the standard RapidEye cloud-mask products.

3.4 Multi-temporal Change Detection for Landslide Identification

The developed multi-temporal change detection approach is based on the analysis of temporal trajectories of the NDVI derived from the pre-processed multi-temporal data stack. These temporal trajectories are obtained for every pixel across the time span of the entire data archive and thus allow analysis of vegetation cover changes over time rather than solely assess the absence of vegetation in a single dataset or the loss of vegetation between two datasets. These specific temporal footprints of vegetation changes (see Figure 3) enable identification of landslide events due to the short-term character of destruction of the vegetation cover caused by landslide events as well as longer-term re-vegetation rates as effects of the disturbance of the surface cover caused by slope failures.

In Figure 3 the derivation of the NDVI temporal trajectories representing the temporal surface cover changes in the analyzed time-period is exemplary illustrated for a landslide complex within the study area. In order to demonstrate typical temporal vegetation changes occurring over a longer time-period, the example is based on a multi-temporal NDVI database comprising image data acquired by the ASTER and RapidEye sensor between the years 2002 and 2012. Figure 3 shows the NDVI trajectory of a landslide which occurred in April 2004. The slope failure caused a severe surface disturbance and a destruction of the vegetation cover which is reflected in the abrupt decrease of the NDVI values between June 2003 and June 2004. After the landslide event the NDVI values remain low because of a slow re-vegetation rate caused by the disturbance and dislocation of soil which makes the landslide body susceptible to erosion and reactivation processes affecting the dislocated masses. This resulting typical shape of the NDVI trajectory is used to distinguish landslides from other temporal vegetation changes which can not be related to landslide processes, such as the other five land cover classes shown in Figure 3. Such a comprehensive assessment and comparison of temporal change patterns represented by the trajectories is not possible by state-of-the-art

change detection approaches which are limited to mono or bi-temporal analysis for landslide identification. The trajectories of Figure 3 show that it is difficult to distinguish between harvested fields and landslide-related vegetation change if only 2 or 3 multi-temporal datasets are considered in the analysis. The resulting so-called false-positives can be minimized by analyzing a longer time series of multi-temporal image data.

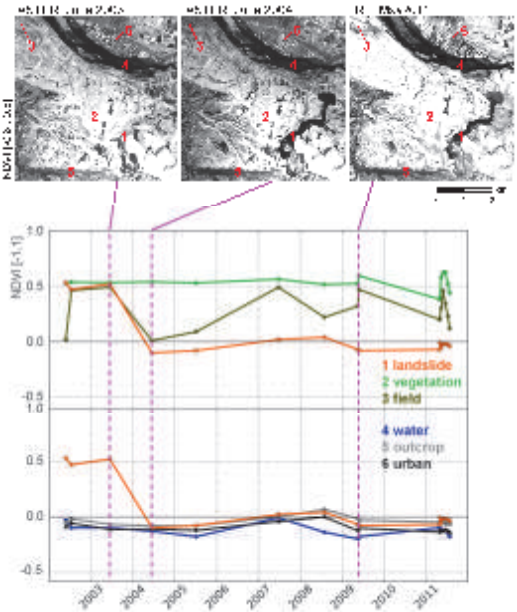


Figure 3: NDVI-based temporal trajectories representing typical vegetation changes related to land cover classes

In the following sections the developed approach for analyzing the temporal vegetation changes related to the occurrence of landslide activity will be described in more detail. The combined pixel- and object-based approach is divided into three main steps: 1) bi-temporal change detection, 2) segmentation based on the bi-temporal change result and 3) object-based multi-temporal change detection. The principle of the developed landslide identification approach is described for two subsequent multi-temporal image dataset. Then, this principle is applied to all subsequent image data pairs contained in the entire multi-temporal data stack resulting in objects of surface cover changes related to slope failures for all available time steps. Thus, the developed approach (Figure 4) allows the creation of a dynamic landslide inventory.

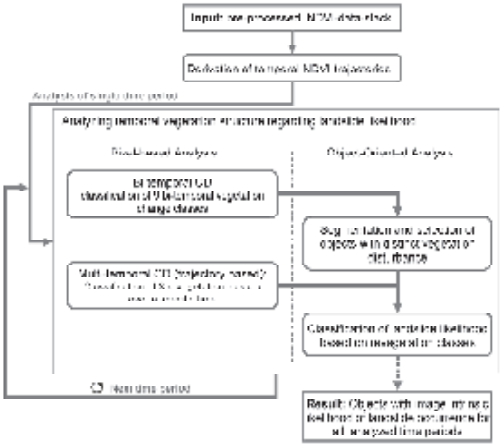


Figure 4: Multi-temporal landslide identification approach based on temporal NDVI trajectories

3.4.1 Multiple Threshold Based Bi-temporal Change Detection

Taking into account that a slope failure removes or degrades the until then largely undisturbed vegetation cover, in a first step, the NDVI-based temporal trajectories are analyzed in respect to abrupt vegetation loss between two dates. For this purpose the NDVI is analyzed in any subsequent bi-temporal image pair resulting in the parameters *pre* and *post*. Combined with a normalized difference index of the NDVI of the two images, these parameters are used to classify the vegetation change regarding to landslide likelihood. In this context, a strong abrupt vegetation loss resulting in bare soil represents the ideal case in form of a high likelihood for landslide occurrence (index very high and $post < 0$). However, most of the natural processes do not create such distinct surface conditions. In contrast, they are rather fuzzy in their spatial and temporal dispersion as well as in their process intensity. Therefore, the classification of landslide likelihood is performed on a scale of 9 classes (see Figure 6c). This way, less pronounced surface cover changes related to landslides can also be recognized, e.g. if the vegetation cover is already less dense before landslide occurrence. This is often the case, if slope failures occur as secondary movements in already displaced masses which have not fully been re-vegetated yet. This multiple threshold change detection approach also reduces the susceptibility for radiometric differences contained in the NDVI values (see section 3.3).

3.4.2 Segmentation of Bi-temporal Landslide Likelihood Datasets

The pixel-based multiple threshold change detection described in section 3.4.1 enables a first view on areas where a vegetation disturbance took place which is likely to be caused by a slope failure. Based on this bi-temporal landslide likelihood dataset segmentation is performed in order to derive vegetation loss objects. This segmentation procedure is subdivided into two steps. First, all pixels which are assigned to any of the existing landslide likelihood classes and are spatially connected form individual segments. These segments represent objects which contain not only ideal change pixels but pixels belonging to any of the 9 classes. In a second step, a selection of these objects is performed by analyzing the proportion of the consisting landslide likelihood classes in a specific segment. In the result of this selection, objects with a strong vegetation disturbance are kept and objects containing a (too) high percentage of pixels characterized by non-ideal vegetation changes are eliminated. Thus, this approach ensures the derivation of landslides as single objects even if they

represent partly not the optimal case of strong vegetation disturbance such as landslides occurring in a sparsely vegetated older landslide body or landslides whose runout zones cover an already harvested field.

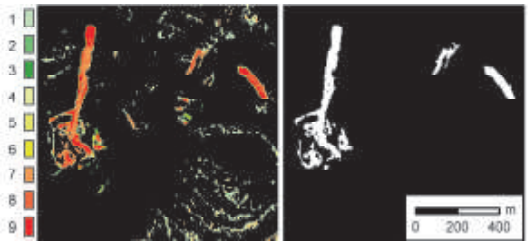


Figure 5: Segmentation based on bi-temporal change (left: subset of figure 6c; right: segmentation result)

3.4.3 Object-oriented Multi-temporal Change Detection

Despite the described multiple thresholding and subsequent object selection, a lot of initially identified objects remain which are characterized by bi-temporal vegetation changes similar to the ones typical for landslides. In this context, the main problems are fields that are vegetated in the pre-event image and harvested in the post-event image, exemplarily shown in the south western part of the subset in Figure 6.

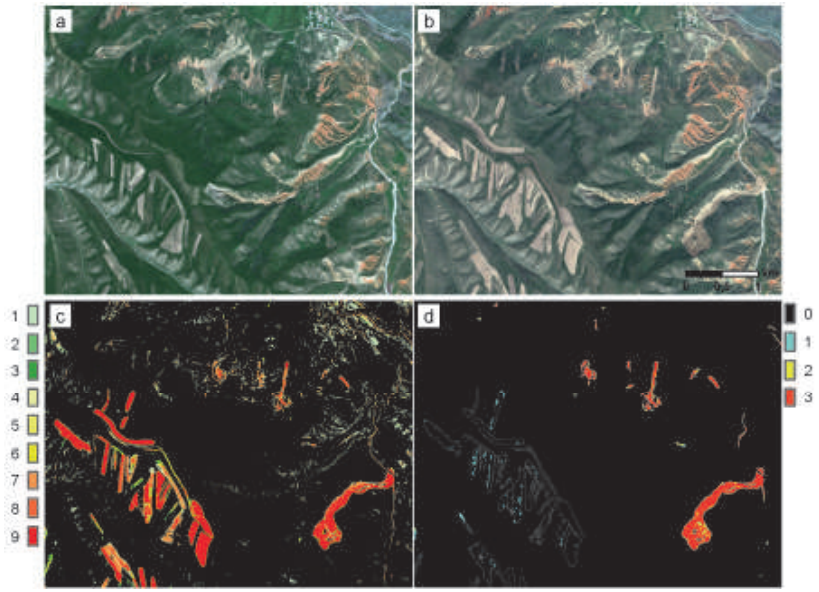


Figure 6: Analyzed vegetation changes between RapidEye data from 2009-05-26 (a) 2012-06-04 (b). c: Bi-temporal change (9 classes of vegetation loss: 1 low to 9 high); d: Multi-temporal change shown for all bi-temporal derived objects depicted as white polygons (4 re-vegetation classes: 0 very fast to 3 very slow)

In order to eliminate these false positives, the typically long-lasting re-vegetation process of landslides is considered. Based on the NDVI temporal trajectories four different classes of re-vegetation rates are distinguished in a specific time period after the bi-temporal event using a pixel-based classification approach. The considered time period can be adapted to the specific local and/or regional conditions influencing the re-vegetation rates (e.g., climate and soil properties). Based on the classification of re-vegetation, the landslide segments derived from bi-temporal analysis are further characterized according to their landslide likelihood. Figure 6d shows that fields are mainly characterized by higher re-vegetation rates (classes 0 or 1) than landslides (classes 2 or 3). This object-oriented classification based on the proportions of the re-vegetation classes leads to the final result of an image-intrinsic likelihood that this object represents a landslide.

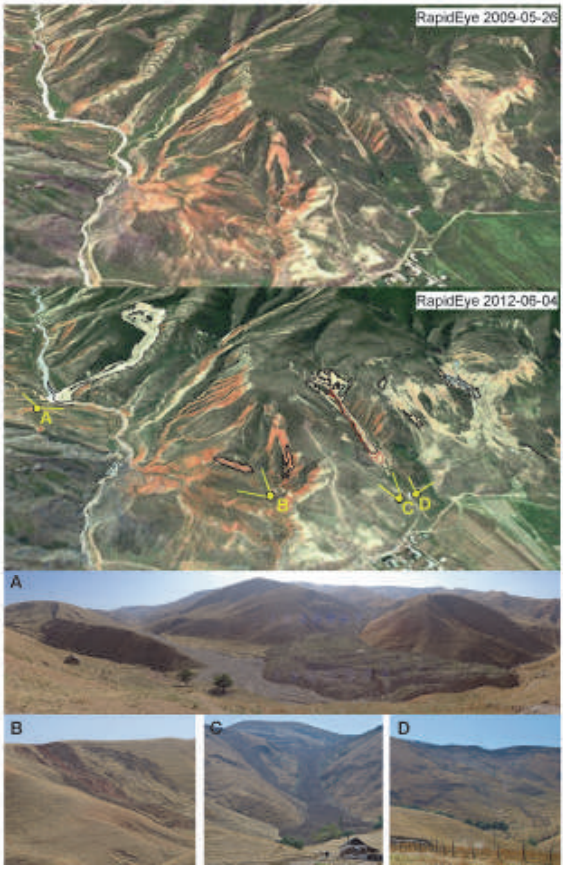


Figure 7: Mapping results in a 4 x 5 km subset of the study area. Detected landslide polygons depicted in a perspective view (black: 2010, red: 2012) and visited in a field survey in September 2012

4 Results

Results of the developed approach are exemplarily shown in Figure 7 for a 4 x 5 km subset of the study area which is also represented by Figure 6. Figure 7 depicts landslides as vectorized objects which have occurred in 2010 and 2012 overlaid on a perspective view of two RapidEye datasets. It can be seen that landslides of different sizes, shapes and in different stages of development (fresh failures, reactivation, relocation) can be automatically derived. Figure 7a shows a recent failure that happened in 2010 which is approx. 1.3 km long and up to 400 m wide and results in an accumulation zone of about 20 m thickness at the toe of the landslide. At the location of this landslide the RapidEye image of 2009 shows a closed vegetation cover and no clearly visible sign of an active landslide body can be detected. Besides these relatively fresh failures the approach also allows the detection of reactivated landslides and relocation of landslide material within recent landslide bodies covered by sparse vegetation. In contrast to the fresh failures the resulting vegetation disturbance is less distinct but still identifiable by the developed multi-temporal change detection approach. Figure 7b for example shows an enlargement of a landslide that has already existed in 2009. In the 2010 (black polygon) and 2012 (red polygon) images two reactivations of this already existing landslide had been detected resulting in an enlargement of the crown area of the landslide and displacement of material further down the slope. Figure 7 also contains some small polygons in the right part representing relocation of material caused by erosion processes within an active landslide complex. Figure 7c displays an example occurred within weakly consolidated sedimentary rocks whereas all other examples have occurred within loess. These results show that the developed approach based on multi-temporal RapidEye data is suitable for mapping recent landslide activity under varying natural conditions of land cover, lithology and landslide activity.

In a next step this approach has been applied to the entire multi-temporal RapidEye image database comprising the whole study area for the years from 2009 to 2012. For every analyzed time period a shapefile has been produced containing the image-intrinsic likelihood of a landslide occurring within this period. A visual evaluation of the derived objects resulted in the selection of about 250 objects most likely representing landslide events. The size of these objects is ranging between 500 and 250,000 square meters and the total area affected by these landslide amounts to 5.5 million square meters. About half of these events happened between 2009 and 2010. For the same period of time (2009-2012) the Ministry of Emergency Situations of Kyrgyzstan has only reported 40 events.

The obtained results have been verified during a four-week field survey which took place in September 2012. During this survey about 100 of the 250 detected landslides have been visited and in almost all of these cases the field checks confirmed recent surface changes related to landslides. Exceptions were primarily caused by the extraction of construction material (e.g., clay and gravel) at the bottom of hillslopes representing artificial mass movements. This field survey has revealed that the developed approach is capable of automatically detecting different kinds of landslides caused by a variety of slope processes, such as rotational and translational landslides and debris flows under diverse natural conditions. Thus, it could be shown that the approach can be used for reliable landslide detection at a regional scale.

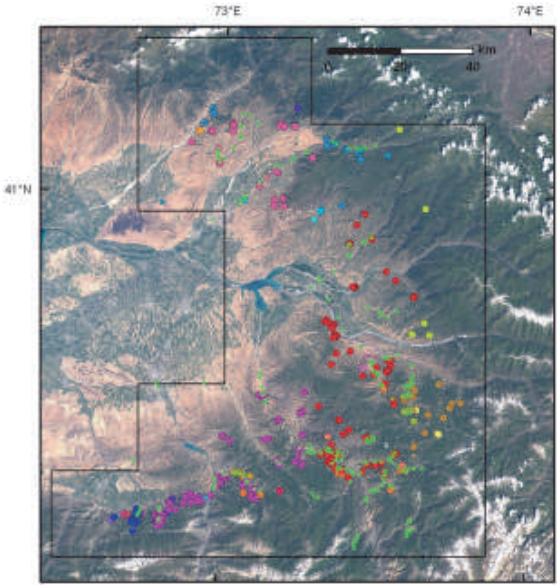


Figure 8: Automatically detected landslides between 2009 and 2012 (different colored circles: landslide of different movement date; stars: GPS points of the field survey in September 2012)

5 Discussion and Outlook

The developed approach using multi-temporal RapidEye data allows automated detection of landslides in an object-based form enabling the systematic derivation of a multi-temporal landslide inventory for large areas required for objective hazard and risk assessment. Although compared to previous years process activity had been rather low during the analyzed time period between 2009 and 2012 and no specific triggering event has been known, a total of about 250 events of landslides could be automatically detected. This shows the constant ongoing process activity occurring independently from distinct triggering events in this area. This situation emphasizes the need for multi-temporal landslide inventories in regions dominated by complex slope failures, such as the Eastern rim of the Fergana Basin. In a next step the developed approach will be applied to a longer-term time series containing multi-temporal image data acquired by a variety of optical sensor starting from 1986 whereas annual coverage for the complete region of interest is available from 1996 until today.

Despite the to a large degree correct automated detection of slope failures, in some cases visual evaluation of the obtained results has revealed false positives. They are mostly caused by remaining geometric mismatches between the multi-temporal image data due to non-systematic geometric errors resulting from the original orthorectification process that can not be corrected using the shift-based co-registration procedure. In some cases false positives are caused because the temporal vegetation change of other land covers are similar to those related to landslides. However, these cases represent a small minority within all of the detected objects. Future work will investigate the potential of including additional spatial data, such as DEM information in the analytical process in order to further improve the reliability of automated detection of landslides. The resulting automated detection system will

be a useful tool for deriving spatially differentiated quantitative information about process activity at a regional scale.

Acknowledgements

This work was funded by the German Federal Ministry of Research and Technology (BMBF) within the framework of PROGRESS (Potsdam Research Cluster for Georisk Analysis, Environmental Change and Sustainability). We also thank the German Aerospace Agency (DLR) for providing RapidEye data by the RESA (RapidEye Science Archive) program (grant number 424).

References

BEHLING, R., ROESSNER, S., SEGL, K., ROGASS, C., WETZEL, H. U. & KAUFMANN, H. 2012: Automatische geometrische Koregistrierung multitemporaler Satellitendaten zur Inventarisierung von Hangrutschungen. In: Borg, Daedelow, und Johnson. (Eds.), Vom Algorithmus zum Produkt - 4. RESA Workshop: 35-50.

COPPIN, P., JONCKHEERE, I., NACKAERTS, K., MUYS, B. & LAMBIN, E., 2004: Digital change detection methods in ecosystem monitoring: a review. International Journal of Remote Sensing, 25(9): 1565-1596.

FELL, R., COROMINAS, J., BONNARD, C., CASCINI, L., LEROI, E. & SAVAGE, W., 2008: Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. Engineering Geology, 102(3-4): 85-98.

GUZZETTI, F., MONDINI, A. C., CARDINALI, M., FIORUCCI, F., SANTANGELO, M. & CHANG, K., 2012: Landslide inventory maps: New tools for an old problem. Earth-Science Reviews 112 (1–2): 42–66.

IBATULIN, C. V., 2011: Monitoring of landslides in Kyrgyzstan. Ministry of Emergency Situations of the Kyrgyz Republic, Bishkek: 145 pp. (in Russian).

Jianya, G., Haigang, S., Guorui, M. & Qiming, Z., 2008: A review of multi-temporal remote sensing data change detection algorithms. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37 (B7): 757–762.

LU, D., MORAN, E., HETRICK, S. & LI, G., 2010: Land-Use and Land-Cover Change Detection. Advances in Environmental Remote Sensing: Sensors, Algorithms, and Applications, 7: 273.

ROESSNER, S., WETZEL, H.-U., KAUFMANN, H. & SARNAGOEV, A., 2005: Potential of satellite remote sensing and GIS for landslide hazard assessment in Southern Kyrgyzstan (Central Asia). Natural Hazards, 35(3): 395–416.

VAN WESTEN, C.J., CASTELLANOS, E. & KURIAKOSE, S.L., 2008: Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. Engineering Geology, 102(3-4), S.112-131.

Schlussfolgerungen aus der Anwendung der Geo-Fernerkundung in geowissenschaftlichen Projekten in SW-Deutschland

Hans Joachim Franzke[1] & Hans-Ulrich Wetzel[2]

Zusammenfassung: Es werden Beispiele von fernerkundlichen Arbeiten aus dem südwestdeutschen Raum beschrieben, die als Zuarbeiten zu ingenieurgeologischen, lagerstättenkundlichen und seismotektonischen Projekten angelegt waren. Es wird hervorgehoben, dass geowissenschaftliche Erkenntnisgewinne vor allem durch die enge Verknüpfung der Fernerkundung mit geologischen Arbeiten gewonnen werden können. Hierbei ist ein stufenweiser Ablauf, beginnend mit Übersichtsbearbeitungen und anschließender Detailbearbeitung sowie die ständige Rückkopplung zur „Bodengeologie" erforderlich.

1 Aufgabe

Die Autoren waren wiederholt in Projekte eingebunden, welche sich mit Grundlagen der Lagerstättenbildung, Regionaltektonik und speziellen Problemen der Bruchtektonik (Neotektonik, Seismotektonik) auseinander zu setzen hatten (FRANZKE et al, 2011 und 2012, in dieser Zeitschrift). Dies erlaubt einige generalisierende Schlussfolgerungen zu den Möglichkeiten und gegenwärtigen Grenzen in der Anwendung der Fernerkundung zur Lösung einiger komplexer geowissenschaftlicher Fragestellungen. In Übereinstimmung mit Ergebnissen aus anderen Gebieten hat sich bestätigt, dass die Fernerkundung nur begrenzt unabhängig vom übrigen geowissenschaftlichen Methodenspektrum erfolgreich agieren kann. Schlüssige Ergebnisse kommen vor allem dann zustande, wenn die Fernerkundung mit weiteren geowissenschaftlichen Untersuchungen im interdisziplinären Teamwork sinnvoll verknüpft wird. Ohne diese Einbindung hängen die Resultate meist buchstäblich „in der Luft" und können nicht ausreichend in die Geologie eingebunden werden. Einsatzmöglichkeiten der Fernerkundung für die Bewertung von großen Bruchzonen der Erdkruste bezüglich ihres seismotektonischen Potentials, ihrer Bedeutung für die Fluid-Migration bei Mineralisationsprozessen sowie bei der Analyse der tektonisch-historischen Entwicklung im südwestdeutschen Raum sollen nachfolgend an einigen Beispielen präsentiert und in ihrem Aussagewert beurteilt werden.

2 Fernerkundliche Vorbereitung

Für jedes geowissenschaftliche Projekt ist es hilfreich (1. Schritt), eine aufgabenbezogene regionale Analyse in Form von Übersichtsbearbeitungen an Standard-Materialien (TM, ERS1-2, ASTER, SPOT, JERS) durchzuführen. Dies gewährleistet es, dass das Gebiet nicht von vornherein „durch das Schlüsselloch" betrachtet wird und die stets notwendige regionale Einbindung des Arbeitsbereichs gewährleistet wird.

[1] Hans Joachim Franzke, Hirschler Weg 11, D-38678 Clausthal-Zellerfeld, Tel.: 05323-82140, E-Mail: franzke.clz@gmail.com
[2] Dr. Hans-Ulrich Wetzel, Deutsches GeoForschungsZentrum Potsdam, Telegrafenberg, D-14473 Potsdam, Tel.: 0331-2881194, E-Mail: wetz@gfz-potsdam.de

In Abb. 1 ist eine Übersichtskarte zur regionalen Bruchtektonik SW-Deutschlands gezeigt, welche von den Autoren aus TM- und ERS-1-2-Daten generiert wurde. Diese Karte zeigt praktisch alle Bruchlinien, die auch in der „Geologischen Karte der Bundesrepublik Deutschland 1 : 1Mio zu sehen sind, d. h. die fernerkundliche Kartierung bildet die regional bekannten Bruchzonen nahezu komplett ab und geht in einigen Bereichen darüber hinaus. Dieses Beispiel kann auch als Testfall bzw. als Vergleichsobjekt dafür dienen, welches Aussageniveau beim Einsatz geeigneter Fernerkundungsdaten durch die Bearbeitung erfahrener und mit einem Gebiet vertrauter Geowissenschaftler in ähnlichen Gebieten des humiden Klimabereichs des mittleren und nördlichen Europa erreicht werden kann.
Diese Übersicht bildete einen geeigneten Rahmen für weiterführende Detailarbeiten zu ingenieurgeologischen, lagerstättenkundlichen und allgemein-geologischen Fragestellungen.

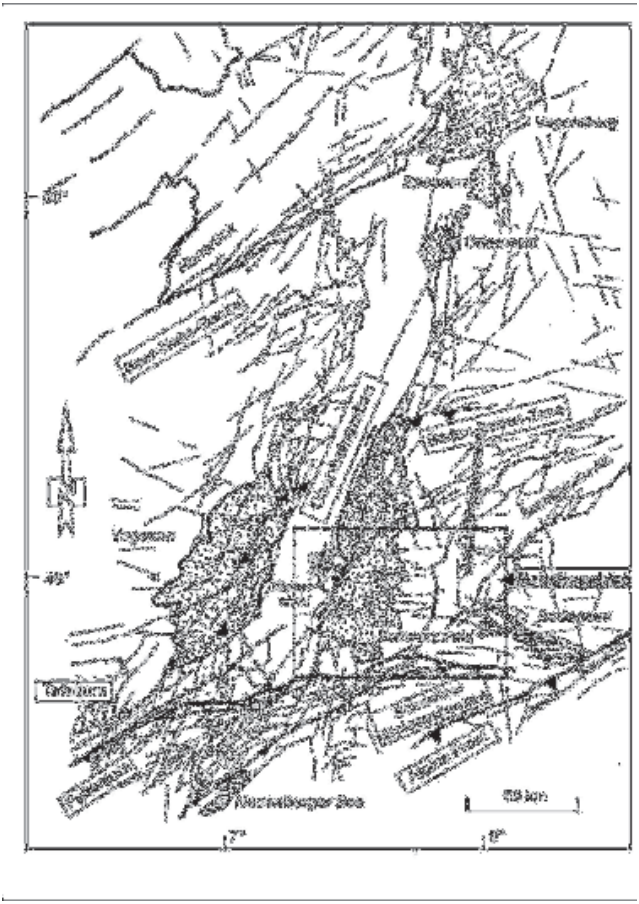


Abb. 1: Bruchkarte des südwestdeutsch-nordschweizerischen Raums, kompiliert aus SRTM-Daten. LLZ - Lalaye-Baden-Baden-Zone (Variszische Scherzone).

3 Strukturgebundene Mineralisationsprozesse

Im DFG-Schwerpunkt-Projekt „Intraformationale Lagerstättenbildung“ in der Mitte der 90er Jahre ging es u. a. um tektonische Kontrolle weiträumiger Fluidmigrationen und deren Mineralkonzentration in Bruchzonen des mitteleuropäischen Bruchschollenfeldes. Das weiträumig aufgeschlossene Grundgebirge des Schwarzwalds diente dabei als Schlüsselgebiet (Abb. 2). Da Bruchzonen bei der traditionellen geologischen Kartierung zumeist nur abschnittsweise erfasst werden können, stand die Aufgabe, die Lücken im Bruchgitter zu schließen und durch die Kartierung zusätzlicher Strukturen das Bruchgitter zu komplettieren, um die strukturgeologische Situation im Sinne der obigen Fragestellung fundierter bewerten zu können.

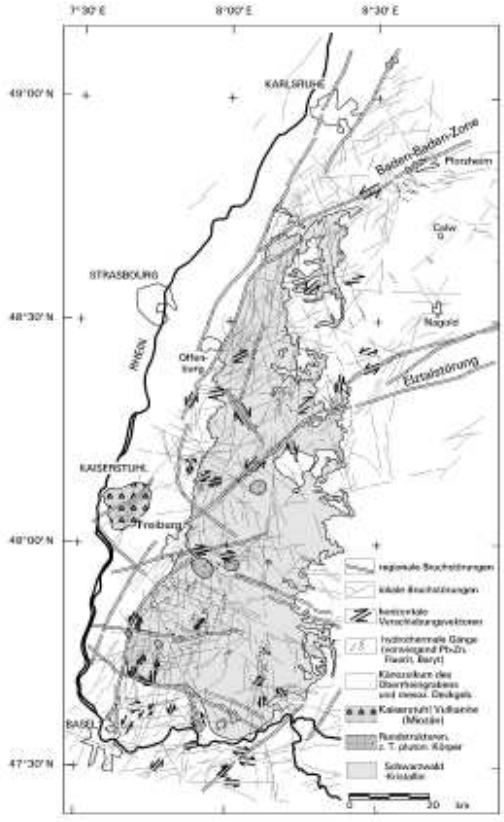


Abb. 2: Interpretationskarte, erstellt aus einem Satz von Satelliten-Daten (TM, SPOT, JERS, ERS-2) vom Gebiet des Schwarzwalds. Die großen Bruchzonen sind hervorgehoben. Zusätzlich zu den bekannten Strukturen treten

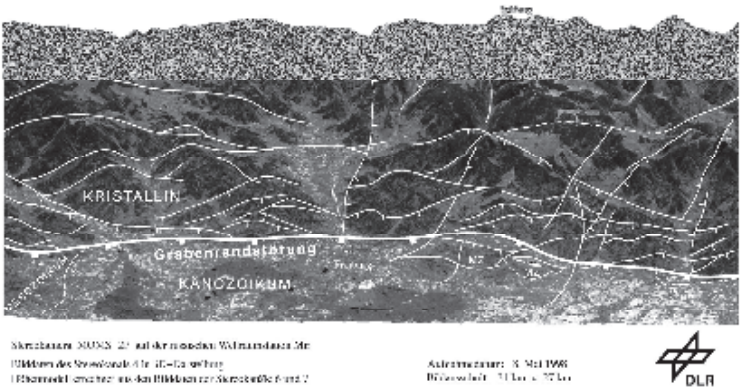


Abb. 3: Randzone des Oberrheingrabens (Blick nach Osten) in der unmittelbaren Umgebung von Freiburg. MZ – Mesozoikum. Das Höhenmodell verdanken wir der DLR. Das Aufgabeln der Randstörung nach Nord (links) beginnt am linken Bildrand.

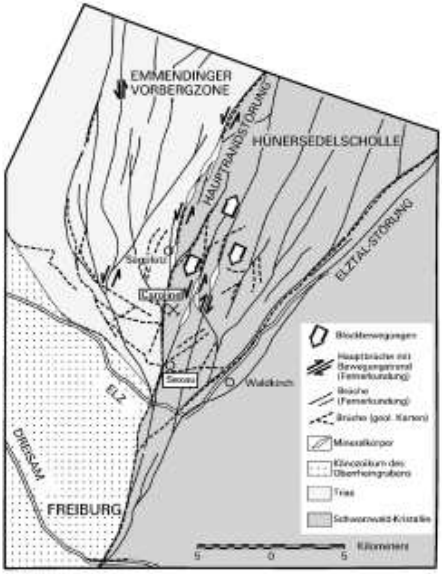


Abb. 4: Aufsplitterung der Randzone des Oberrheingrabens N von Freiburg i. Br. Erläuterung im Text.

Die Ergebnisse zeigten, dass sich die Mineralisationen nicht nur unmittelbar in den Störungszonen abgeschieden haben, sondern dass die großen steilen Frakturzonen des Grundgebirges des Schwarzwaldes auch Zirkulationszellen hydrothermaler Lösungen eingrenzen und damit die Fixierung der Mineralisationen in bestimmten Blocksegmenten favorisieren.

Fernerkundliche Detailstudien konnten in Verbindung mit bodengestützten Strukturanalysen die phasenhafte Verlängerung von Brüchen und deren nachfolgende Mineralisationen anhand bestimmter Mineralparagenesen nachweisen. Am Ostrand des Oberheingrabens nördlich von Freiburg i. Br. konnte das Aufspleißen der bei Freiburg noch relativ eng gebündelten Grabenrandzone (Abb. 3) nach Norden dokumentiert werden (Abb. 4), was im Gebiet des Freiamt Sexau N von Freiburg zum Stau von hydrothermalen Lösungen in einer Migrationsfalle mit intensiver Mineralausscheidung führte. Dies wird als wesentliche Ursache für die Lokalisation des dortigen Mineralgangfeldes (Polysulfide, Karbonate, Baryt-Fluorit) angesehen (FRANZKE et al. (2003).

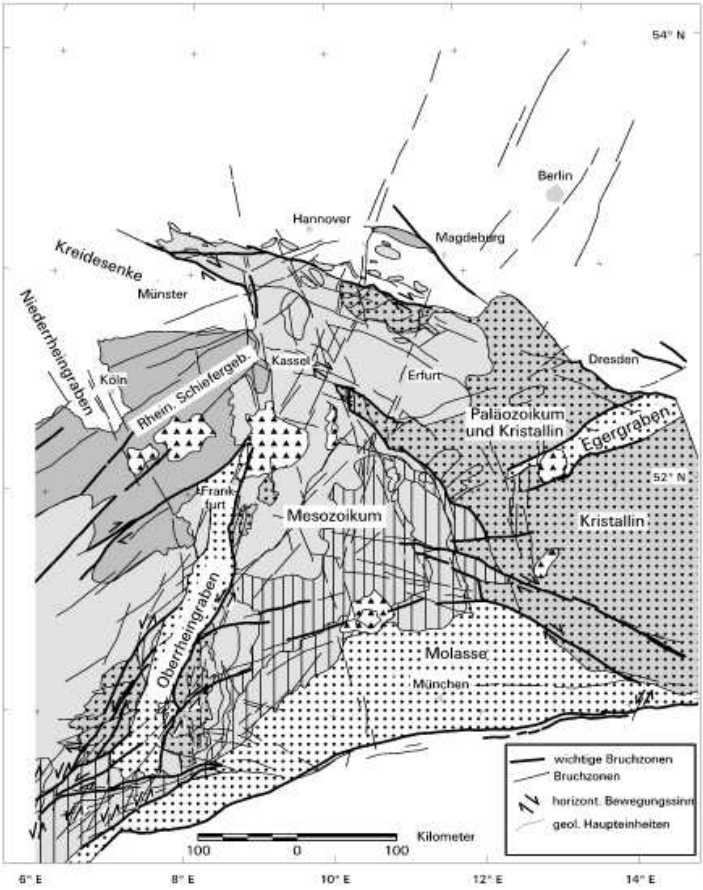


Abb. 5: Geologische Umsetzung der Interpretation von Radardaten (ERS-1 und 2) in eine Bruchkarte von Deutschland (vereinfachte Version). WETZEL & FRANZKE (2001).

4 Neotektonik-Seismotektonik

Eine weitere Anwendung der Fernerkundung für angewandte Fragestellungen im südwestdeutschen Raum ergab sich mit der Zuarbeit zur seismotektonischen Rayonierung Mitteleuropas im Rahmen des deutschen Forschungsprojekts zur Vermeidung von Naturkatastrophen (Projekt DFNK). Die dabei auf der Basis von Satellitendaten erstellte Bruchkarte (Abb. 5) bildete die Basis der seismologischen Zonierung Mitteleuropas von GRÜNTHAL et al. (2009) als ein Ergebnis dieses Projekts.
Mittels der Kombinierung von Fernerkundung mit bodengestützten strukturgeologischen Arbeiten konnte die Verknüpfung der heutigen Seismizität im Bereich des 9°-Ost-Lineaments (N-S verlaufende Albstadt-Scherzone zwischen Stuttgart und dem Bodensee) mit querenden Störungszonen (Hohenzollerngraben) untermauert werden (Abb. 6).

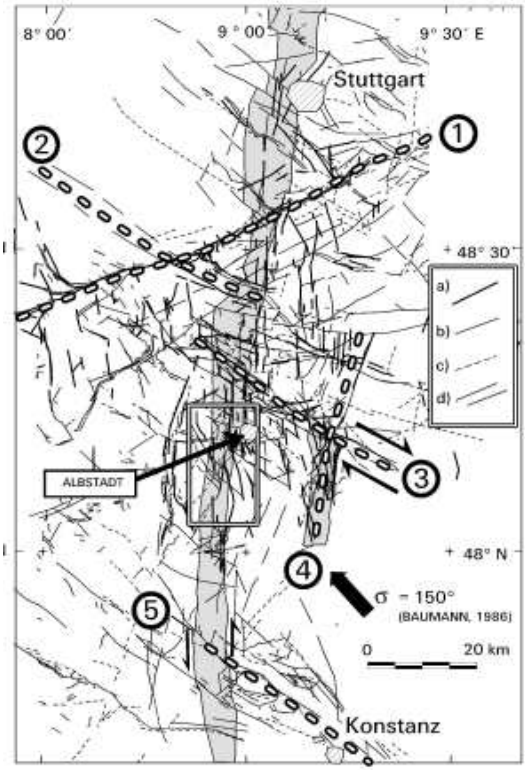


Abb. 6: Struktur-Interpretation des zentralen Bereichs des 9°-Ost-Lineaments. Indikationen im Störungsgitter (Felduntersuchungen) belegen die Reaktion von Teilen des Störungsgitters auf die Einspannung durch das heutige Stressfeld (dicker Pfeil = Schubrichtung um 145-150°). Eingekreiste Zahlen: 1- Schwäbisches Lineament, 2- NW-Se- Störungen bei Rottenburg,, 3-Hohenzollern-Grabenzone 4-Laucha-Zone (N-S), 5- Albstadt-Scherzone (N-S). a)- große Bruchstörungen, b)-kleinere Bruchstörungen, c)-vermutete Brüche, d)- von Brüchen eingegrenzte Zonen, z. B. Grabenbrüche oder Störungsgruppe.

Literatur:

FRANZKE, H. J., WERNER, W. & WETZEL, H.-U. (2003): Die Anwendung von Satellitenbilddaten zur tektonischen Analyse des Schwarzwalds und des angrenzenden Oberrheingrabens.- Jh. Landesamt f. Geologie, Rohstoffe und Bergbau Baden-Württemberg, 39: 25-54; Freiburg i. Br.

FRANZKE, H.J.; WETZEL, H.-U. & MADRITSCH, H. (2011): Verknüpfung von strukturgeologischer Fernerkundung mit Geländeaufnahmen im Grenzbereich zwischen Oberrheingraben, Südschwarzwald und Faltenjura. - In: Seyfert, E. (Eds.), Geodaten - eine Ressource des 21. Jahrhunderts : Vorträge ; 29. [i.e. 31.] Wissenschaftlich-Technische Jahrestagung der DGPF, 13. - 15. April in Mainz, 271-278.| EDOC: 17729 |

FRANZKE, J. H.; WETZEL, H.-U. & KNIEPER, M. (2012): Einsatzmöglichkeiten der Fernerkundung für struktur- und ingenieurgeologische Fragestellungen in gering aufgeschlossenen Kristallingebieten (Süd-Schwarzwald).. Erdblicke - Perspektiven für die Geowissenschaften, 32. Wissenschaftlich-Technische Jahrestagung der DGPF (Potsdam 2012), 217-224.| EDOC: 18630 |

FRANZKE, H-J; WETZEL, H.-U. & WERNER, W. (2000): Das Strukturmuster des Schwarzwaldes aus der Sicht bruchtektonisch interpretierter TM- und ERS-1-Aufnahmen. Publikationen der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung, 7, 489-496. EDOC: 1369 |

GRÜNTHAL, G., BOSSE, C. & STROMEYER, D. (2009) Die neue Generation der probabilistischen seismischen Gefährdungseinschätzung der Bundesrepublik Deutschland: Version 2007 mit Anwendung für die Erdbeben-Lastfälle der DIN 19700:2004-07 „Stauanlagen“. Scientific Technical Report STR 09/07, Deutsches GeoForschungsZentrum, Potsdam, 81 pp.

FRANZKE, H-J & WETZEL, H.-U. (2001): Geologische Interpretation eines ERS-1 Radarmosaiks von Deutschland. Publikationen der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung, 10, 503-510.| EDOC: 2523 |

FRANZKE, H-J; WETZEL, H.-U. & WERNER, W. (2001): Welche Informationen können uns Fernerkundungsdaten zu Paläo-Stressfeldern und zu Mineralisationsprozessen im Schwarzwald geben?. Publikationen der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung (DGPF), 9, ISSN 0942-2870, 340-353. | EDOC: 3446 |

Structural analysis of Martian impact craters using several remote sensing data

GERWIN WULF[1*], MICHAEL POELCHAU[1] & THOMAS KENKMANN[1]

Die Bildung von Impaktkratern ist ein hochdynamischer und komplexer geologischer Vorgang, der einer der fundamentalsten Prozesse im Sonnensystem darstellt. Aufgrund von Verwitterung und Erosion ist die strukturgeologische Untersuchung terrestrischer Einschlagskrater oft nur eingeschränkt möglich. Die hohe Qualität, Auflösung und Verfügbarkeit von Fernerkundungsdaten des Mars ermöglicht ein nie zuvor da gewesenes Maß an Präzision im Hinblick auf geologische Untersuchungen. In dieser Arbeit präsentieren wir detaillierte strukturgeologische Analysen eines Marskraters auf der Basis von hochauflösenden Fernerkundungsdaten. Die Ergebnisse zeigen, dass die interne Struktur eines Impaktkraters Hinweise auf die Einschlagsrichtung zulässt.

1 Introduction

The majority of impacts on planetary bodies occur at an oblique impact angle to the target surface, where the incidence angle follows a Gaussian probability distribution with a mean value of 45° (GILBERT 1893; SHOEMAKER 1962). Although oblique impacts are prevalent occurrences, the crater shape remains circular for impact angles above 10-15° from the target surface (GAULT & WEDEKIND 1978; BOTTKE et al. 2000) and thus can rarely give implications for the impact direction or angle. The distribution of the ejecta blanket, on the other hand, is a distinctive indicator for oblique impacts. It loses its radial symmetry at angles below 45-35° and at lower angles forms "forbidden" zones and "butterfly" patterns (GAULT & WEDEKIND 1978, HERRICK & HESSEN 2006). Unfortunately, on Earth most ejecta blankets are eroded and cannot be used to determine the impact direction for terrestrial craters. Hence, other potential indicators for an oblique impact are needed if the impact trajectory is to be determined in terrestrial craters.
The internal structure of central uplifts has also been interpreted to be influenced by an oblique impact angle. The subsurface structures of eroded central uplifts of some terrestrial impact craters show a preferential, non-radial orientation of folding, faulting and stacking of layered (and originally subhorizontal) bedrock that implicates a preferred transport direction during the crater formation process, as seen in Upheaval Dome (Utah, USA) (KENKMANN et al. 2005), Spider (Australia) (SHOEMAKER & SHOEMAKER 1996), Gosses Bluff (Australia) (SCHERLER et al. 2006), Matt Wilson (Australia) (KENKMANN & POELCHAU 2009) and Jebel Waqf as Suwwan (Jordan) (KENKMANN et al. 2010). Kenkmann and Poelchau (2009) suggest that this imbrication is caused by remnant horizontal momentum transferred from the impacting projectile to the target during an oblique impact.

1) Albert-Ludwigs-Universität Freiburg, Institut für Geowissenschaften - Geologie, Albertstr. 23b, 79104 Freiburg, *E-Mail: gerwin.wulf@geologie.uni-freiburg.de

To date, the observed structural trends could be directly connected to an impact direction in the case of Matt Wilson crater due to its elliptical crater shape (KENKMANN & POELCHAU 2009), with upturned bedding striking perpendicular to the long axis of the crater ellipse, which is parallel to the impact trajectory. In order to verify the connection between an oblique impact trajectory and the observed preferential, non-radial deformation in central uplifts, further "ground truth" examples are needed in which an independent indicator of the impact direction, i.e., the asymmetric ejecta blanket, is preserved.
Martian impact craters now offer an ideal opportunity to verify both terrestrial observations and numerical models of the central uplift. The increasing quality, resolution and availability of remote sensing data of the Martian surface, in particular High Resolution Imaging Science Experiment (HiRISE) images (up to 25 cm/pixel, (MCEWEN et al. 2007)), permit the application of structural geology methods to Martian craters with a high measure of precision.
In this work we present mapping results and structural analyses of the central uplift of a Martian impact crater based on HiRISE imagery. The methods used for structural mapping and the implications the results have for oblique impact craters are discussed below.

2 Methods

The Mars Global GIS dataset from the NASA Planetary Data System (PDS) was downloaded and implemented into ESRI ArcGIS to find relevant craters and to compile our crater catalog. In the process, MOLA (Mars Orbiter Laser Altimeter) shaded relief maps, THEMIS (Thermal Emission Imaging System) infrared, and Viking mosaics were used to identify and collect suitable craters with oblique ejecta patterns. HiRISE images of the central uplifts were screened for the amount of layered bedrock that can be mapped, using the IAS viewer, a JavaScript-based software tool provided on the HiRISE website.
Arima crater (15.84°S 296.33°E) on Mars met all three criteria for this project, i.e. an oblique ejecta pattern, exposed layered bedrock in the central uplift and HiRISE coverage. All available image data of the crater, including THEMIS, HRSC (High Resolution Stereo Camera), CTX (Context Camera) and HiRISE, were processed by using USGS's Integrated System for Imagers and Spectrometers (ISIS) to obtain an overview for subsequent mapping in ArcGIS. Strike and dip were inferred from HiRISE images in a qualitative fashion based on the morphological expression of outcropping layers in plan view (Fig. 1). For example, highly inclined layers are recognized due to the strong linear trend of outcropping, regardless of local changes in topography. Less steeply inclined beds have a visible curvature that is a result of the intersection of the bedding plane with the local topography; this curvature becomes more pronounced with smaller dip angles. Bedding strike is more easily and precisely determined for steeply inclined layers. The influence of regional morphological trends (hillsides, ridges, etc.) on apparent strike and dip was taken into account when possible. For this purpose, DTMs and derived contour maps with the highest possible resolution (upon availability MOLA, HRSC or HiRISE) were used in addition to the panchromatic HiRISE image data (Fig. 1). Resolution of the MOLA DTMs was 463 m/px, HRSC DTMs had 150-175 m/px and HiRISE DTMs had 1m/px resolution.

Strike values are estimated to have an error <10° for beds inclined at an angle of 75° or steeper, while the error for flatter beds is expected to gradually increase. Strike of subhorizontal bedding (<15-25° dip) was generally not measured due to the high degree of uncertainty.

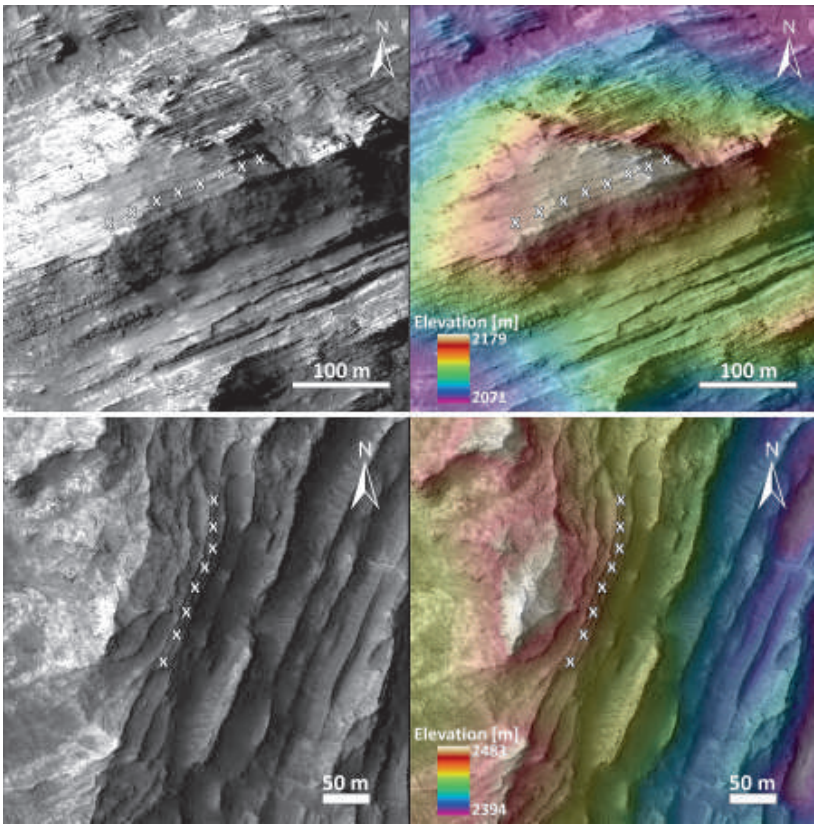


Fig. 1: Two examples of layered bedrock in HiRISE images that were used to map structural trends (upper images from Martin crater (DT1EA_002353_1585_003408_1585), lower images from Arima crater (DT2EA_005201_1640_005913)). Superposed elevation values on the right are from a DTM generated from a HiRISE image pair (top DTEEC_002353_1585_003408_1585, bottom DTEEC_005201_1640_005913). Bedrock in the top image is steeply dipping to vertical and strikes WSW-ENE. The trace of bedrocks is straight and largely unaffected by the topography. Bedrock in the lower image is moderately inclined (~30°), dips to the WNW and strikes NNE-SSW. Bedding outcrops show curvature related to the topography.

Strike values were then plotted in a rose diagram with a 10° step size. The average value for all strike data in the central uplift was then calculated as the vector mean of two dimensional orientation data. Care was taken that strike measurements did not suffer from lighting bias. The images in Fig. 1 additionally emphasize that the solar position is irrelevant for structural measurements. For example, strike of bedrock in the top image only deviates from the sub-solar azimuth by ~15° (i.e., lighting is nearly parallel to the strike, suggesting poor visibility), and yet the layering is easily determined.

We are therefore certain that lighting had no effect on a 1 m/px DTM, which was pre-processed and released to the PDS by the HiRISE team (MATTSON et al., 2011), covered the western half of Arima crater and was incorporated into the ArcGIS map. Strike and dip values were geometrically calculated using the Layertools extension for ArcGIS (KNEISSL et al. 2010). Care was taken by selecting only planar layers that did not appear to be faulted or folded. These quantitative data were then compared to strike values that were visually estimated in the same region with the method described above. The quantitative measurements and qualitative estimations are in good agreement with each other; the resulting trends are the same within an error of ±12%. The three-dimensional method had the advantage of enabling a much more precise measurement of dip angles, thus these strike and dip values could be plotted in a stereographic projection.

The orientation of faults in the central uplift was also mapped based on morphological and structural criteria, using for example abrupt discontinuations or offsets of structural trends of layered bedrock, or the contact of layered units oriented at oblique angles to each other. The orientation of these fault lines was also plotted in rose diagrams.

Finally, the impact trajectory azimuth was estimated from the craters ejecta blanket, based on the extent of the non-radial distribution of ejecta, the blanket's bilateral symmetry, and the position of the uprange forbidden zone.

3 Results

Arima crater is a central pit crater located at 15.84°S 296.33°E, directly south of Valles Marineris in Thaumasia Planum. The single layer radial (SLERd) crater is 55.3 km in diameter and contains a ~11 km diameter central pit (Fig. 2a) (according to the revised Catalog of Large Martian Impact Craters; BARLOW 2006). THEMIS nighttime IR images best illustrate a distinctly asymmetrical ejecta blanket with a "forbidden" zone in the W sector of the crater (Fig. 2b). The ejecta pattern has been slightly thinned by erosion, but the bulk of the original extent and distribution is still identifiable using and combining THEMIS IR, THEMIS VIS, HRSC and CTX image data. The crater can therefore clearly be characterized as an oblique impact crater whose projectile came from WNW (azimuthal angle of ~290°). About 90% of the rimmed floor pit and its surroundings are covered by HiRISE imagery. Additionally, a 1 m/px HiRISE DTM is available for the western part of the structure (Fig. 3a). The estimated apparent strike values show a preferred strike trend of ~30° over the majority of the central pit (Figs. 2c, d). The data density decreases in the easternmost part of the central pit due to the absence of exposed layered bedrock.

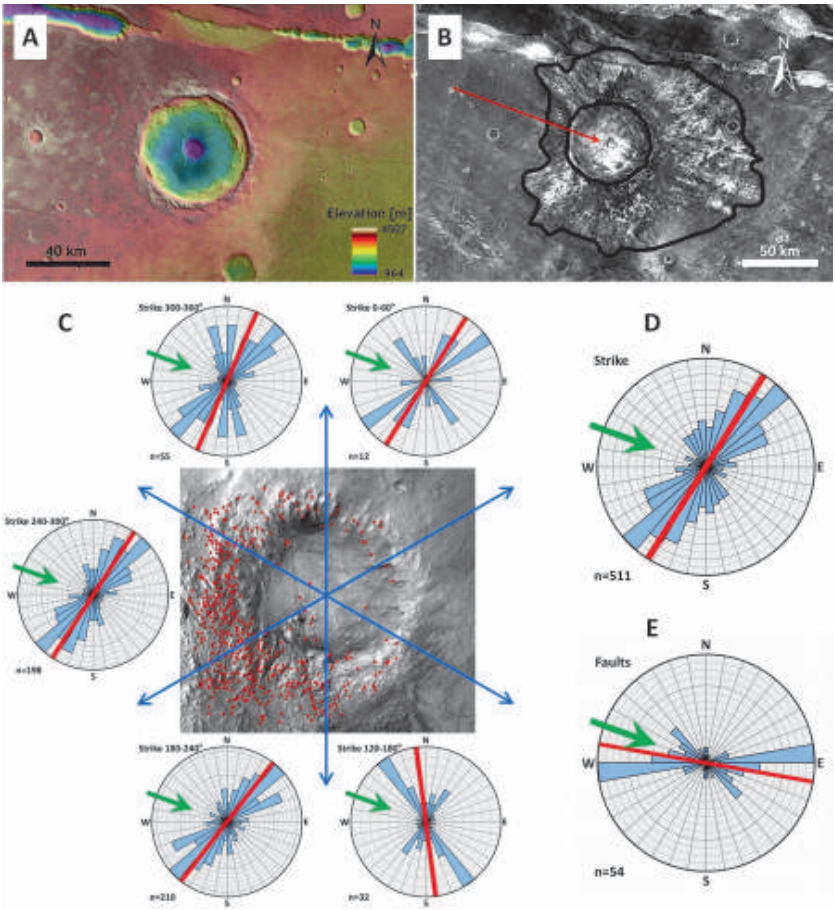


Fig. 2: (A) Topographical overview of Arima crater (15.84°S 296.33°E) in Thaumasia Planum south of Valles Marineris (HRSC DTM h0482_0000_dt4 over HRSC nadir image h0482_0000_nd4). (B) The THEMIS nighttime IR data show an asymmetric ejecta pattern with a "forbidden zone" in the W sector indicating an oblique impact from WNW (290°). (C) Strike trends of layered bedrock measured in HiRISE imagery (PSP_005913_1640, PSP_007047_1640) show little variation throughout the central pit region of the crater and are nearly perpendicular to the assumed impact direction from WNW (green arrows in rose diagrams). (D) Apparent strike values are perpendicular to the impact direction (green arrow), while faults (E) trend roughly parallel to the impact direction.

The HiRISE DTM of the western part of the central pit was used to measure true strike and dip values using the Layertools extension for ArcGIS. The measured true strike orientations confirm the preferred trend of ~30-40° of the apparent strike values over the majority of the western part of the central pit. The strata predominantly dip WNW with shallow dip angles of ~25° (Fig. 3a). The measured dip directions give evidence of a significant symmetry axis that trends at 290° and is dominant over the entire central pit (Figs. 3a, b). The orientation of the main faults also shows a preferred WNW-ESE trend (Fig. 2e).

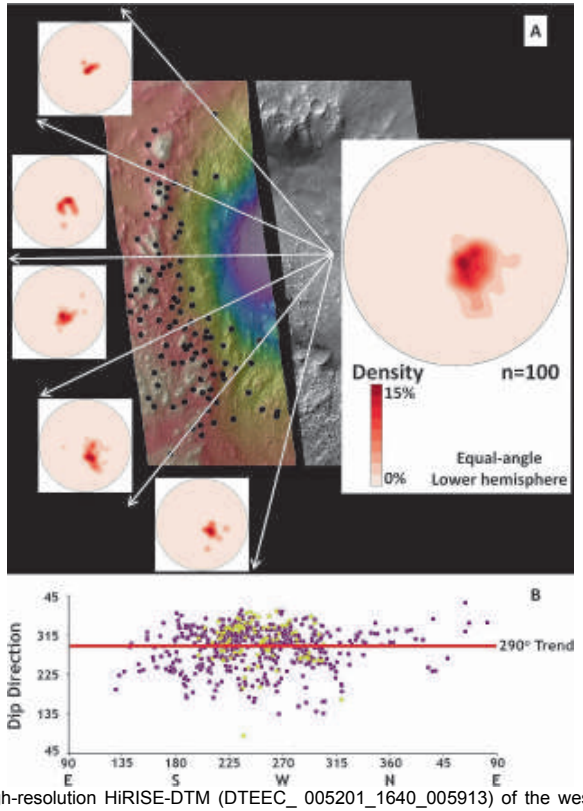


Fig. 3: (A) A high-resolution HiRISE-DTM (DTEEC_ 005201_1640_005913) of the western part of the structure allows the measurement of "true" strike and dip values, which all dip to the NW at ~25°, as seen in stereographic projections of the measured surface poles. (B) The two results for bedding orientation derived in 2C (purple dots) and 3A (yellow dots) are in good agreement with each other when the dip direction of measured beds is plotted against the azimuthal position relative to the crater center. 3D DTM based strike measurements thus confirm the validity of strike values derived from 2D single HiRISE images in this study.

4 Discussion and Conclusions

The flat surface of Thaumasia Planum strongly indicates that the layered bedrock involved in the cratering processes was originally in a horizontal to sub-horizontal position. Most layers must have experienced a strong rotation from a horizontal position during the central uplift formation and are therefore ideal tracers for structural deformation. Formation of central uplifts during the modification stage of impact cratering is the result of the inward and upward movement of rock towards the crater center and may result, at least under ideal conditions, in either concentric strike of the tilted bedrock in more distal parts of the central uplift or in radial strike in the central portion of the uplift due to constrictional faulting and folding (KENKMANN 2002). The analyzed central uplift of Arima crater deviates considerably from this ideal structure of radially symmetric behavior showing a preferred orientation of layered bedrock. The dominant strike orientation is nearly perpendicular to the impact direction inferred from the asymmetric ejecta blanket of this crater. Additionally, the strata show a preferred orientation of faults parallel to the impact trajectory. Our study documents that the uplifted strata in the surrounding of the central pit of Arima crater are not affected in strike or dip by the crater pit, nor is a concentric downward faulting or downward flexure of strata observable. This study demonstrates that crater pit formation must postdate crater modification and formation of the central uplift.

Arima crater on Mars was selected to study the effects of oblique impacts on the subsurface structure of this crater. The impact trajectory was inferred from the asymmetric ejecta blanket. Full HiRISE coverage of the craters interior parts allowed a detailed structural analysis of strata and fault orientations. The impact trajectory inferred from the asymmetric ejecta blanket and the orientation of structural elements show a close correlation:

(i) the strike of bedrock in the central part of the crater floor is predominantly perpendicular to the impact trajectory,

(ii) the majority of faults trend parallel to the impact trajectory.

The presented structural features can thus be used to unravel impact trajectories. Criterion (i) is the dominant feature indicative for impact trajectories. Criteria (ii) should not be used independently but is likely to be associated with (i). The formation of preferred orientations of strata strike in central uplifts is at least as sensitive as the asymmetry of ejecta blankets. These findings are in agreement with observations from terrestrial and other Martian impact craters (WULF et al. 2012)

5 Acknowledgements

Financial support for this study came from the German Research Foundation (DFG), grant KE 732/11-1 and -2.

6 References

BARLOW, N. G., 2006: Status report on the ‘‘Catalog of Large Martian Impact Craters’’, version 2.0. Lunar Planet. Sci. XXXVII, 1796 (abstract).

BOTTKE, W. F., LOVE, S.G., TYTELL, D., GLOTCH, T., 2000: Interpreting the elliptical crater populations on Mars, Venus, and the Moon. Icarus 145, 108-121.

GAULT, D. E. & WEDEKIND, J. A., 1978: Experimental studies of oblique impact. Proc. Lunar Sci. Conf. IX, 3843-3875.

GILBERT, G. K., 1893: The Moon’s face, A study of the origin of its features. Bulletin of the Philosophical Society of London 12, 241-292.

HERRICK, R. R. & HESSEN, K. K., 2006: The planforms of low-angle impact craters in the northern hemisphere of Mars. Meteor. Planet. Sci. 41, 1483-1495.

KENKMANN, T., 2002: Folding within seconds. Geology 30, 231-234.

KENKMANN, T., JAHN, A., SCHERLER, D., IVANOV, B. A., 2005: Structure and formation of a central uplift: A case study at the Upheaval Dome impact crater, Utah. In: Kenkmann, T., Hörz, F., Deutsch, A. (Eds.), Large meteorite impacts III, Boulder, Colorado, GSA Special Paper 384, pp. 85-115.

KENKMANN, T. & POELCHAU, M. H., 2009: Low-angle collision with Earth: The elliptical impact crater Matt Wilson, NT, Australia. Geology 37, 459-462.

KENKMANN, T., REIMOLD, W.U., KHIRFAN, M., SALAMEH, E., KHOURY, H., KONSUL, K., 2010: The complex impact crater Jebel Waqf as Suwwan in Jordan: Effects of target heterogeneity and impact obliquity on central uplift formation. In: Gibson, R.L., Reimold, W.U. (Eds.), Large meteorite impacts and planetary evolution IV, Boulder, Colorado, GSA Special Paper 465, pp. 471-487.

KNEISSL, T., VAN GASSELT, S. & NEUKUM, G., 2010: Measurement of strike and dip of geologic layers from remote sensing data - new software tool for ArcGIS. Lunar Planet. Sci. XLI, 1640 (abstract).

MATTSON, S., et al., 2011. Release of HiRISE Digital Terrain Models to the Planetary Data System. Lunar Planet. Sci. XLII, 1608 (abstract).

MCEWEN, A.S., et al., 2007. Mars Reconnaissance Orbiter’s High Resolution Imaging Science Experiment (HiRISE). J. Geophys. Res. 112, E05S02.

SCHERLER, D., KENKMANN, T., JAHN, A., 2006: Structural record of an oblique impact. Earth and Planetary Science Letters 248, 28-38.

SHOEMAKER, E. M., 1962: Interpretation of lunar craters. In: Kopal, Z. (Eds.), Physics and astronomy of the Moon, Academic Press, New York, pp. 283-359.

SHOEMAKER, E. M. & SHOEMAKER, C. S., 1996: The proterozoic impact record of Australia, AGSO J. Aust. Geol. Geophys. 16, 379-398.

WULF, G., POELCHAU, M. H. & KENKMANN, T., 2012: Structural Asymmetry in Martian Impact Craters as an Indicator for an Impact Trajectory. Icarus 220, 194-204.

An Approach for Automated Registration of Hyperspectral Images for Boresight Calibration

UWE KNAUER[1] & UDO SEIFFERT[2]

Abstract: Parametric geocoding is a standard procedure for orthorectification of hyperspectral airborne scanner data. Boresight parameters are estimated by using the image coordinates of precisely known ground reference points. To fully automate the processing of the recorded hyperspectral data ground reference points should be selected and matched by an algorithm without user interaction. In this article a two-step procedure is presented. First, direct geocoding of the data is done without ground reference points. This step is solely based on the data of the inertial measurement unit with differential GPS. Next, the hyperspectral data is normalized and pixels are clustered with an unsupervised method. Image blocks of an edge image of the resulting regions are matched with a high resolution edge image of a digital orthophoto of the calibration site. The found correspondences are used as ground control points in the second geocoding step.

1 Introduction

In this paper a robust algorithm for automated matching of low resolution hyperspectral and high resolution RGB orthophotos is presented. The matching results are used as ground control points for parametric geocoding of hyperspectral images. This automates the boresight calibration of the camera system. This is an important task for fast generation of geospatial data products based on remotely sensed data because it eliminates the need for user interaction.

Different approaches for automatic point selection exist. Automatic tie point generation for aerial triangulation is used for the registration of stereo images (SCHENK, 1997). As in (SKALOUD, 2007) described, the roofs of buildings can be detected in LiDAR images and matched across different flight tracks as a substitute for using precisely known ground reference points. Different tie point suitability indices and validation parameters for hyperspectral satellite images have been investigated in (MARCAL, 2007). In (XIONG, 2009) a super point detection algorithm based on interest point detection for matching of control points in satellite stereo image pairs has been presented.

For our photo mission flights the runway and the facilities of the base station airport were chosen as calibration targets. High resolution digital orthophotos are matched with short wave infrared (SWIR) hyperspectral images to find optimal boresight parameters. However, matching of images from different spectral ranges taken at different times is a challenging task. In (KNAPP, 2012) a method for automatic tie point selection for the fusion of hyperspectral and LiDAR images is presented which uses features in the overlapping spectral regions of different spectrometers. Their method cannot be applied for our datasets because such an overlap does not exist for a single sensor.

1) Uwe Knauer, Fraunhofer IFF, Biosystems Engineering, Joseph-v.-Fraunhofer-Str. 1, 39106 Magdeburg; E-Mail: uwe.knauer@iff.fraunhofer.de
2) Udo Seiffert, Fraunhofer IFF, Biosystems Engineering, Joseph-v.-Fraunhofer-Str. 1, 39106 Magdeburg; E-Mail: udo.seiffert@iff.fraunhofer.de

Instead, the use of edge and line features seems to be adequate because they are potentially visible in the reference RGB image and the SWIR hyperspectral image as they typically belong to material boundaries. Edge detection in the hyperspectral image is based on unsupervised clustering of the spectral data while the aerial photo is transformed with a traditional gradient based algorithm. By this, the method benefits both from the high resolution of current aerial photography and the high spectral resolution of current hyperspectral imaging devices. The calibration procedure consists of two steps. First, a coarse geocoding is done which is based on the positioning data of the inertial measurement unit only. Using the flat ground assumption with an average ground level for the observed region seems to be sufficient. Second, the algorithm searches corresponding ground control points. These are used to repeat geocoding. Then, a real digital elevation model is used to model the terrain. First results for the boresight calibration of five recent flights using the same calibration site are given.

2 Methods and Datasets

The Biosystems Engineering Group at Fraunhofer IFF operates an airborne hyperspectral SWIR imager from Norsk Elektro Optikk (NEO). For orthorectification of the hyperspectral images the software PARGE is used. This toolkit has been developed and is distributed by the company ReSe Applications Schläpfer. To fully automate the process workflow of PARGE a set of ground control points and their coordinates within the hyperspectral datacube must be provided.

2.1 Data Aquisition

The hyperspectral imager is operated in different airplanes using a stabilized mount (GSM3000) and an inertial measurement unit (Novatel SPAN-CPT with DGPS receiver). Hyperspectral images are recorded on solid state disks by the data acquisition unit from NEO. This device also performs the simultaneous recording of the IMU data. The lever arm to DGPS antenna is measured for each setup before flight.

2.2 Parametric Geocoding

According to the PARGE manual the standard workflow of geocoding consists of the following steps (SCHLÄPFER, 2011):

- Data preparation
    - Define image data
    - Define sensor model
    - Load IMU/DGPS data
- Preparing Digital Elevation Model (DEM)
- Import of ground control points (GCP)
- Estimation and setting of offsets
    - Roll and pitch, heading, FOV, height, …
- Main Processing

  - Calculation of the Image Geometry Map
  - Scan Angles for ATCOR
- Final Processing
  - Geocoded data cube
  - Geocoded RGB

Typically, this workflow is processed manually by selecting and configuring the parameters of each step within the user interface of the PARGE software. Additionally, PARGE offers an API for automation of the workflow. This interface and PARGE itself are based on the IDL programming language.

2.3 Automated Selection and Matching of Ground Control Points

First, the PARGE workflow runs without GCPs. This provides a rough estimate of the orthorectified image. Second, the coordinates of candidate control points are randomly selected. This selection is done in the coordinate system of the unaligned image. Hence, the coordinates are the line number and the pixel number. They are projected to UTM coordinates using the image geometry map which is the result of the first main processing step. Fig. 1 shows a set of candidate points and their projection onto the georeferenced orthophoto.

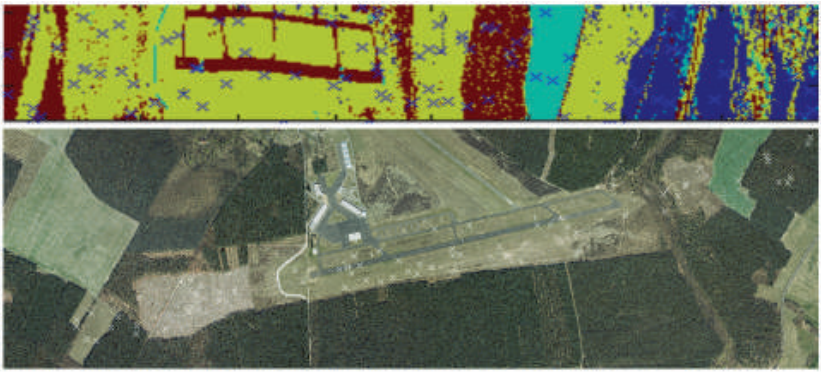


Fig. 1: Randomly chosen candidate control points in raw coordinate system (top) and projected to UTM coordinates with PARGE (bottom)

Next, the hyperspectral data is clustered using the k-means algorithm (LLOYD, 1982). However, before clustering a normalization of illumination is done. Normalization is not a complete ATCOR procedure, but estimates the spectra of sunlight and corrects the overall intensity. Then, all pixels are classified to belong to one of the resulting k classes. The boundaries of the connected regions in the resulting label map are extracted to be used for matching with the edge

image from a RGB reference image. To suppress small regions, k is set to a small value (e.g. k=6) and only the largest regions are used for edge extraction.

The edge map of the RGB reference is generated with a gradient based algorithm. However, the parameters of the so called Canny edge detector (CANNY, 1986) are adapted to the gradients of the current image. Adaption is controlled by the number of edge pixels in the hyperspectral edge map as shown in Fig. 2.

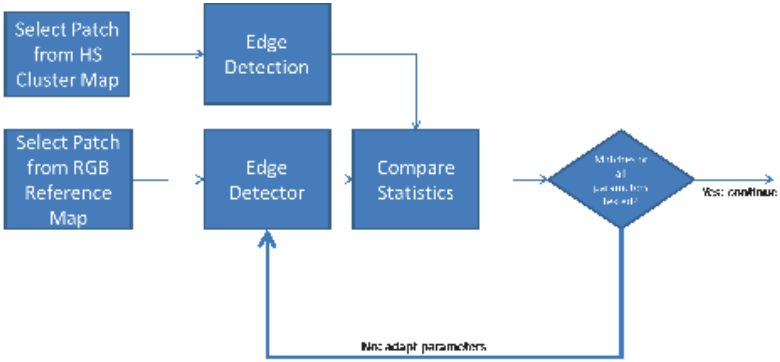


Fig. 2: Flowchart with feedback loop for selection of edge detection parameters

Fig. 3 shows a typical result of edge detection for the reference image and the hyperspectral image. The matching between the RGB based edge map and the hyperspectral edge map is done by two dimensional correlation between image patches centered on the randomly chosen candidate points.

Finally, the number of reference points is reduced based on the appearance of the edge maps as well as based on outlier detection for the estimated translation vectors between the RGB image patches and the hyperspectral edge map patches.

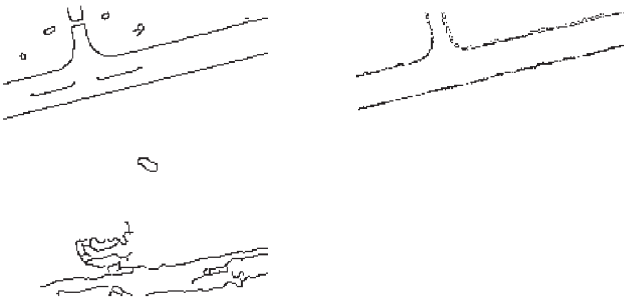


Fig. 3: Edge images generated from RGB reference orthophoto (left) and from the label image of the clustered hyperspectral image (right)

2.4 Datasets

A hyperspectral image of the airfield was recorded as part of each photo mission. For testing the automated boresight calibration approach five of these recordings are analyzed.

The reference image is a digital orthophoto of the base station airport with a ground resolution of 20 cm per pixel.

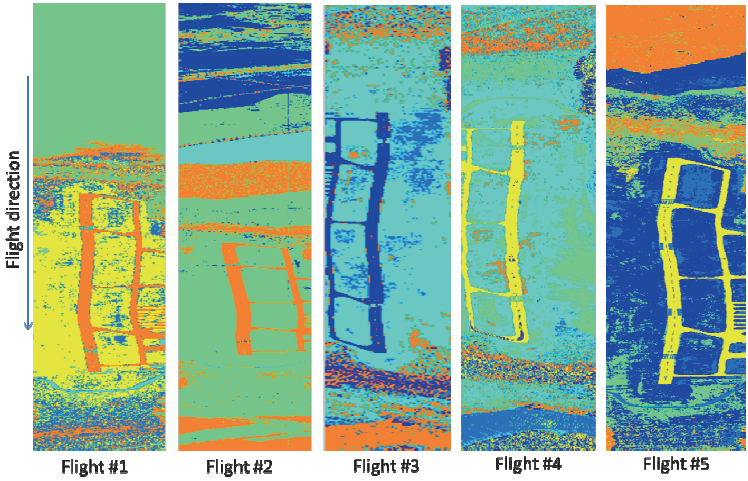


Fig. 4: Results of clustering the hyperspectral data of the calibration track for different flight missions

3 Results and Discussion

The objective of unsupervised clustering of the hyperspectral data is to derive a high quality edge image of the observed area. Therefore, it is important to extract the boundaries of ground structures such as the runway and streets. On the other hand, edges which are related to the vegetation should be treated as irrelevant. Fig. 4 shows the segmentation results based on unsupervised clustering for five flights over the calibration site. For compactness of presentation the not rectified images are show. The results show that boundaries of the ground structures can be successfully segmented for all flights and hence can be used for matching.

Fig. 5 visualizes the distributions of the length of translation vectors between matching pairs of ground control points for flights #3 and #5. Translation vectors are used to update the UTM coordinates of the reference points in the hyperspectral image. The vectors length is a measure for the error of direct geocoding with PARGE using IMU/DGPS data only. However, also errors made by the proposed matching algorithm contribute to the shown distributions and require an outlier detection. The results show clearly that quality of direct geocoding is different for all flight missions.

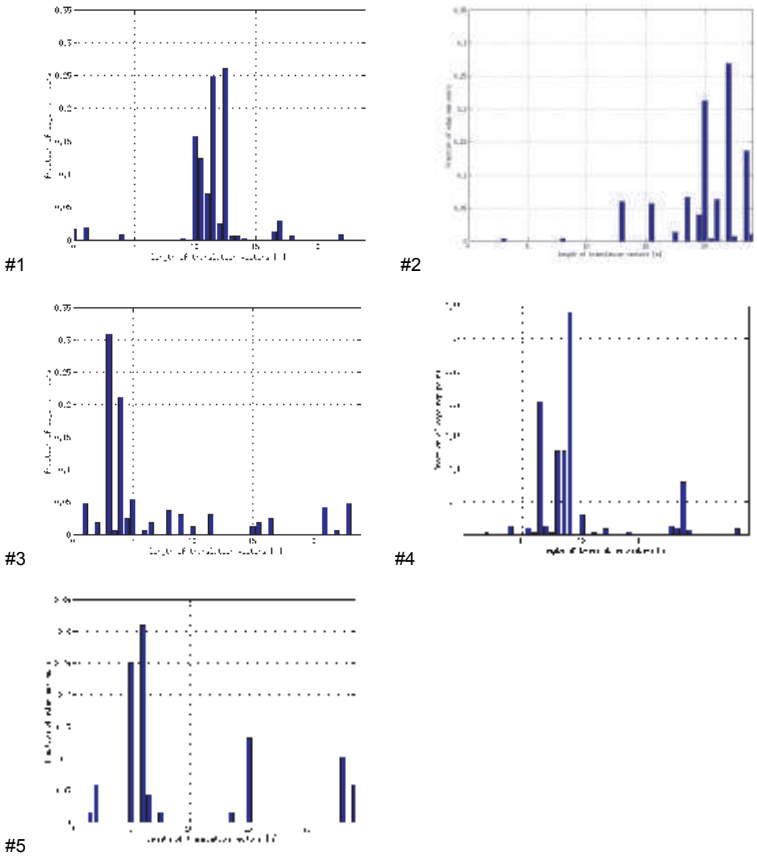


Fig. 5: Distributions of the length of translation vectors between the coordinates of matching points

Finally, the calculated correspondences are used as ground control points in the PARGE workflow. Now, the boresight parameters for each flight can be estimated and later be used as offset values for the direct geocoding of all the other flight tracks.
In Fig. 6 results of geocoding without ground control points are compared with geocoding using the proposed algorithm.

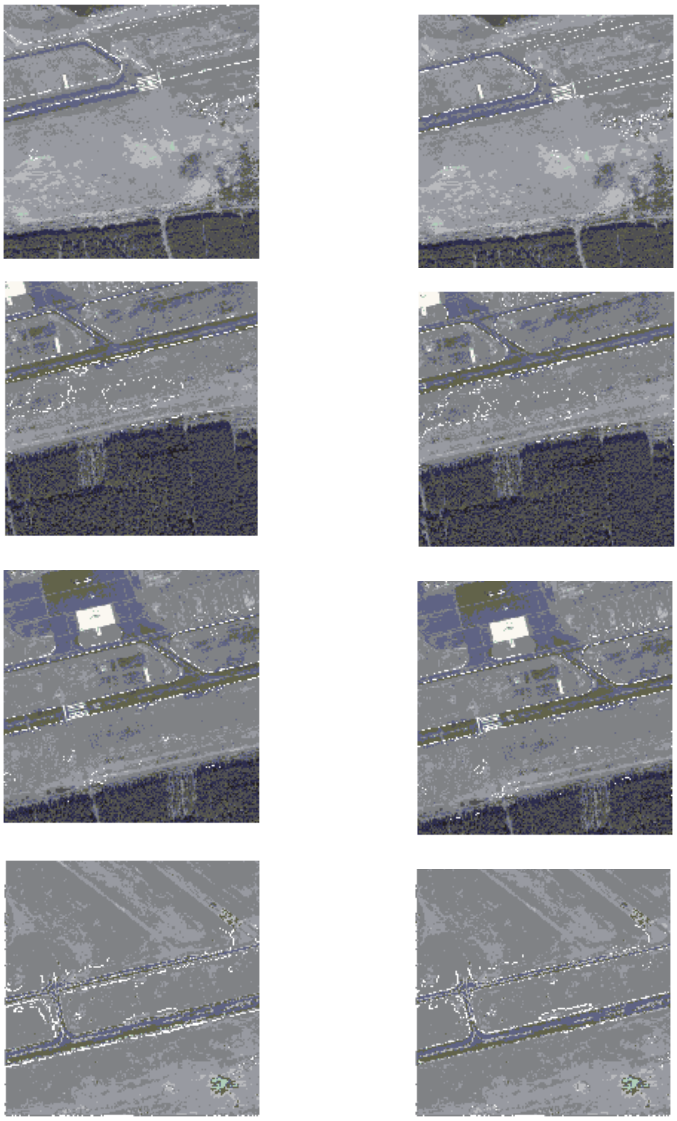


Fig. 6: Before (left column) and after (right column) matching of ground reference points

4 Conclusion

A fully automated geocoding procedure is needed in applications which require fast generation of maps based on remotely sensed data. When a final data product has to be provided shortly after landing to allow a decision making for instance in agriculture, forestry, or security applications a trade-off between the highest possible accuracy and the use of fast but less accurate methods exists. We presented such a robust algorithm for automated matching of randomly chosen ground control points in low resolution hyperspectral data and high resolution RGB orthophotos. The results show that the approach provides a satisfying level of accuracy compared to the error which is made by using direct geocoding based on the IMU/DGPS data only. However, the article describes work in progress. Therefore, modifications of the algorithm as well as different metrics for performance evaluation have to be investigated.

5 References

CANNY, J., 1986: A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 8, 6, 1986, pp. 679-698.

KNAPP, D.E., 2012: Fusing Hyperspectral and LiDAR data from CAO-VSWIR for Increased Data Dimensionality. AGU fall meeting, San Francisco, 2012.

LLOYD, S. P., 1982: Least squares quantization in PCM. IEEE Transactions on Information Theory. 28, 1982, pp. 129–137, doi:10.1109/TIT.1982.1056489.

MARCAL, A.R.S., 2007: Automatic identification and validation of tie points on hyperspectral satellite images from CHRIS/PROBA. SCIA'07 Proceedings of the 15th Scandinavian conference on Image analysis, pp. 553-561.

SCHENK, T., 1997: Towards Automatic Aerial Triangulation. ISPRS Journal of Photogrammetry & Remote Sensing 52, 1997, pp. 110-121.

SCHLÄPFER, D., 2011: Parametric Geocodig Orthorectification for Airborne Scanner Data. User Manual, Version 3.1, page 47.

SKALOUD, J. & SCHAER, P., 2007: Towards Automated LiDAR Boresight Self-calibration. 5th International Symposium on Mobile Mapping Technology, Padova (Italy), May 29-31, 2007.

XIONG, Z., 2009: Technical Development for Automatic Aerial Triangulation of High Resolution Satellite Imagery. Technical Report 268, UNB Geodesy and Geomatics Engineering 2009.

Genauigkeit der inversen Navigation bei hoher Passpunktsichtbarkeit zur Absolutpositionierung eines geführten Messsystems

THOMAS DARR[1], CARSTEN GÖTZ[1], SEBASTIAN TUTTAS[1], TOBIAS BECKER[2] & UWE STILLA[1]

Zusammenfassung: Im Beitrag wird die Genauigkeit in der Bestimmung der äußeren Orientierung einer Einzelkamera durch räumlichen Rückwärtsschnitt in Abhängigkeit unterschiedlicher Parameter in einer fiktiven Messumgebung untersucht. Darauf aufbauend wird ein Multi-Kamerasystem betrachtet, das mit vier im 90°-Winkel zueinander montierten Kameras mehrere Raumrichtungen beobachtet. In einem praktischen Versuch mit dem Multi-Kamerasystem werden die realen Abweichungen der äußeren Orientierung bestimmt und mit den theoretisch erreichbaren Genauigkeiten verglichen. Während die Simulation Abweichungen im Bereich unter einem Zehntel Millimeter in der Position liefert, zeigt der praktische Versuch um bis zum Faktor fünf höhere Abweichungen. Der Grund hierfür liegt in der beim praktischen Versuch eingeschränkten Sichtbarkeit von Referenzpunkten.

1 Einleitung

1.1 Motivation

Bei der Produktion in der Automobilindustrie nimmt die Qualitätskontrolle einen hohen Stellenwert ein. Zur Geometriemessung werden neben Koordinatenmessmaschinen zunehmend photogrammetrische 3D-Messsysteme verwendet. Diese besitzen den Vorteil, dass sie 3D-Informationen flächenhaft erfassen. Im Projekt „MultiGO“ (Multifunktionale Messzelle zur Geometrie- und Oberflächenprüfung), an dem die Bayerische Forschungsstiftung, die BMW Group, das Fachgebiet Photogrammetrie und Fernerkundung der TU München und die Steinbichler Optotechnik GmbH beteiligt sind, wird mit Hilfe eines optischen Messsensors die Geometrie von Fahrzeugteilen überprüft. Derzeit wird dafür ein aufwendiger Rahmen um das Messobjekt konstruiert, der die photogrammetrische Verknüpfung der Einzelaufnahmen erlaubt. Bei „MultiGO“ ist es ein Teilziel durch inverse Navigation diesen individuellen Kulissenaufbau zu reduzieren und stattdessen ein globales Referenzpunktfeld zu verwenden. Abb. 1 zeigt den Aufbau der Messzelle, in der an der Wand die Referenzpunkte angebracht sind. Die Orientierung des Messsensors wird mit Hilfe dieser bestimmt.

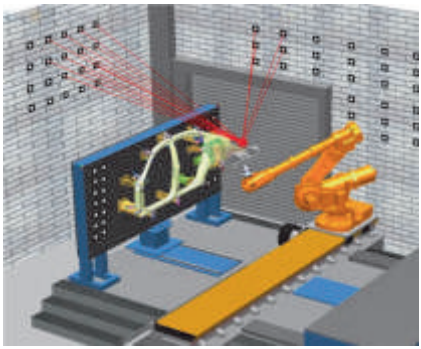


Abb. 1: Aufbau der Messzelle

1) Photogrammetrie & Fernerkundung, Technische Universität München, Arcisstraße 21, 80333 München, http://www.pf.bv.tum.de
2) BMW AG, Petuelring 130, 80788 München

Die Schwierigkeit in der Umsetzung dieses Konzepts besteht in den hohen Anforderungen an die Orientierungsgenauigkeit des Sensors. Außerdem muss die Genauigkeit der angebrachten Referenzpunkte, z. B. an Wand und Boden, über lange Strecken und Zeit gewährleistet werden. Um eine hohe Genauigkeit zu erreichen wird ein aus vier Kameras bestehendes Kamerasystem verwendet. Dieses blickt in vier verschiedene Raumrichtungen, um möglichst viele Referenzpunkte erfassen zu können.

1.2 Verwandte Arbeiten

In LUHMANN (2009) wurden die sechs Freiheitsgrade der äußeren Orientierung und deren Genauigkeiten von einem Objekt (Lokator) innerhalb eines Referenzsystems mit Hilfe einer Kamera bestimmt. Im Gegensatz zur hier vorliegenden Arbeit, bei der die Genauigkeit der Kameraposition und -orientierung von Interesse ist, liegt der Schwerpunkt der Untersuchung in der Genauigkeit der Lokatorposition. Diese Methode findet z.B. Anwendung bei der Crashtest-Analyse. FRAHM et al. (2004) untersuchten ein Multi-Kamerasystem, welches eine feste und bekannte Anordnung der Kameras aufweist, sodass die relativen Orientierungen dieser Kameras berechnet werden können. Bei MUHLE et al. (2008) wurde ein ähnliches Themengebiet behandelt: Es wurden zwei fest verbundene Stereokamerasysteme untersucht, die sich starr im Raum bewegen. Die letzten beiden Ansätze gehen nicht näher auf die erreichbare Genauigkeit im industriellen Umfeld ein. GÖTZ et al. (2011) evaluierten mehrere Konfigurationen von Kamera und Referenzpunkten, um die in der Industrie geforderten Genauigkeiten zu erhalten. Ebenso wie GÖTZ et al. (2011) ist das Ziel dieser Arbeit, dass die Genauigkeit der Kameraorientierung den Anforderungen der industriellen Fertigung genügen soll. Im Gegensatz zu HANEL et al. (2013) wird in diesem Beitrag die Grenzleistung des Multi-Kamerasystem im dichten Referenzpunktfeld des Außenraums untersucht.

2 Methodik

In dem vorliegenden Beitrag wird zuerst die Genauigkeit einer Kameraorientierung bei Variation der beeinflussenden Größen (siehe Abschnitt 3) untersucht. Darauf aufbauend wird die Genauigkeit eines Multi-Kamerasystems betrachtet. Die Vorgehensweise für die theoretischen Untersuchungen ist folgende:

1. Betrachten eines Referenzpunktfeldes
2. Berechnen der Bildpunkte aus den sichtbaren Referenzpunkten
3. Berechnung der äußeren Orientierung und deren Genauigkeit durch den räumlichen Rückwärtsschnitt

2.1 Einzelkamera

Als Ansatz für den räumlichen Rückwärtsschnitt werden die *Kollinearitätsgleichungen* als Verbesserungsgleichungen nach einer vermittelnden Ausgleichung verwendet. Da die Koordinaten der Referenzpunkte nicht fehlerfrei bestimmt werden können, werden diese zusätzlich in die Ausgleichung eingeführt. Es wird die Notation nach KRAUS (1996) verwendet:

$$
\begin{aligned}
x'_i + \hat{v}_{x'_i} &= x'_0 - c \cdot \frac{\hat{r}_{11} \cdot (\hat{X}_i - \hat{X}_0) + \hat{r}_{21} \cdot (\hat{Y}_i - \hat{Y}_0) + \hat{r}_{31} \cdot (\hat{Z}_i - \hat{Z}_0)}{\hat{r}_{13} \cdot (\hat{X}_i - \hat{X}_0) + \hat{r}_{23} \cdot (\hat{Y}_i - \hat{Y}_0) + \hat{r}_{33} \cdot (\hat{Z}_i - \hat{Z}_0)} \\
y'_i + \hat{v}_{y'_i} &= y'_0 - c \cdot \frac{\hat{r}_{12} \cdot (\hat{X}_i - \hat{X}_0) + \hat{r}_{22} \cdot (\hat{Y}_i - \hat{Y}_0) + \hat{r}_{32} \cdot (\hat{Z}_i - \hat{Z}_0)}{\hat{r}_{13} \cdot (\hat{X}_i - \hat{X}_0) + \hat{r}_{23} \cdot (\hat{Y}_i - \hat{Y}_0) + \hat{r}_{33} \cdot (\hat{Z}_i - \hat{Z}_0)} \\
X_i + \hat{v}_{X_i} &= \hat{X}_i \\
Y_i + \hat{v}_{Y_i} &= \hat{Y}_i \\
Z_i + \hat{v}_{Z_i} &= \hat{Z}_i
\end{aligned}
\tag{1}
$$

Für das stochastische Modell wird im Weiteren von unkorrelierten Beobachtungen n ausgegangen. Die Kovarianzmatrix der Beobachtungen $\mathbf{Q}_{bb}$ entspricht somit einer $n \times n$-Diagonalmatrix. In diesem Abschnitt werden die quadrierten Standardabweichungen, die an die Bild- und Referenzpunkte angebracht werden, als Varianzen auf der Hauptdiagonale eingesetzt. In Abschnitt 2.2 werden die gegebenen Varianzen der Referenzpunkte verwendet. Ausgehend vom Funktionalmodell (1) können die auf die Genauigkeit der äußeren Orientierung einflussnehmenden Größen identifiziert werden. Die unmittelbar in den Gleichungen vorkommenden Größen sind die Kamerakonstante *c*, die Bildpunkte *x'*, *y'* und die Referenzpunkte (*X, Y, Z)*. Bei den Bild- und Referenzpunkten spielt sowohl die Anzahl der sichtbaren Punkte, als auch deren Genauigkeit eine Rolle. Ebenso kommen in den Gleichungen die Translationsparameter der äußeren Orientierung (X_0, Y_0, Z_0) direkt vor. Es kann also der Abstand der Kamera zur Referenzpunktebene betrachtet werden. In Abschnitt 3 wird der Einfluss der Kamerakonstanten *c* näher untersucht.

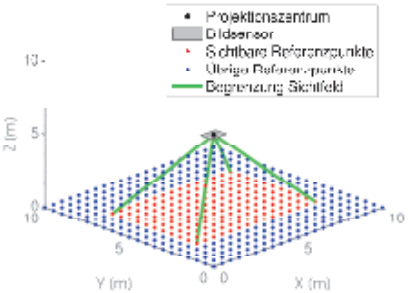


Abb. 1: Virtuelle Messzelle zur Untersuchung der Orientierungsgenauigkeit einer Kamera

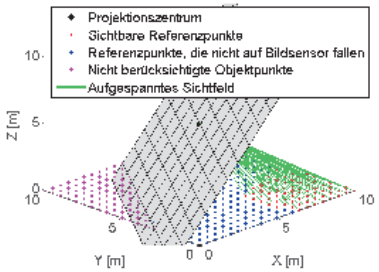


Abb. 2: Aufnahmesituation zur Berechnung der Bildpunkte

Für diese Untersuchungen wird eine virtuelle Messzelle (siehe Abb. 2) verwendet. Die Parameter für die Position, das Sichtfeld der Kamera, die Anzahl der sichtbaren Referenzpunkte, etc. sind bei den folgenden Betrachtungen nicht fest, sondern werden je nach Abhängigkeit der Kamerakonstanten *c* verändert.
Die Kamera in Abb. 2 blickt senkrecht nach unten auf ein Referenzpunktfeld in gleichmäßigem Raster. Die grünen Linien markieren die Eckpunkte des Kamerasichtfeldes, in dem die rot darge-

stellten Punkte auf den Bildsensor (graues Rechteck) fallen. Im Gegensatz zu Abb. 1 ist hier eine ideale Messkonfiguration dargestellt.
In Abb. 3 ist die Berechnung der Bildpunkte aus den vorhandenen Referenzpunkten dargestellt. Diese wird benötigt, falls das Punktfeld nicht senkrecht betrachtet wird, was in der Realität der Normalfall ist. Mit Hilfe der *Hesse'schen Normalform* der Bildebene können diejenigen Referenzpunkte identifiziert werden, die – bezogen auf den Koordinatenursprung – auf der gleichen Seite der Ebene wie das Projektionszentrum liegen. Abb. 3 zeigt die vorgenommene Einteilung der Referenzpunkte, die nicht auf den Bildsensor fallen (in Magenta dargestellt) und von der Berechnung der Bildpunkte ausgeschlossen werden. Die Koordinaten der Referenz- und Bildpunkte müssen somit zuerst berechnet werden, um in das Funktionalmodell (1) eingeführt werden zu können.

2.2 Multi-Kamerasystem

Nun wird das System von einer einzelnen Kamera zu einem Multi-Kamerasystem erweitert. Durch Messung mit mehreren Kameras kann der Raum in mehreren Raumrichtungen beobachtet werden. Kennt man die einzelnen Positionen der Kameras zueinander, kann eine höhere Genauigkeit der Position des Gesamtsystems durch Ausnutzung der Redundanz erreicht werden. Für die Berechnung der Orientierung des Gesamtsystems wird das Koordinatensystem einer Kamera als Referenzkoordinatensystem gewählt.
Um Aussagen über die Orientierungsgenauigkeit des Multi-Kamerasystems entlang ihres Messpfades zu erhalten, wird folgendermaßen vorgegangen: Als erstes werden die Positionen und Ausrichtungen der Kameras in einem Modellkoordinatensystem relativ zur Referenzkamera definiert. Mit Hilfe der vorgegebenen Orientierungen (=Messpfad) der Referenzkamera im übergeordneten Referenzkoordinatensystem können sämtliche Kameras in dieses transformiert werden. Zur Vereinfachung werden die Positionen im Referenzkoordinatensystem als fehlerfrei angesehen. Das transformierte Modellkoordinatensystem entspricht somit den Sollwerten in der später folgenden *bedingten Ausgleichung*. Die äußere Orientierung einer einzelnen Kamera berechnet sich nach der in Abschnitt 2.1 dargestellten Methode. Im Gegensatz zur simulierten Messzelle sind die Standardabweichungen der Referenzpunkte fest vorgegeben. Die Ergebnisse (Istwerte) werden mit den bekannten Orientierungen im Referenzkoordinatensystem (Sollwerte) anschließend als Beobachtungen in das Funktionalmodell (2) eingeführt, wobei nur die Translationen benötigt werden. Es werden nur diejenigen Kameras berücksichtigt, die mindestens vier Referenzpunkte sehen und sich somit mit Hilfe der Ausgleichungsrechung positionieren können. Das allgemeine Funktionalmodell lautet:

$$\hat{\mathbf{X}}_{0_{K_R,I}} - \hat{\mathbf{X}}_{0_{K_i,I}} - (\mathbf{X}_{0_{K_R,S}} - \mathbf{X}_{0_{K_i,S}}) = \mathbf{0}$$

mit

$i = 2 \ldots n$ Kameras

I = Istwerte

S = Sollwerte

K = Kamera

R = Referenzkamera

(2)

3 Theoretische Untersuchungen zu den Einzelkameras

Für die Simulation werden die Parameter an den praktischen Versuch (Abschnitt 4.2) angelehnt. So wurde die Abmessung des Sensors in der Breite zu 8,8 mm und in der Höhe zu 6,6 mm gewählt. Die Referenzpunkte sind frei gewählt und fehlerfrei. Durch die mathematische Abbildung durch die *Kollinearitätsgleichungen* dieser auf den Bildsensor sind die Bildpunkte ebenfalls fehlerfrei. Um eine realistische Situation für die Bild- und Referenzpunkte zu schaffen, werden diese mit künstlichen und zufälligen Standardabweichungen versehen. Die typische Standardabweichung beträgt für die Referenzpunkte 0,1 mm und für die Bildpunkte 0,5 µm. Um generelle Aussagen zu machen, werden mehrere Durchläufe gerechnet und das Ergebnis über diese gemittelt.

3.1 Kamerakonstante *c* bei veränderlichem Bildmaßstab

Es wird untersucht, wie sich die Genauigkeit der Kameraorientierung verhält, wenn die Kamerakonstante *c* im Bereich von 6 mm bis 20 mm variiert. Da nur die Kamerakonstante verändert wird und die räumliche Position der Kamera gleich bleibt, ändert sich zwangsläufig der Bildmaßstab *M*.

Bei zunehmender Kamerakonstante wird der Bildausschnitt kleiner und somit können weniger Punkte gesehen werden. Damit die Anzahl der beobachteten Referenzpunkte die Untersuchungsbedingungen nicht verändern, wird als Zwischenschritt mit der Kamera ein Testpunktfeld verwendet. Anschließend werden die gerade noch sichtbaren Referenzpunkte an den Ecken des Feldes identifiziert und es wird ein neues Punktfeld mit gleich vielen Referenzpunkten aufgespannt. Die Abstände der Referenzpunkte variieren hierbei je nach Größe des Kamerasichtfeldes, d. h. bei größerer Kamerakonstante werden die Abstände kleiner.

Mit Hilfe dieser neu gesetzten Referenzpunkte werden nun durch die *Kollinearitätsgleichungen* die entsprechenden Bildpunkte berechnet. Die Bild- und Referenzpunkte werden nach dem Funktionalmodell (1) als Beobachtungen für eine Ausgleichung nach vermittelnden Beobachtungen verwendet.

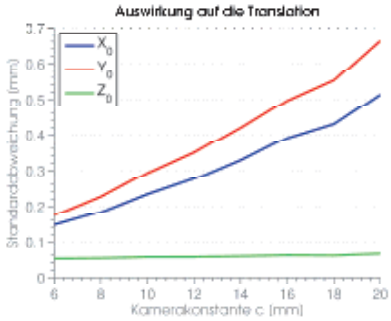


Abb. 3: Auswirkung auf die Translation bei variierender Kamerakonstante und veränderlichem Bildmaßstab

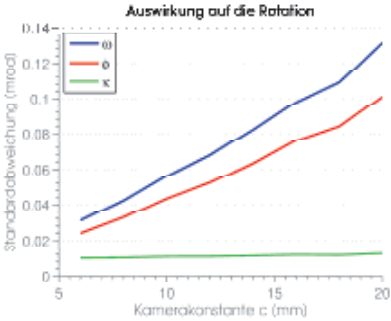


Abb. 4: Auswirkung auf die Rotation bei variierender Kamerakonstante und veränderlichem Bildmaßstab

Die Abb. 4 und Abb. 5 zeigen, dass die Genauigkeit bei zunehmender Kamerakonstante abnimmt. Die Blickrichtung der Kamera in Richtung der *Z*-Achse bewirkt, dass die Standardabweichung der Translation in dieser Richtung bereits sehr gering ist und kaum von einer größer werdenden Kamerakonstante beeinflusst wird. Durch die Aufnahmerichtung folgt auch, dass sich der Rotationswinkel κ analog zu Z_0 verhält. Die Standardabweichungen der übrigen Parameter steigen bei größerer Kamerakonstante nahezu linear an. Grund hierfür ist der bereits erwähnte kleiner werdende Abstand der betrachteten Referenzpunkte. Die Strahlen, welche die Referenzpunkte mit dem Projektionszentrum verbinden, schneiden sich bei kleinerem Sichtfeld zunehmend schleifend. Dadurch sinkt die Genauigkeit der Positionsbestimmung der Kamera. Auf die Genauigkeit der Rotationswinkel ω und φ wirkt sich das kleiner werdende Sichtfeld ebenfalls negativ aus: Die geringe räumliche Ausdehnung der Referenzpunkte – und damit auch der Bildpunkte auf dem Sensor – erlaubt keine genauere Bestimmung der Kameraorientierung.
Durch die Sensorabmessungen bedingt, verhalten sich die Translation in X- und Y -Richtung leicht unterschiedlich. Die Abstände in X-Richtung sind größer als in Y -Richtung, sodass erstere etwas besser bestimmt wird. Die Rotationswinkel verhalten sich zur Translation umgekehrt.

3.2 Kamerakonstante *c* bei konstantem Bildmaßstab

Der Einfluss des Bildmaßstabs wird nun eliminiert, indem die Position des Projektionszentrums in Abhängigkeit der Kamerakonstante verändert wird.
Mit Hilfe der *Hesse'schen Normalform* kann der Abstand des Projektionszentrums zur Referenzpunktebene bestimmt werden. Dadurch kann der Bildmaßstab für jede Kamerakonstante abgeleitet werden. Mit der Wahl eines Referenzmaßstabs, dem Lotfußpunkt des Projektionszentrums auf der Ebene und dem normierten Normalenvektor der Ebene wird das Projektionszentrum in Abhängigkeit der Kamerakonstante verschoben. Im vorliegenden Fall blickt die Kamera exakt senkrecht auf das Punktfeld, sodass die Kamera in Richtung der positiven Z-Achse verschoben werden kann. Dadurch werden in jeder Position gleich viele Referenzpunkte gesehen.
Anschließend werden analog zu Abschnitt 3.1 die Genauigkeiten der Translation und Rotation berechnet.

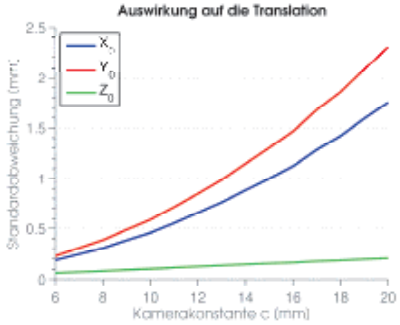


Abb. 5: Auswirkung auf die Translation bei variierender Kamerakonstante und gleichem Bildmaßstab

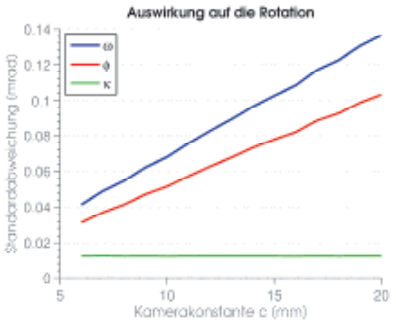


Abb. 6: Auswirkung auf die Rotation bei variierender Kamerakonstante und gleichem Bildmaßstab

Als Ergebnis für die Orientierungsgenauigkeiten erhält man die Plots in Abb. 6 und Abb. 7.
Die Standardabweichungen zeigen auf den ersten Blick ein ähnliches Ergebnis wie in Abb. 4 und Abb. 5. Die Translation in X- und Y -Richtung steigt bei zunehmender Kamerakonstante, wohingegen die Translation in Z-Richtung sehr gut bestimmt wird und beinahe unabhängig von der Variation der Kamerakonstante ist. Der Grund für das Verhalten der Translation in Z ist ebenfalls derselbe wie bei der Variation der Kamerakonstante c mit veränderlichem Bildmaßstab. Die Standardabweichung der Rotationswinkel steigt auch beinahe linear mit erhöhter Kamerakonstante an. Die Genauigkeit der Translationsparameter sinkt quadratisch bei einer größeren Kamerakonstante.
Weitwinkelobjektive mit geringer Kamerakonstante haben somit einen Vorteil bei der Orientierungsbestimmung.

4 Untersuchungen zum Multi-Kamerasystem

4.1 Vorgehensweise

Es wird eine bestimmte Anordnung der Kameras auf dem Multi-Kamerasystems ausgewählt, anhand einer Simulation näher untersucht und durch einen praktischen Versuch in einer Messzelle im BMW-Werk München evaluiert.
Bei den Untersuchungen zum Multi-Kamerasystem werden die Koordinaten einer Referenzkamera im übergeordneten Koordinatensystem vorgegeben. Anhand dieses gegebenen Messpfades wird eine Anordnung der Kameras in 90°-Winkeln hinsichtlich der Orientierungsgenauigkeit betrachtet. Die Umgebung, in der das Kamerasystem untersucht wird, ist eine virtuelle Messzelle, die an die Gegebenheiten der Messzelle im BMW-Werk München angepasst wurde und in der die Referenzpunkte frei gewählt wurden. Zur Vereinfachung werden in der Messzelle keine Verdeckungen berücksichtigt.

4.2 Versuchsaufbau

Die praktischen Untersuchungen für diese Arbeit wurden in einer Messzelle im BMW-Werk München durchgeführt. Der Aufbau der Messzelle ist in Abb. 8 dargestellt. Am Flansch eines Industrieroboters wurde für diesen Versuch ein maßstabs- und detailgetreuer Gehäusedummy des für das Projekt „MultiGO“ entwickelten Sensorsystems der Firma Steinbichler Optotechnik GmbH verwendet. Auf die Oberseite des nachgebildeten Sensors wurden vier Kameras angebracht. Drei Kameras verwenden Objektive mit einer Kamerakonstante von $c = 6{,}1$ mm, die vierte Kamera besitzt ein Objektiv mit $c = 8$ mm. Die Anordnung der Kameras ist in Abb. 10 dargestellt. Im Innern des Dummys befinden sich die Stromversorgung der Kameras sowie die Datenübertragung der Bilder per LAN zu einem Laptop.

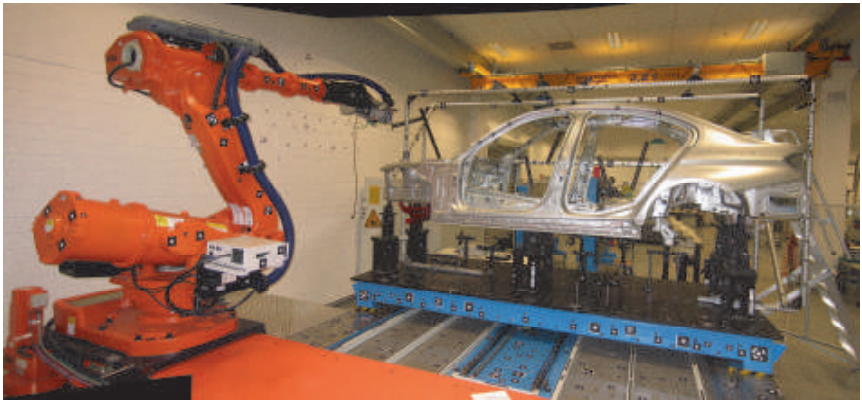


Abb. 7: Aufbau der Messzelle des praktischen Versuchs

Als Vorbereitung für die eigentliche Messung wurden eine Kalibrierung aller vier Kameras, eine photogrammetrische Einmessung der Referenzpunkte im Referenzkoordinatensystem der Messzelle und eine Optimierung der Schärfentiefe (Fokussierung) der Kameras mit Hilfe des Siemenssterns durchgeführt.
Nachdem nun diese vorbereitenden Maßnahmen abgeschlossen wurden, konnte der Versuch durchgeführt werden. Insgesamt wurden 18 Positionen im Außenraum der Karosserie (siehe Polygon Abb. 9) angefahren. Zum einen sind im Außenraum kaum (z.B. durch den Roboterarm) Verdeckungen vorhanden, zum anderen können durch die Ausrichtung der Kameras in vier Raumrichtungen auch Punkte auf der gegenüberliegenden Seite der Messzelle gesehen werden. Die Positionen wurden ausgewählt, da der Geometriesensor an diesen Stellen geometrische Eigenschaften der Karosserie (z.B. Löcher) vermisst.
Die Auswertung des Versuchs liefert die äußeren Orientierungen des Kamerasystems an den Messpositionen, sowie die zugehörigen Genauigkeiten.

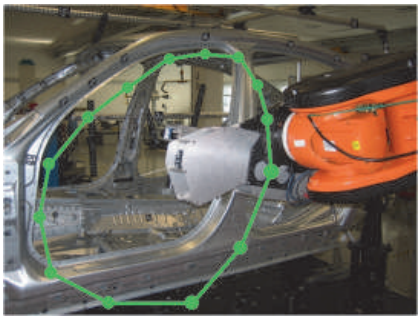


Abb. 8: Messpfad entlang der Karosserie

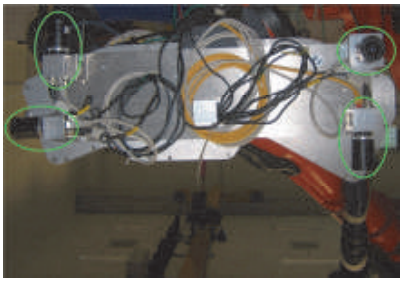


Abb. 9: Anordnung der Kameras (grün eingekreist) auf dem Sensordummy

4.3 Ergebnisse

In Abschnitt 3 wurde das Verhalten der Orientierungsbestimmung bei Variation der beeinflussenden Größen einer einzelnen Kamera untersucht. Nun werden die Untersuchungen auf ein Mehrkamerasystem erweitert. Die Blickrichtungen von vier Kameras werden mit 90°-Winkeln zueinander montiert (siehe Abb. 10). Durch diese Anordnung soll für die dreidimensionale Messzelle eine möglichst hohe überlappungsfreie Abdeckung erreicht werden. Um Genauigkeiten für die Position des Kamerasystems in Bezug auf eine Referenzkamera zu erhalten, entspricht der in Abschnitt 2.2 beschriebenen Vorgehensweise.

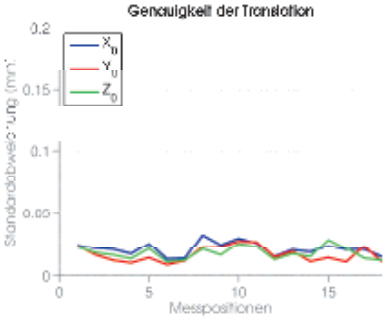


Abb. 10: Theoretische Translationsgenauigkeit des Kamerasystems an den Messpositionen

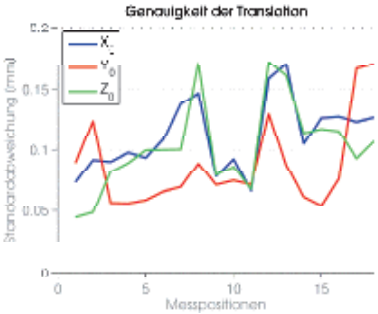


Abb. 11: Translationsgenauigkeit des Kamerasystems im praktischen Versuch

Der Vorteil dieser Konfiguration (siehe Abb. 11) ist, dass die Genauigkeit der Translation mit der des praktischen Versuchs (siehe Abb. 12) verglichen werden kann. Die Größenordnung der Genauigkeiten der Translation des simulierten Kamerasystems ist um den Faktor 3-5 besser als beim praktischen Versuch. Der primäre Grund hierfür ist die vorgenommene Vereinfachung, dass keine Verdeckungen berücksichtigt werden. Im praktischen Versuch ist die Sicht der Kameras durch Objekte eingeschränkt: Die erste Kamera schaut beispielsweise an der zwölften Messposition (siehe Abb. 11 und Abb. 12) von oben auf das Messobjekt. Es ist ersichtlich, dass die Referenzpunkte, die unter dem Fahrzeug auf der Messplatte befestigt sind, nicht gesehen werden können. In der gerechneten Simulation ist dies jedoch möglich, sodass die Ausgleichung zur Bestimmung der Translationsgenauigkeit eine höhere Redundanz hat. Ferner wird im praktischen Versuch ebenfalls der Translationsparameter Y_0 an den Zwischenpositionen am besten bestimmt.
Es zeigen sich jedoch auch Unterschiede: Während bei der simulierten Translationsgenauigkeit die Standardabweichungen in der 12. Position sinken, steigen diese im praktischen Versuch, durch die fehlende Orientierung der 4. Kamera an. Eine mögliche Erklärung ist, dass die Daten des Versuchs mit einer kommerziellen Software (AICON 3D STUDIO) ausgewertet wurden. Die genaue Auswertemethode ist unbekannt, es kann beispielsweise für die Ausgleichung des Gesamtsystems ein anderes Funktionalmodell verwendet worden sein. Des Weiteren ist davon auszugehen, dass die Fehler der Bildpunkte nicht zufällig hinzugefügt werden. Während bei den Versuchen zufällige Fehler zu den Bildpunkten hinzugefügt werden, sind diese Zuschläge bei der Software nicht bekannt.

5 Zusammenfassung und Ausblick

Bei den Untersuchungen der Einzelkamera hat sich mit Variation der Kamerakonstante gezeigt, dass kleinere Kamerakonstanten eine höhere Genauigkeit erreichen. Weitwinkelobjektive haben somit einen Vorteil bei der Orientierungsbestimmung. Die Anordnung der Kameras beim Multi-Kamerasystem wurde so gewählt, dass sie einen guten Vergleich mit dem praktischen Versuch bietet. Es hat sich gezeigt, dass bei der Simulation die Positionsbestimmung im Bereich unter einem zehntel Millimeter möglich ist. Jedoch werden dabei keine Verdeckungen berücksichtigt, welche das Ergebnis deutlich verschlechtern. Um diesen Einfluss zu verringern, kann untersucht werden, ob z.B. ein kleineres Bauteil, das weniger Referenzpunkt verdeckt, eine Verbesserung der Genauigkeit der Orientierungsparameter bewirkt. Weiterhin könnte bei kleineren Bauteilen eine kleinere Messzelle die Homogenität des Referenzpunktfeldes verbessern.

6 Danksagung

Das Projekt „MultiGO“ wurde von der Bayerischen Forschungsstiftung (AZ-876-09) gefördert. Die Autoren bedanken sich bei der BMW AG, Werk München, für die freundliche Unterstützung zur Durchführung der Messung in der Messzelle. Unser Dank gilt auch Herrn Gericke und Herrn Meyer von der Firma Steinbichler Optotechnik GmbH, Neubeuern, für die Bereitstellung von Hardware.

7 Literaturverzeichnis

FRAHM, J.-M.; KÖSER, K.; KOCH, R., 2004. Pose Estimation for Multi-camera Systems. In: *Lecture Notes in Computer Science Vol. 3175: Pattern Recognition*, Springer, 2004, pp. 286-293.

GÖTZ, C.; TUTTAS, S.; HOEGNER, L.; EDER, K.; STILLA, U., 2011: Accuracy evaluation for a precise indoor multi-camera pose estimation system. In: Stilla U et al (Eds) PIA11. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (3/W22): 97-102

HANEL, A.; GÖTZ, C.; TUTTAS, S.; BECKER, T.; STILLA, U.; 2013: Orientierungsgenauigkeit eines Multi-Kamerasystems bei Roboterbewegungen in einer Fahrzeugkarosserie. Tagungsband der 33. Wissenschaftlich-Technischen Jahrestagung der DGPF, Band 22

KRAUS, K., 1996. Photogrammetrie – Band 2: Verfeinerte Methoden und Anwendungen. Bonn: Ferd. Dümmlers Verlag

LUHMANN, T., 2009: Precision potential of photogrammetric 6DOF pose estimation with a single camera. ISPRS Journal of Photogrammetry and Remote Sensing, 64(3), pp. 275-284.

MUHLE, D.; ABRAHAM, S.; HEIPKE, C.; WIGGENHAGEN, M., 2008. Automatische Orientierung von zwei gemeinsam bewegten Stereosystemen ohne gegenseitige Korrespondenzen. In: *Photogrammetrie - Laserscanning - Optische 3D-Messtechnik: Beiträge der Oldenburger 3D-Tage 2008*, Wichmann, 2008, pp. 186-193.

Konzeption einer Methode zur quantitativen Erfassung von Bodenerosion durch integrierte Aufnahme mit einer Mikrodrohne und einem terrestrischen Laserscanner

ANETTE ELTNER[1]

Zusammenfassung: Der Bodenabtrag aufgrund oberflächlich abfließenden Niederschlags wird durch komplexe Prozesse bestimmt. Zur non-invasiven Erfassung dieser Vorgänge eignen sich berührungslose Messverfahren. Bisher wurden photogrammetrische Methoden kaum eingesetzt, um Bodenerosion – außerhalb des Labors – hochaufgelöst zu messen. Im Rahmen des vorgestellten Projektes wird auf drei Testflächen in Andalusien und in Sachsen die Bodenoberfläche vor und kurz nach starken Niederschlagsereignissen erfasst. Zur Aufnahme der Geländeoberfläche werden Geländemodelle aus einem mit einer Mikrodrohne (UAV) aufgenommenen Bildverband aus und einer terrestrischen Laserscanneraufnahme verknüpft. Die Befliegung der Testfläche erfolgt mittels einer UAV mit integriertem GPS und IMU sowie einer installierten handelsüblichen Amateurkamera. Neben der Geländeerfassung durch photogrammetrische Triangulation wird ein terrestrischer Laserscanner (TLS) angewandt, der zur Verbesserung der Aufnahmegeometrie auf einem Hochstativ steht. Beide Messmethoden ergänzen sich gegenseitig.
Für einen multi-temporalen Vergleich der gemessenen Bodenoberfläche wird ein stabiles lokales Referenzsystem definiert. Vermarkte Passpunkte und retroreflektierende Zylinder dienen gleichzeitig der Verknüpfung der akquirierten Drohnen- und Laserscannerdaten zueinander sowie zwischen verschiedenen Zeitepochen. Die Vorbereitung der Datenakquise sowie ein Vergleich erster erzeugter Geländemodelle werden vorgestellt.

1 Einleitung

Bodenerosion beschreibt einen durch Wind und Wasser induzierten Abtragungsprozess des Bodens und kann volkswirtschaftlich langfristige Folgen aufgrund der Degradation von landwirtschaftlichen Nutzflächen haben. Die Mergellandschaften Andalusiens besitzen infolge der Bodeneigenschaften und des Niederschlagverhaltens (saisonale Starkregenereignisse) hohe Erosionsraten (u.a. FAUST & SCHMIDT, 2009; BRACKEN & KIRKBY, 2005; POESEN & HOOKE, 1997). Die kurzfristige Reliefveränderung ist als solche meist nicht wahrnehmbar, sollte aber keinesfalls bezüglich ihrer langfristigen Wirkung unterschätzt werden.
Zur Messung der kurzfristigen Ereignisse wird die Bodenoberfläche hochaufgelöst erfasst. Berührungslose photogrammetrische Verfahren werden mit einem mm-genauen Tracerverfahren mittels Radionuklid (WALLING et al., 2009) verknüpft. Seitens der Photogrammmetrie kommt eine UAV, ausgestattet mit einer handelsüblichen Amateurkamera auf einem aktiv stabilisierenden Kamerahalter und ein TLS, installiert auf einem Hochstativ zum Einsatz. UAVs sind im Vergleich zur bemannten Luftbildphotogrammetrie kostengünstig, flexibel und können mittels programmierter Flugmuster die Testflächen wiederholt erfassen. Sie verzeichnen in den letzten Jahren ein stark wachsendes Interesse (u.a. EISENBEIß, 2004; KIRCHGÄSSNER et al., 2010; CRAMER & HAALA, 2012; NEITZEL & KLONOWSKI, 2012; RAU et al., 2012).

UAV und TLS als Varianten zur Geländeerfassung ergänzen sich. Einerseits werden Verdeckungseffekte beim TLS (z.B. infolge tiefer Erosionsrinnen) und Fehlerfortpflanzungen bei großen Bildverbänden mit wenig Passpunkten kompensiert. Andererseits werden die Vorteile beider Methoden – hoher Automatisierungsgrad des TLS und hohe mögliche Genauigkeit bei niedrigen Flughöhen der UAV – genutzt. Zum Vergleich der Geländemodelle generiert einerseits mit UAV- und andererseits mit TLS-Daten, aber besonders zur multi-temporalen Veränderungsdetektion wird ein stabiles lokales Referenznetz aufgebaut, dass die Testflächen geringstmöglich beeinflusst. Die Möglichkeit der Verknüpfung von TLS und UAV wurde bereits demonstriert (EISENBEIß & ZHANG, 2006). In dieser Arbeit wird eine Genauigkeit unter einem Zentimeter angestrebt – realisierbar durch sehr niedrige Flughöhen und kleinräumig gewählte Untersuchungsflächen.
In der Bodenerosionsforschung werden UAVs in Verbindung mit TLS noch nicht eingesetzt. Generell werden berührungslose photogrammetrische Verfahren zur Erosionserfassung überwiegend unter Laborbedingungen angewandt bzw. werden sehr kleine Testparzellen betrachtet (RIEKE-ZAPP & NEARING, 2005; JESTER & KLIK, 2005; HAUBROCK et al., 2009). Während Laserscanner vereinzelt Anwendung finden (HANCOCK et al., 2005; SCHMID et al., 2004), dienen UAV bisher zur Detektion von Bodenerosion größerer Magnituden (D'OLEIRE-OLTMANNS et al., 2012). Diese Arbeit soll das Potenzial der kombinierten Nutzung beider Methoden zur hochaufgelösten Erfassung der Bodenoberfläche aufzeigen.

2 Verwendete Systeme zur Geländeerfassung

Zur hochaufgelösten Erfassung der Bodenoberfläche werden in Andalusien (im Raum Sevilla) zwei rechteckige Testflächen (40 x 50 m²) – Campo 1 und Campo 2 – mit unterschiedlichen Bodeneigenschaften definiert. Campo 2 besitzt im Unterschied zu Campo 1 einen Boden, der deutlich tonhaltiger ist. Neben dem mediterranen Raum wird eine weitere Testfläche (20 x 30 m²) im Mittelsächsischen Lösshügelland (bei Döbeln) festgelegt. Das hier anzutreffende Substrat Löss besitzt vergleichbare Erosionseigenschaften wie der untersuchte Mergel in Spanien, jedoch ist die Niederschlagscharakteristik der Region eine andere. Einzelne Starkregenereignisse sind wesentlich seltener und es herrscht im Gegensatz zum mediterranen Klima ein ganzjährig fallender Niederschlag vor.

2.1 UAV

Zur Geländeerfassung wird mit der Mikrodrohne „AscTec Falcon 8" (Abbildung 1) der Firma Ascending Technologies geflogen. Die UAV ist mit acht Propellern auf zwei Motorschienen ausgestattet. Die Redundanz im Antrieb erhöht die Sicherheit des Flugsystems. Die Flugzeit beträgt bis zu 18 Minuten, kann aber je nach Windverhältnissen deutlich kürzer ausfallen. Die UAV kann bis zu 15 ms^{-1} fliegen. Wird ein programmiertes Flugmuster abgeflogen beträgt die Fluggeschwindigkeit 3 ms^{-1}. Für die autonome Befliegung bestimmter Flugrouten wird ein auf der UAV-Plattform installiertes integriertes GPS- und IMU-System genutzt.
Die UAV ist mit einem aktiv stabilisierenden Kamerahalter versehen. Das bedeutet, neben Aufnahmestörungen infolge von Vibrationen des Flugkörpers, werden Nick- und Rollbewegungen kompensiert. Bei sehr niedrigen Flughöhen ist dies wichtig, da bereits geringe

Verdrehungen um die ω- oder φ-Achse der UAV-Plattform – z.B. aufgrund von Wind – starke Auswirkung auf die Überlappung der Bilder hat. Für die spätere Aerotriangulation des aufgenommenen Bildblocks muss für die Näherungswerte der Orientierung des Projektionszentrums lediglich der κ-Winkel beachtet werden und Verdrehungen um die verbleibenden Achsen betragen näherungsweise Null.
Mit einer Panasonic Lumix DMC-LX 3 erfolgt die Aufnahme der Bilder. Die Sensorgröße ist vom Hersteller mit 1/1.63´´-CCD angegeben. Die Kamera besitzt ein Zoomobjektiv. Die Brennweite beträgt minimal 5.1 mm und maximal 12.8 mm. Im Projekt wird mit der festen Brennweite von 5.1 mm gearbeitet. Das Kleinbildäquivalent ist 24 mm bzw. 50 mm. Bei einem Seitenverhältnis von 3648 zu 2736 Pixeln lässt sich somit eine Sensorgröße von 7.36 x 5.52 mm² mit einer Pixelgröße von 2 µm berechnen. Die Kamera besitzt einen optischen Bildstabilisator, der jedoch während der Befliegung deaktiviert wird, um ein stabiles Verhältnis der Linsen zum Sensor zur gewähren.



Abb. 1: UAV „Falcon 8“; TLS auf dem Hochstativ

2.2 TLS

Neben der UAV wird das Gelände mit einem terrestrischen System erfasst. Der TLS Riegl LMS Z420i ist ein Panoramascanner, der das Impulslaufzeitverfahren verwendet und eine Reichweite von bis zu einem Kilometer besitzt. Die Messgenauigkeit liegt unter einem Zentimeter. Die Testflächen werden von allen vier Seiten gescannt, um Schatteneffekte weitestgehend zu vermeiden. Die Winkelauflösung wird mit 0.024° gewählt, damit bei einer Erfassung von den vier Seiten ein theoretischer Punktabstand von einem Zentimeter im näheren Bereich des Scanfelds (Entfernung < 25 m) und ca. 2 cm am gegenüberliegenden Testfeldrand gewährt ist. Der Scanner wird auf ein Hochstativ in einer Höhe von etwa 4 m installiert (Abbildung 1), denn für diese Arbeit werden relativ flache Hänge erfasst. Durch den erhöhten Blickwinkel soll die Geometrie des Schnitts des Laserspots mit dem zu erfassenden Objekt verbessert und somit Ungenauigkeiten, infolge unzureichender Reflexion oder falscher Distanzwerte durch Mehrfachreflexion (BOEHLER et al., 2003) verringert werden.

2.3 Definition eines Referenzsystems

Eine wesentliche Zielstellung des Projektes ist es, die Geländeveränderungen multi-temporal hochaufgelöst zu betrachten. Dafür muss ein stabiles lokales Referenzsystem errichtet werden, dass zusätzlich die Verknüpfung der UAV- mit den TLS-Daten erlaubt. Das Referenzsystem wird mittels 60 bzw. 100 cm langer Vermarkungsstangen, die bis zum Anschlag in den Boden eingebracht werden, realisiert. Sie werden im regelmäßigen Abstand um die Testflächen platziert. Da die Testflächen weitestgehend ungestört bleiben sollen, wird in der Fläche nur ein, das Feld halbierender, schmaler Pfad genutzt, um zusätzliche Höhenpasspunkte zu setzen. Auf die Vermarkungsrohre werden stabile selbstgebaute passgenaue Zielmarken (Abbildung 2) gesteckt, die nach der Messkampagne wieder entfernt werden. Bei der Aerotriangulation finden Holzplatten mit weißen Kreisen auf schwarzem Hintergrund ihre Anwendung. Für die Registrierung der Laserscans kommen retroreflektierende Zylinder zum Einsatz. Auf die Zylinder werden zusätzlich ebenfalls weiße Kreise mit einem schwarzen Kontrastring installiert. Somit können die Zylinder für die Aerotriangulation und die Laserscanregistrierung gemeinsam genutzt werden. Eine Veränderung der Positionen der Vermarkungsrohre ist nicht gänzlich auszuschließen, besonders für die tonige Testfläche Campo 2, auf der stärkere Quell- und Schrumpfungsprozesse und somit Bodenveränderung bis in größere Tiefen erwartet werden. Daher werden zur langfristigen Kontrolle die Passpunkte zusätzlich tachymetrisch in ein weiteres übergeordnetes lokales Netz, markiert mit Betonnägeln auf betonierten Flächen in der näheren Umgebung, eingemessen.

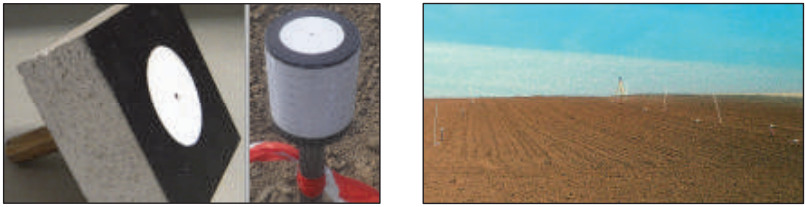


Abb. 2: Zielmarken für die Orientierung im lokalen Referenzsystem; Ausschnitt des Testfelds in Sachsen

3 Datenakquise und -verarbeitung

Vor dem eigentlichen Einsatz der UAV erfolgt eine Flugplanung mittels spezieller Software des Drohnenherstellers. Bei den Testflächen in Spanien beträgt die Flughöhe 10 m und in Sachsen 8 m. Längs- und Querüberdeckung ergeben jeweils 80% und 60%. In Tabelle 1 sind die sich daraus ergebenden Bildverbandparameter dargestellt. Die Bildblöcke müssen in allen Fällen aufgrund der begrenzten Akkuleistung unterbrochen und mit neuem Akku fortgesetzt werden. Letzte Einstellungen der Flugmatrix – absolute Lage des ersten Aufnahmepunktes und Orientierung des Bildblocks – müssen vor Ort erfolgen, da Kartenmaterial mit ausreichendem Maßstab für die kleinen Testflächen für die Arbeit nicht gegeben ist. Die Erfassung von Campo 1 erfolgte unter erschwerten Bedingungen. Aufgrund stärkeren Windes sind häufig unscharfe Bilder im Bildblock enthalten und einzelne Flugstreifen stark zueinander verdreht.

Tabelle 1: Parameter der Flugplanung

	Brenn-weite [mm]	Maß-stabszahl	Bodenauf-lösung [mm]	Bildseite a (Gelände) [m]	Bildseite b (Gelände) [m]	Basis [m]	Streifen-abstand [m]	Bilderan-zahl pro Streifen	Streifen-anzahl
Spanien	5.1	1961	3.96	14.42	10.82	2.16	5.77	25	7
Sachsen	5.1	1569	3.17	11.54	8.65	1.73	4.62	19	5

Bei der verwendeten Amateurkamera ist von einer instabilen inneren Genauigkeit auszugehen. Um die Reproduzierbarkeit der Kameraparameter zu untersuchen, wird sie vor dem Einsatz im Gelände mehrmals kalibriert. Zusätzlich wird die Kamera direkt vor und nach jedem Flug mittels eines mobilen Kalibrierfelds kalibriert (Abbildung 3). Die Werte der Verzeichnung sind zwar hoch, bleiben aber zwischen den verschiedenen Aufnahmevarianten nahezu konstant. Die Veränderungen der Lage des Bildhauptpunktes sowie der Kamerakonstante vor und nach jedem Flug besitzen hingegen einen wesentlich stärkeren Einfluss auf die Genauigkeit der Objektpunkte.

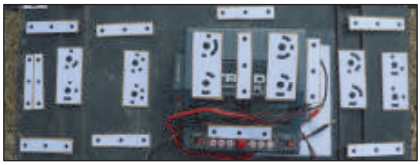


Abb. 3: mobiles Kalibrierfeld

Die Orientierung der Bildblöcke erfolgt mit LPS. Aufgrund der stabilisierenden Kamerahalterung, die eine Nadiraufnahme und Einhaltung der geplanten Überlappung erleichtert, und der autonomen Ansteuerung der Bildaufnahmepositionen kann eine Aerotriangulationskette erfolgen, wie sie bei Aufnahmen mit Flugzeugen üblich ist. Die Bündelblockausgleichung (BBA) erfolgt einstufig. Als Näherungswerte für die Projektionszentren werden die vom System aufgezeichneten GPS-Lagekoordinaten verwendet. Die Genauigkeit der GPS-Höhenkoordinate – Schwankungen im Meterbereich – ist bei der niedrigen Flughöhe nicht ausreichend. Ersatzweise kommen die barometrisch gemessenen Höhen zur Anwendung. Es wird mit entzeichneten Bildern gearbeitet. Zur automatischen Verknüpfungspunktsuche verwendet LPS die Kreuzkorrelation zur Grobbestimmung und anschließend LSM (Least-Square-Matching) zur sub-pixelgenauen Messung der Bildkoordinate. Mit den Originalbildern ist, aufgrund der starken Verzeichnung sowie der begrenzten Genauigkeit des GPS/IMU, besonders in den oberen Hanglagen bei abnehmender Überlappung die Punktzuordnung v.a. zwischen Bildern verschiedener Streifen nicht erfolgreich. Die Bildkoordinaten der Passpunkte werden mit einer iterativen sub-pixelgenauen Anpassung von Ellipsen – umgesetzt durch hauseigene Software – automatisch ermittelt. Das Ergebnis der Ausgleichung ist für alle drei Testflächen in Tabelle 2 dargestellt. Die Gesamtgenauigkeit liegt zwischen 0.5 und 1 Pixel, d.h. zwischen 1.6 und 4 mm im Objektraum.

Tab. 2: Statistik der Residuen der Verbesserungen der Passpunkte nach der BBA

	Lüttewitz (s0 = 0.55 Pixel)			Campo 1 (s0 = 1.06 Pixel)			Campo 2 (s0 = 0.49 Pixel)		
	v_X[mm]	v_Y [mm]	v_Z [mm]	v_X [mm]	v_Y [mm]	v_Z [mm]	v_X [mm]	v_Y [mm]	v_Z [mm]
Mittelwert	-0.5	0	-0.4	0.3	-0.3	0.2	0	-0.1	-0.2
Stabw.	1.7	1.3	1.5	1.6	1.4	2.3	1.5	1.5	4

Die Referenzierung der einzelnen Scanpositionen der Laserscans wird im ersten Schritt mittels der in die Vermarkungsrohre eingefügten Zielmarken sowie zusätzlicher temporär installierter Zylinder vorgenommen. Danach erfolgt die Übertragung der vereinigten Punktwolke in das lokale Referenzsystem mit den gesteckten Zylindern mittels einer weiteren Helmert-Transformation. Die Genauigkeiten liegen zwischen 4 und 6 mm. Die Referenzkoordinaten der gesteckten Zylinder werden am Mittelpunkt der Deckfläche gemessen. Da die automatische Referenzierung der Scanpositionen mit den Zylindermittelpunkten geschieht, müssen die Referenzkoordinaten der Zylinder um die Zylinderhöhe reduziert werden. Im Gelände ist kein lotrechtes Einschlagen aller Vermarkungsrohre gewährleistet. Während die Auswirkung der alleinigen Reduzierung der Höhe für die Koordinatenübertragung bei schrägen Zylindern für die Höhe gering ist, ist sie für die Lage umso bedeutender.

4 Vergleich der UAV- und TLS-Daten

Mit den Werten der errechneten äußeren Orientierung, der kalibrierten inneren Orientierung und den entzeichneten Bildern wird in LPS ein erstes Geländemodell erstellt. Die Punktdichte beträgt durchschnittlich 1 Punkt/15 cm² und ist wesentlich niedriger als beim TLS. Innerhalb dieser Präsentation wird primär die Performance des triangulierten Blocks zur Punktwolke des TLS über die gesamte Testfeldfläche betrachtet. Dafür sind die gegebenen Punktdichten ausreichend. Dennoch muss für die weitere hochaufgelöste Geländemodellierung mit weiteren Matching-Algorithmen (z.B. HIRSCHMÜLLER et al., 2008) verfahren werden.
Der Vergleich der TLS- mit den UAV-Daten wird mit der Open-Source Software CloudCompare realisiert (Abbildung 5). Dabei werden die Distanzen in Z-Richtung zwischen der zu testenden Punktwolke (LPS) und der Referenzpunktwolke (TLS) nach der Chamfer-Methode berechnet (HEZEL et al., 2002). Bei der kleinen Testfläche in Sachsen liegen die Abweichungen innerhalb des Genauigkeitsbereichs des TLS. In Spanien werden die Differenzen jedoch größer. Verschiedene Ursachen müssen aufgelistet werden. Die Kamerakalibrierungen vor und nach dem Flug zeigen wesentlich größere Unterschiede für die Testflächen in Spanien im Vergleich zu Sachsen. Außerdem sind die Flughöhen in Spanien größer und somit wirkt sich die instabile Kamera zudem stärker auf das geometrische Modell aus. Des Weiteren sind die Testflächen Campo 1 und Campo 2 größer und die gesetzten Höhenpasspunkte im mittleren Streifen des Feldes sind anscheinend nicht ausreichend, um den Bildblock zu stabilisieren. Höhen aus LPS werden in der Mitte der leicht konvexen Hänge unterschätzt. Bei der nächsten Messkampagne werden weitere Höhenpasspunkte ausgelegt und zusätzliche Querstreifen geflogen. Einen weiteren Einfluss auf die unterschiedlichen Genauigkeiten in Spanien und Sachsen hat die auf Campo 1 und 2 vorhandene Vegetation. Nur die östlichen Feldhälften wurden geeggt, während auf den Westlichen Getreidestoppeln stehen blieben. Auch hier sind die Höhen der TLS-Punkte aufgrund unzureichender Punktzuordnung bei den bewachsenen Flächen wesentlich größer.

Abschließend ist für Campo 1 festzustellen, dass sich die erschwerten Aufnahmebedingungen in den Genauigkeiten des Geländemodells – generiert aus den UAV-Daten – widerspiegeln. Dennoch sind die Differenzen zwischen TLS und UAV auf den vegetationsfreien Flächen in Spanien vielversprechend.

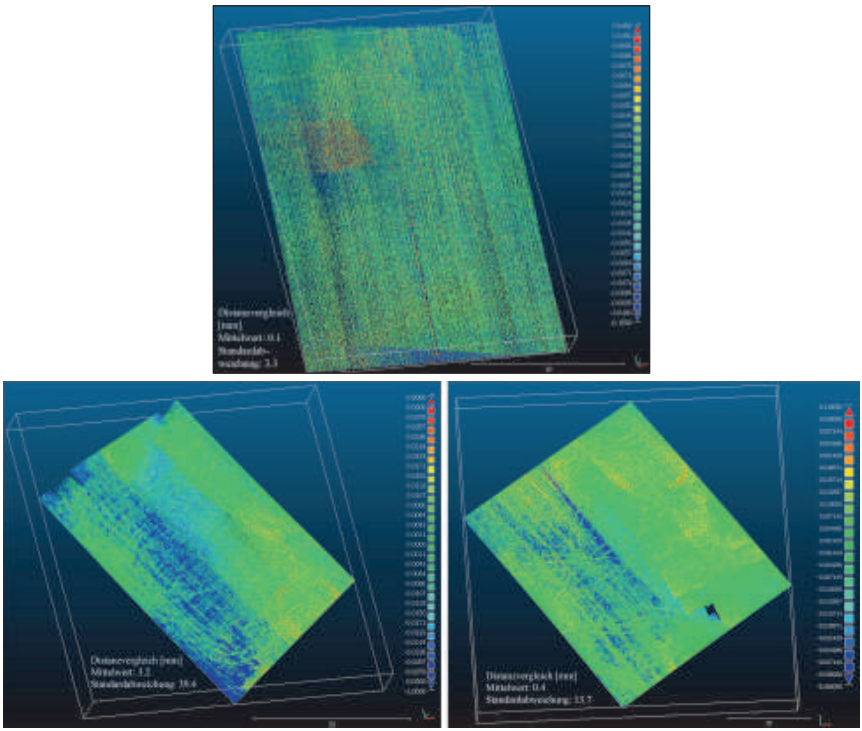


Abb. 5: Differenz Punkte der Geländemodelle erzeugt mit LPS und TLS für die Testflächen in Sachsen (oben) und Spanien (unten links Campo 1, unten rechts Campo 2)

5 Ausblick

Ein erster Vergleich der UAV und TLS-Daten konnte demonstrieren, dass es möglich ist mit dem definierten lokalen Referenzsystem Genauigkeiten unter einem Zentimeter zu erreichen. In wieweit die hochaufgelöste Geländeerfassung zur Bodenerosionsmodellierung beitragen kann, werden weitere Messkampagnen mit multi-temporalen Vergleichsdaten zeigen.
Die Verwendung weiterer Software ist nötig. Neben der einstufigen BBA wird in einem nächsten Schritt zweistufig mit verschiedener Open-Source Software (Beispiele siehe KERSTEN et al., 2012) gearbeitet, in denen z.B. der SIFT-Operator (LOWE, 1999) implementiert ist. Eine deutlich

höhere Verknüpfungspunktdichte ermöglicht die Selbstkalibrierung und Verwendung der verzeichneten Originalbilder.

6 Literaturverzeichnis

BOEHLER, W.; BORDAS VICENT, M. & MARBS, M. (2003): Investigating Laser Scanner accuracy. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV – 5/C15, 696-701.

BRACKEN, L. J. & KIRKBY, M. J. (2005): Differences in hillslope runoff and sediment transport rates within two semi-arid catchments in southeast Spain. Geomorphology, **68** (3-4), 183-200.

CRAMER, M. & HAALA, N. (2012): Genauigkeitspotential der photogrammetrischen Bildauswertung für Daten unbemannter Luftfahrzeuge. DGPF Tagungsband, **21**, 428 - 440.

D'OLEIRE-OLTMANNS, S.; MARZOLFF, I.; PETER, K. D. & RIES, J. B. (2012): Unmanned Aerial Vehicle (UAV) for Monitoring Soil Erosion in Morocco. Remote Sensing, **4**, 3390-3416.

EISENBEIß, H. (2004): A mini Unmanned Aerial Vehical (UAV): System Overview and Image Acquisition. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36 – 5/W1.

EISENBEIß, H. & ZHANG, L. (2006): Comparison of DSMs generated from Mini UAV Imagery and Terrestrial Laser Scanner in a Cultural Heritage Application. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36 – 5, 90-96.

FAUST, D. & SCHMIDT, M. (2009): Soil erosion processes and sediment fluxes in a Mediterranean marl landscape, Campiña de Cádiz, SW Spain. Z. Geomorphologie N.F., **52** (2), 247-265.

HAUBROCK, S.-N.; KUHNERT, M.; CHABRILLAT, S.; GÜNTNER, A. & KAUFMANN, H. (2009): Spatiotemporal variations of soil surface roughness from in-situ laser scanning. Catena, 79, 128 - 139.

HANCOCK, G.R.; CRAWTER, D.; FITYUS, S.G.; CHANDLER, J. & WELLS, T. (2008): The measurement and modelling of rill erosion at angle of repose slopes in mine spoil. Earth Surf. Process. Landf., **33**, 1006 - 1020.

HEZEL, S.; KUGEL, A.; MÄNNER, R. & GAVRILA, D. M. (2002): FPGA-based Template Matching using Distance Transforms. Proc. of the IEEE Symposium on Field-Programmable Costum Computing Machines, 89 - 97.

HIRSCHMÜLLER, H. (2005): Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. IEEE Conference on Computer Vision and Pattern Recoginition, June 2005, San Diego, CA, USA, Volume 2, 807 - 814.

JESTER, W. & KLIK, A. (2005): Soil surface roughness measurement-methods, applicability, and surface representation. Catena, **64**, 174 - 192.

KERSTEN, T.; LINDSTAEDT, M.; MECHELKE, K. & ZOBEL, K. (2012): Automatische 3D-Objektrkonstruktion aus unstrukturierten digitalen Bilddaten für Anwendungen in Architektur, Denkmalpflege und Archäologie. DGPF Tagungsband, **21**, 137 - 148.

KIRCHGÄSSNER, U.; PUTZE, U.; SCHÖNERMARK, M.V. & HAALA, N. (2010): Anforderungen an die Auswertung UAV-gestützter Fernerkundungsdaten. DGPF Tagungsband, **19**, 597 - 605.

LOWE, D. G. (1999): Object recognition from local scale-invariant features. Proceedings of the International Conference on Computer Vision, **2**, 1150 - 1157.

NEITZEL, F. & KLONOWSKI, J. (2011): Mobile 3D mapping with a low-cost UAV system. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38 – 1/C22.

POESEN, J. & HOOKE, J. M. (1997): Erosion, flooding and channel management in Mediterranean Environments of southern Europe. Progress in Phys. Geogr. **21** (2), 157 - 199.

RAU, J.; JHAN, J. P.; LO, C. F. & LIN, Y. S. (2012): Landslide Mapping Using Imager Acquired by a fixed-wing UAV. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38 – 1/C22.

RIEKE-ZAPP, D. & NEARING, M. (2005): Digital Close Range Photogrammetry for Measurement of Soil Erosion. The Photogrammetric Record, **20** (109), 69 - 87.

SCHMID, T.; SCHACK-KIRCHNER, H. & HILDEBRAND, E. (2004): A case study of terrestrial laser scanning in erosion research: calculation of roughness and volume balance at a logged forest site. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36 – 8/W2, 114 - 118.

WALLING, D. E.; SCHULLER, P.; ZHANG, Y. & IROUMÉ, A. (2009): Extending the timescale for using beryllium 7 measurements to document soil redistribution by erosion. Water Resources Research, **45**.

Orientierungsgenauigkeit eines Multi-Kamerasystems bei Roboterbewegungen in einer Fahrzeugkarosserie

ALEXANDER HANEL[1], CARSTEN GÖTZ[1], SEBASTIAN TUTTAS[1], TOBIAS BECKER[2] & UWE STILLA[1]

Zusammenfassung: Eine hochgenaue Positionsbestimmung eines Multi-Kamerasystems über Rückwärtsschnitte ist nur in Bereichen mit vielen Referenzpunkten möglich. Liegen nur wenige Referenzpunkte vor, können aus bereits orientierten Bildern Neupunkte über Vorwärtsschnitte eingemessen werden. In derartig schwach besetzen Bereichen werden alle bis zu diesem Zeitpunkt eingemessenen Punkte wiederum genutzt, um weitere Aufnahmen zu orientieren. Diese Strategie wird nun abwechselnd eingesetzt, um die Orientierung aller entlang einer Trajektorie aufgenommenen Bilder zu bestimmen. Im folgenden Beitrag wird der Ansatz in einer simulierten Umgebung und unter realen Messbedingungen untersucht. Besondere Aufmerksamkeit ist dabei der Fortpflanzung der nach der vermittelnden Ausgleichung resultierenden Genauigkeit gewidmet. Bei den Versuchen im industriellen Umfeld konnten meistens ausreichend Neupunkte eingemessen werden, um die Trajektorie zu bestimmen. Es hat sich gezeigt, dass der Fehler der Orientierungen im Bereich der eingemessenen Neupunkte gegenüber dem Bereich mit dichtem Referenzpunktfeld nicht übermäßig angestiegen ist.

1 Einführung

1.1 Motivation

Für Geometrie- und Oberflächenmessungen werden Inspektionssensoren mit einem Roboter nicht nur im Außenraum um eine Fahrzeugkarosserie bewegt, sondern auch in deren Innenraum geführt (Abb. 1). Für manche Anwendungen ist die Kenntnis deren Position und Orientierung aus den Einstellungen des Roboters jedoch nicht genau genug. Deswegen sollen diese mit einem zusätzlichen Multikamerasystem hochgenau bestimmt werden.

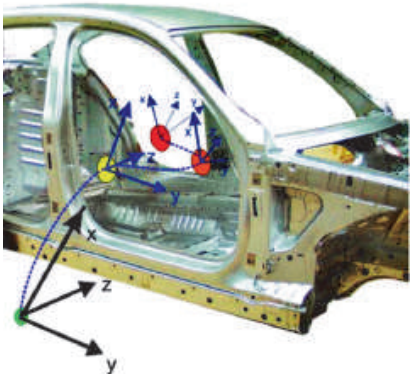


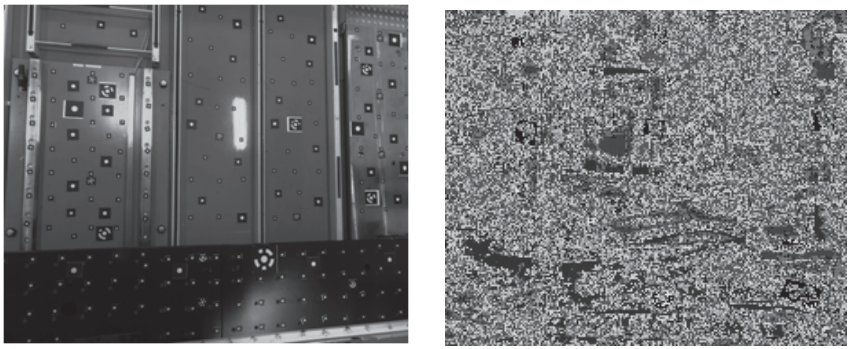
Abb. 1: Beispiel für die Messsituation im Außen- und Innenbereich mit vier Kamerapositionen.

Bei einer photogrammetrischen Vorgehensweise zur Bestimmung der äußeren Orientierung aufgrund von Referenzmarken spielen deren Sichtbarkeit und Verteilung eine wichtige Rolle für die erreichbaren Genauigkeiten und Zuverlässigkeiten. So kann man für den Außenraum um das Messobjekt von einer

1) Photogrammetrie & Fernerkundung, Technische Universität München, Arcisstraße 21, 80333 München, http://www.pf.bv.tum.de
2) BMW AG, Petuelring 130, 80788 München

sehr hohen Anzahl an a priori eingemessenen Passpunkten mit guter geometrischer Verteilung ausgehen. Als Beispiel ist in Abbildung 2a der Boden einer Messumgebung mit mehreren kodierten und unkodierten Referenzmarken unterschiedlicher Größe zu sehen. In solchen Bereichen ergeben sich typischerweise nur kleine Fehlerellipsen (in Abb. 1 durch grüne Ellipse angedeutet).

Bewegt man sich mit dem Sensorsystem hingegen in die Karosse hinein, ist die Messsituation deutlich ungünstiger. Aufgrund von Einschränkungen bei der Anzahl und Platzierung von Marken, erreicht man schlechtere Ergebnisse der Orientierungsbestimmung (in Abb. 1 durch rote Fehlerellipse angedeutet). Abbildung 2b zeigt den Innenbereich einer Fahrzeugkarosse, in der im Sichtbereich der Kamera nur fünf Marken zu sehen sind.



a b

Abb. 2: Bildaufnahmen durch das Kamerasystem. a) Messumgebung mit eingemessenen Passpunkten auf der Bodenplatte, b) Fahrzeuginnenraum mit markierten Verknüpfungspunkten

Sofern derartige schwach besetzte Bereiche entlang der untersuchten Trajektorie vorliegen, kann nicht mehr für alle Kamerapositionen garantiert werden, dass die Orientierung über Rückwärtsschnitte berechnet werden kann. Da die Aufnahmen zur Bildung einer Trajektorie verknüpft werden müssen, ist somit auch deren lückenlose Bestimmung nicht mehr gewährleistet. Dies birgt schließlich die Gefahr, dass für die Messungen der Inspektionssensoren die Positionen nicht mehr bestimmt werden können. Daher wird in diesem Beitrag ein erweiterter Ansatz zur Bestimmung der Orientierung untersucht.

1.2 Verwandte Arbeiten

Grundlegende Untersuchungen zum Rückwärtsschnitt befinden sich in WROBEL (2001). Insbesondere wird im dortigen Beitrag das Prinzip der projektiven und perspektivischen Kollineation eingeführt, wobei letzteres im Rahmen dieses Beitrags herangezogen wird. Des Weiteren werden dort die verschiedenen Standard-Orientierungsaufgaben voneinander abgegrenzt.
Verschiedene Herangehensweisen für industrielle Anwendungsbeispiele der Photogrammetrie schildert FRASER (1988), wobei im Gegensatz zu diesem Beitrag dort auch Methoden der analytischen Photogrammetrie eingeflossen sind. Dahingegen geht SCHEWE (1987) konkret auf die Anwendung der Karosserievermessung ein und verwendet hierfür mit InduSURF ein photogram-

metrisches Messsystem. In Abgrenzung hierzu wird in diesem Beitrag kein marktreifes System verwendet, sondern die Untersuchungen auf elementare Hardware und selbstentwickelte Ansätze gestützt.
Basisinformationen über den Aufbau eines Multi-Kamerasystems werden behandelt in GÖTZ et al. (2011). Im dortigen Beitrag wird insbesondere ein System aus drei Kameras detailliert evaluiert. Dessen Erweiterung um eine zusätzliche Kamera findet Anwendung im praktischen Versuch, der im weiteren Verlauf dieses Beitrags beschrieben wird.
Verwandt mit der hier behandelten Fragestellung ist das Problem eines Roboters, gleichzeitig seine Orientierung zu bestimmen und die Umgebung zu beschreiben. Grundlagenarbeit für das als SLAM (Simultaneous Localization And Mapping) bezeichnete Verfahren haben beispielsweise FACCHINETTI et al. (1995) und WELLS et al. (1996) geleistet. Eine Abgrenzung zur hier behandelten Fragestellung ergibt sich aus der häufigen Einbindung von Zusatzsensoren wie zum Beispiel einer IMU in das Messsystem, während hier nur ein Kamerasystem zum Einsatz kommt. Gemeinsam ist jedoch das Ziel, die eigene Position und nicht die eines abgebildeten Objektes zu bestimmen. Im Gegensatz zu DARR et al. (2013) wird im vorliegenden Beitrag das Genauigkeitsverhalten nicht nur im Bereich mit vielen Referenzpunkten untersucht, sondern der Übergang von einem derartigen Bereich hin zu einem mit weniger sichtbaren Referenzpunkten betrachtet.

2 Orientierungsverfahren im sichtbeschränkten Raum

Bestimmung der äußeren Orientierung des Kamerasystems über Rückwärtsschnitte und anschließendes Einmessen von Neupunkten über Vorwärtsschnitte bildet die Kernoperationen des in diesem Beitrag erarbeiteten Ansatzes. Mit Hilfe nachstehender Grafik (Abb. 3) und den zugehörigen textuellen Erläuterungen wird dieser im Folgenden vorgestellt.

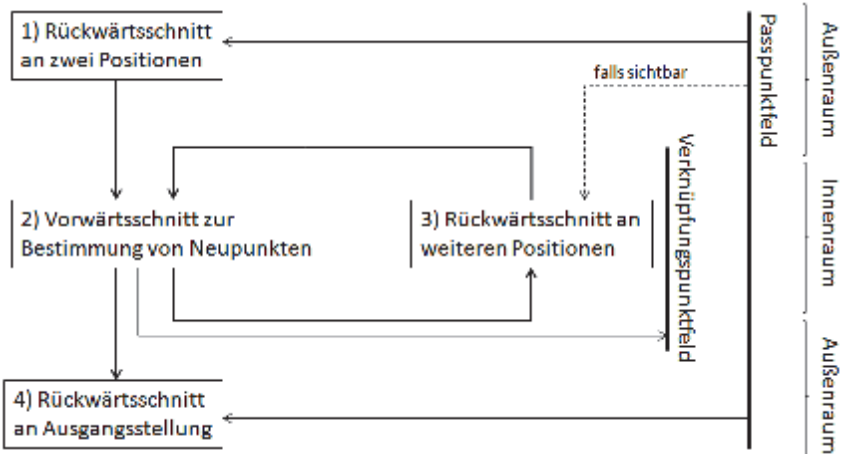


Abb. 3: Schematische Darstellung des Ablaufs der Berechnungen

2.1 Schritt 1 – Positionierung

Zu Beginn des Messpfades (Abb. 3 - 1) liegen Bereiche mit einer hohen Zahl $i = 1...n; n > 3$ von Passpunkten vor, die zu einer Überbestimmung der Orientierung mit hoher Redundanz genutzt werden. Damit ist es dort möglich, Rückwärtsschnitte durch eine Ausgleichung (Gl. 1-3) nach vermittelnden Beobachtungen mit nachfolgendem Funktionalmodell auszuführen. r_{jk} stellt darin die Elemente der Rotationsmatrix dar. In der Darstellung der Gleichungen wird die Notation nach KRAUS (2004) verwendet.

$$\hat{f}_{x_i^{'}} : x_i^{'} + \hat{v}_{x_i^{'}} = x_0^{'} + c^{'} \cdot \frac{\hat{r}_{11}(\hat{X}_i - \hat{X}_0) + \hat{r}_{21}(\hat{Y}_i - \hat{Y}_0) + \hat{r}_{31}(\hat{Z}_i - \hat{Z}_0)}{\hat{r}_{13}(\hat{X}_i - \hat{X}_0) + \hat{r}_{23}(\hat{Y}_i - \hat{Y}_0) + \hat{r}_{33}(\hat{Z}_i - \hat{Z}_0)} \tag{1}$$

$$\hat{f}_{y_i^{'}} : y_i^{'} + \hat{v}_{y_i^{'}} = y_0^{'} + c^{'} \cdot \frac{\hat{r}_{12}(\hat{X}_i - \hat{X}_0) + \hat{r}_{22}(\hat{Y}_i - \hat{Y}_0) + \hat{r}_{32}(\hat{Z}_i - \hat{Z}_0)}{\hat{r}_{13}(\hat{X}_i - \hat{X}_0) + \hat{r}_{23}(\hat{Y}_i - \hat{Y}_0) + \hat{r}_{33}(\hat{Z}_i - \hat{Z}_0)} \tag{2}$$

$$\hat{f}_{X_i} : X_i + \hat{v}_{X_i} = \hat{X}_i \qquad \hat{f}_{Y_i} : Y_i + \hat{v}_{Y_i} = \hat{Y}_i \qquad \hat{f}_{Z_i} : Z_i + \hat{v}_{Z_i} = \hat{Z}_i \tag{3}$$

Das Resultat dieser Berechnungen sind die sechs Parameter der geschätzten äußeren Orientierung ($\hat{X}_0, \hat{Y}_0, \hat{Z}_0, \hat{\omega}, \hat{\varphi}, \hat{\kappa}$) einschließlich deren Genauigkeiten ($\hat{\sigma}_{\hat{X}_0}, \hat{\sigma}_{\hat{Y}_0}, \hat{\sigma}_{\hat{Z}_0}, \hat{\sigma}_{\hat{\omega}}, \hat{\sigma}_{\hat{\varphi}}, \hat{\sigma}_{\hat{\kappa}}$). Bei der Ausgleichung werden durch die Gleichungen (3) zwar auch für die Passpunkte Verbesserungen geschätzt, die vorher gemessenen Koordinaten werden jedoch nicht ersetzt. Der Grund für diese Vorgehensweise liegt darin, dass die Weltkoordinaten zwar als fehlerbehaftete Beobachtungen in die Ausgleichung eingeführt werden sollen, der vorab durchgeführte Bestimmung jedoch als genauer und zuverlässiger angesehen werden muss.

Die Bestimmung von Näherungswerten für die Orientierungsparameter kann beispielsweise über eine Direkte Lineare Transformation LUHMANN (2010) oder anhand der eingestellten Roboterposition abgeleitet werden.

Der so beschriebene Teilschritt muss für mindestens zwei sich überlappende Aufnahmen erfolgreich durchgeführt werden, wie im Folgenden gezeigt wird.

2.2 Schritt 2 – Einmessen der Verknüpfungspunkte

Sobald sich das Messsystem in Bereichen mit nur wenigen oder gar keinen sichtbaren Referenzpunkten mit bekannten Koordinaten bewegt, sinkt die Redundanz und die Möglichkeit zur Berechnung von Rückwärtsschnitten wird eingeschränkt oder gar unmöglich.

Für solche Situationen ist es notwendig, anhand der beiden zuvor orientierten Aufnahmen aus Schritt 1 Neupunkte durch Vorwärtsschnitte zu bestimmen. Der Vorwärtsschnitt wird mit derselben mathematischen Basis wie der Rückwärtsschnitt durch vermittelnde Ausgleichung gelöst. Dabei treten die Gleichungen (1) und (2) dupliziert für jede der beiden involvierten Kameras auf. Jedoch werden die Parameter der äußeren Orientierung nun nicht mehr als Unbekannte mitgeschätzt. Zu beachten ist, dass die verwendeten Bildpunkte beider Kameras denselben Objektpunkt darstellen, was durch Verwenden codierter Marken sichergestellt werden kann.

Näherungswerte für die Koordinaten des Neupunktes erhält man beispielsweise durch Definition zweier Schnittgeraden mit dem Projektionszentrum der jeweiligen Kameras als Aufpunkt und

dem Vektor vom Projektionszentrum zu den Bildkoordinaten als Richtungsinformation. Die Koordinaten des Mittelpunktes der kürzesten Parallaxe zwischen den Geraden bilden den genäherten Wert. Eine Fehlerfortpflanzung nach NIEMEIER (2008) liefert die zugehörigen Gewichte.
Für die resultierenden dreidimensional eingemessenen Punkte wird anhand eines 3D-Punktfehlers nach NIEMEIER (2008) entschieden, ob die eventuell vorhandenen Werte aus vorangegangenen Vorwärtsschnitten überschrieben oder beibehalten werden sollen. Für die erstmalige Berechnung der Neupunkte wird dazu ein hoher Schwellwert des 3D-Punktfehlers manuell angesetzt. Dieser muss so gewählt werden, dass einerseits nur wenige der nur spärlich vorhandenen Punkte ausgeschlossen werden, und anderseits zu schlechte Punkte, die die Orientierbarkeit eventuell gefährden, unberücksichtigt bleiben.

2.3 Schritte 3 und 4 – Iteration

Die in Schritt 2 eingemessenen Neupunkte und die a priori vorhandenen Passpunkte können an einer weiteren Kameraposition genutzt werden, um dort über Rückwärtsschnitt die Aufnahme zu orientieren. Nach diesem Prinzip können weiterführend Vorwärtsschnitte und Rückwärtsschnitte nach oben erläuterter Vorgehensweise berechnet werden. Dies kann solange wiederholt werden, wie neue Bilder aufgenommen werden. Treten Kamerapositionen auf, bei denen die Orientierung nicht bestimmt werden kann, werden diese übersprungen und versucht, mit den darauffolgenden Kamerapositionen fortzufahren. Voraussetzung an diesen ist die Sichtbarkeit von Referenzpunkten.

3 Evaluierung der Methode anhand simulierter und realer Messaufgaben

3.1 Simulation

Vor der Durchführung der praktischen Messversuche werden Simulationen berechnet, um mehrere Aufnahmesituationen zu untersuchen und den erarbeiteten Ansatz zu evaluieren. Bei der Wahl der Aufnahmesituationen werden charakteristische Bedingungen wie beim Eintauchen in die Karosserie berücksichtigt (siehe Abschnitt 1.1). Dabei wird von nur partieller Punktsichtbarkeit und beschränkten Bewegungsmöglichkeiten ausgegangen. Für die Simulation wird statt dem Kamerasystem nur eine Einzelkamera betrachtet.

Die Berechnungsergebnisse sind in Abb. 4 gezeigt. Dargestellt sind sowohl die Standardabweichungen in den einzelnen drei Komponenten $\hat{\sigma}_{\dot{X}_0}, \hat{\sigma}_{\dot{Y}_0}, \hat{\sigma}_{\dot{Z}_0}$ beziehungsweise $\hat{\sigma}_{\dot{\omega}}, \hat{\sigma}_{\dot{\phi}}, \hat{\sigma}_{\dot{\kappa}}$ als auch der 3D-Punktfehler nach NIEMEIER (2008) als skalares Maß für die Positionsgenauigkeit. Dessen Äquivalent für die Rotationsgenauigkeit ist eine Transformation der Rotationsparameter in eine skalare Drehung um eine Raumachse nach FISCHER (2008) und eine anschließende Fehlerfortpflanzung.

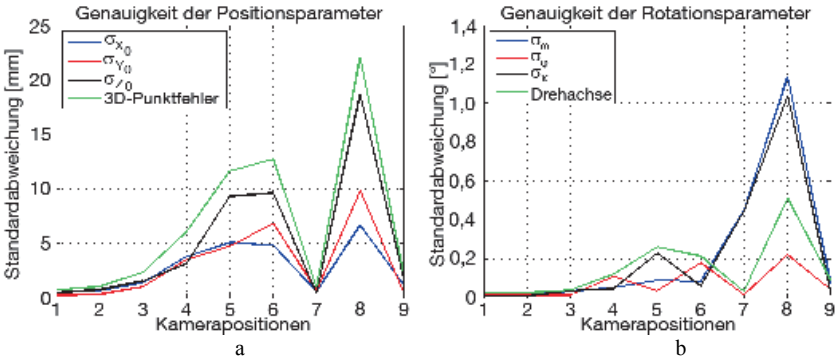


Abb. 4: Genauigkeiten der äußeren Orientierung der Kamera an einem simulierten Kamerapfad. a) Positionsparameter, dargestellt sind die einzelnen Komponenten und ein 3D-Punktfehler nach NIEMEIER (2008), b) Rotationsparameter, dargestellt sind die einzelnen Komponenten und ein skalarer Fehler nach FISCHER (2008)

Es zeigt sich sowohl für die Genauigkeit der Positions- als auch Rotationsparameter zunächst der erwartete Anstieg der Kurven. Die Begründung dafür liegt in der abnehmenden Referenzpunktzahl beim Eintauchen in die Karosserie an Position 3. Ab Mitte des Pfades jedoch steigen die Kurven der Positionsgenauigkeit (Abb. 4a) in stärkerem Maße als zuvor, bevor diese an Kameraposition 7 auf ein Niveau ähnlich der Ausgangssituation abfallen. Dieser Effekt ist darauf zurückzuführen, dass an dieser Position nur Neupunkte einer einzigen Vorwärtsschnitt-Ausgleichung zur Orientierung verwendet werden. Diese sind somit in sich homogen, woraus die geringe, aber nicht-kontrollierte Standardabweichung resultiert. Testweises Hinzufügen eines unabhängig davon bestimmten Passpunktes ergibt Werte, die dem bisherigen Verlauf der Kurven entsprechen und somit die obige Vermutung bestärken. An der darauffolgenden Kameraposition folgt hingegen ein starker Anstieg der Kurven aufgrund der schlechten Geometrie der Bildpunkte. Beendet wird der Genauigkeitsverlauf entlang der Trajektorie in einem Wertebereich ähnlich dem des Ausgangsbildes, da zurück im Außenraum erneut a priori bekannte und hochgenaue Passpunkte vorhanden sind.

Das Genauigkeitsniveau der Rotationen (Abb. 4b) weist im Gegensatz zu dem der Positionen nicht für alle Komponenten einen deutlichen Knick bei Position 7 auf, sondern nur für $\hat{\sigma}_{\dot{\phi}}$. Durch das unterschiedliche Verhalten der Komponenten der Orientierung zeigt sich die Abhängigkeit des Genauigkeitsverhaltens von der Aufnahmegeometrie.

Generell ist ersichtlich, dass die skalaren Genauigkeitsmaße dem Verlauf der einzelnen Komponenten der Standardabweichung folgen. Daraus lässt sich folgern, dass sie als Äquivalent zur Darstellung geeignet sind.

Als Fazit der Simulation zeigt sich, dass der Genauigkeitsverlauf im Wesentlichen den Erwartungen aufgrund der Bewegung vom Außen- in den Innenraum und wieder zurück entspricht. Wie insbesondere Kameraposition 7 gezeigt hat, darf jedoch nicht die Verteilung der Punktkonfiguration und Aufnahmegeometrie außer Acht gelassen werden.

3.2 Praktischer Messversuch

3.2.1 Versuchsaufbau und Durchführung

Der Versuch wurde in einer Messumgebung im Rahmen des Projektes MultiGO aufgebaut. Er besteht aus einem zentral angeordneten und um sieben Achsen beweglichen Roboter. Auf dessen Arm sitzt ein Kameraträger mit vier Kameras (Abb. 5). Diese sind Industriekameras vom Typ „eco655MVCE“ des Herstellers SVS-VISTEK (SVS-VISTEK, 2012), bestückt mit Weitwinkelobjektiven vom Typ „SV-0614H“ (Nr. 1, 3 und 4) beziehungsweise „SV-0814H“ (Nr. 2) von VS Technology (VST, 2012).

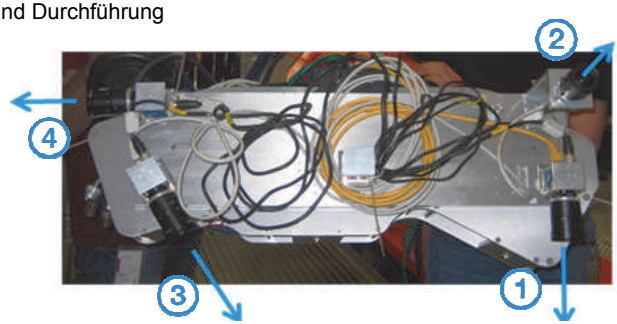


Abb. 5: Anordnung der Kameras auf dem Messsystem. Die Zahlen geben die Nummerierung der Kameras an

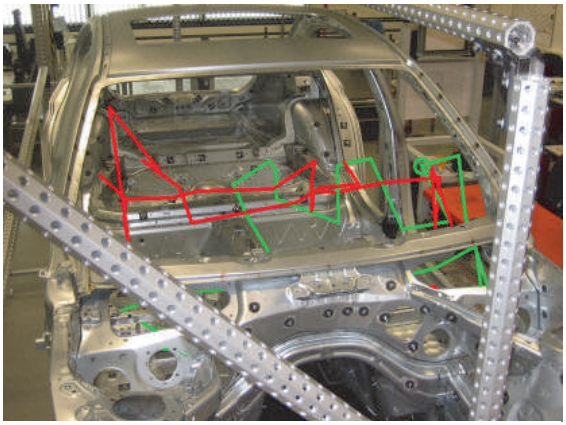


Abb. 6: Messpfad der Stereokameras. Beginn und Ende an durch Kringel markierter Stelle. Kamera 1 in roter, Kamera 3 in grüner Farbe dargestellt

Des Weiteren steht ein dichtes Referenzmarkenfeld zur Kalibrierung und Orientierung der Kameras im Fahrzeugaußenraum zur Verfügung. Das untersuchte Messobjekt (vgl. Abb. 6) ist eine Karosserie des BMW F30 der Baustufe 2 inklusive befestigten Seiten-, Rumpf- und Dachblechen sowie den nötigen Aussparungen.

Zu Beginn der Versuchsdurchführung wurden die in unterschiedlicher Größe vorhandenen Referenzmarken auf Hallenfundament, Rasterplatten und Rahmen angebracht. Anschließend wurden sie zur Berechnung der Referenzkoordinaten hochgenau photogrammetrisch aufgenommen und mit AICON 3D-Studio ausgewertet. Außerdem wurden zur Simulation von Features (z.B. markante Löcher im Blech) magnetische Referenzmarken im Fahrzeuginneren befestigt.

Anschließend wurden als erforderliche Rahmenbedingungen ein Schärfetest der Kameras auf dem Kameraträger durchgeführt und diese auch kalibriert.

Für den Versuch wurde ein vorgegebener Messpfad (Abb. 6), beginnend vom Außenraum in den Innenraum der Karosserie geführt und schließlich wieder zurück geführt. Im Innenraum wurden verschieden ausgerichtete Positionen angefahren. Nach Stoppen der Roboterbewegung an den einzelnen Standpunkten wurden die entsprechenden Bildaufnahmen ausgelöst.

3.2.2 Ergebnisse und Diskussion

Zur Schaffung von möglichst engen Rahmenbedingungen zur Evaluation des vorgestellten Ansatzes wurden nachfolgende Einschränkungen getroffen. Einerseits wurden zur Berechnung von Neupunkten nur Aufnahmen einer Kamera von zwei aufeinanderfolgenden Positionen miteinander kombiniert. Anderseits wurden die Aufnahmen des Mehrkamera-Systems nur innerhalb eines Standpunktes in sich ausgeglichen. Das hierzu verwendete Verfahren basiert auf HANEL (2012).
Die beiden unten zu sehenden Darstellungen zeigen den Genauigkeitsverlauf der vier Kameras entlang des gesamten Pfades getrennt nach Positions- (Abb. 7) und Rotationsparametern (Abb. 8). Dabei werden nicht die einzelnen Parameter dargestellt, sondern die im Abschnitt 3.1 eingeführten skalaren Genauigkeitsmaße. Das Eintauchen in die Karosserie findet statt an Position 4.
Zunächst ist zu erkennen, dass nicht alle Bilder orientiert werden konnten, sondern im ersten Drittel eine Lücke von drei Aufnahmen klafft.

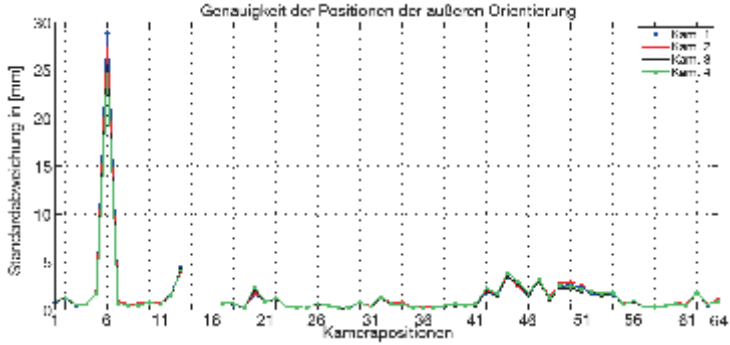


Abb. 7: Positionsgenauigkeit entlang des Messpfades. Dargestellt ist der 3D-Punktfehler nach NIEMEIER (2008) für alle vier Kameras des Multi-Kamerasystems

Dies ist darauf zurückzuführen, dass die vorhandene Punktkonfiguration im Fahrzeuginneren zwar von der Anzahl her ausreichend ist, aber die in diesen Bildern sichtbaren Verknüpfungspunkte an diesem Schritt der sukzessive durchgeführten Prozessierung noch nicht berechnet werden konnten.
Des Weiteren ist ersichtlich, dass bei den Positionsgenauigkeiten ein markanter Ausreißer an Position 6 vorliegt. Dessen schlechte Werte sind darauf zurückzuführen, dass sich nur eine der vier Kameras über Rückwärtsschnitt orientieren konnte und dabei zudem eine schlechte Aufnahmegeometrie vorlag. Die Standardabweichungen der anderen drei Kameras wurden deswegen allein durch die oben erwähnte Mehrkamera-Ausgleichung ermittelt und sind daher an die über

den Rückwärtsschnitt orientierte Kamera gekoppelt. An allen anderen Positionen liegt das Genauigkeitsniveau in einem Bereich von unter einem bis fünf Millimetern, weist also einen weitgehend gleichmäßigen Verlauf auf. Hier bewirkt die Mehrkamera-Ausgleichung, die an jeder Position durchgeführt wurde, ein vergleichbares Genauigkeitsmaß aller Kameras. Dennoch weisen die Kameras Nr. 1 und 3 bei detaillierter Betrachtung aufgrund ihrer Stereoausrichtung ein etwas niedrigeres Wertespektrum als die anderen beiden auf.
Der Genauigkeitsverlauf der Rotationen stellt sich ähnlich dar, so tritt der deutlichste Ausreißer an derselben Stelle wie bei den Positionen auf. Zusätzlich zeigen sich im Bereich vor und kurz nach den überhaupt nicht orientierbaren Kamerapositionen größere Ungenauigkeiten. Dies lässt sich auf die teilweise sehr geringe Neupunktzahl zurückführen, die zu einer höheren Unsicherheit in der Orientierungsbestimmung führt. Auch tritt das Absinken des Genauigkeitsniveaus der Rotationen in den Positionen um Nummer 45 etwas stärker hervor als bei der Positionsgenauigkeit. Dies lässt sich auf die ähnlichen Sichtfenster der Kameras an diesen Positionen zurückführen. Dadurch werden dieselben Punkte zur Orientierungsbestimmung verwendet, womit sich vorhandene Fehler in allen diesen Bildern auswirken.
In beiden Fällen bestätigt sich am Ende des Pfades nach Verlassen des Fahrzeuginnenraums an Position 55, dass ein ähnlicher Wertebereich wie zu Beginn erreicht wird.

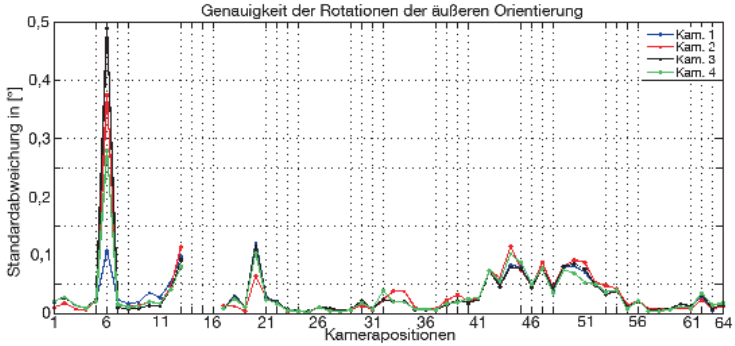


Abb. 8: Rotationsgenauigkeit entlang des Messpfades. Dargestellt ist der skalare Fehler nach FISCHER (2008) für alle vier Kameras des Multi-Kamerasystems

Als Fazit lässt sich folgern, dass der Verlauf der Genauigkeitskurven größtenteils konstant bleibt. Dennoch ist es notwendig, die kritischen Positionen zu analysieren und daraus gegebenenfalls Konsequenzen für den Versuchsaufbau und die Trajektorienwahl zu ziehen. Dennoch bleibt festzuhalten, dass das für die eingangs erläuterte Anwendung gewünschte Genauigkeitsniveau von wenigen Zehntel-Millimetern nicht erreicht werden konnte.

4 Ausblick

Zielsetzung für weitere Arbeiten kann es sein, statt der Magnetmarken im Fahrzeuginneren Features zu messen. Hierfür bieten sich sowohl punkt- als auch linienförmige Varianten an. Dazu ist

es notwendig, geeignete Punkte in der Karosserie zu identifizieren und passende Bildverarbeitungsalgorithmen zu entwickeln. In einem nächsten Schritt müssen diese evaluiert werden, um möglichst zuverlässige Konfigurationen ableiten zu können.

5 Danksagung

Das Projekt MultiGO wurde von der Bayerischen Forschungsstiftung (AZ-876-09) gefördert. Die Autoren bedanken sich bei der BMW AG, Werk München, für die freundliche Unterstützung zur Durchführung der Messungen in der Messzelle. Unser Dank gilt auch Herrn Gericke und Herrn Meyer von der Firma Steinbichler Optotechnik GmbH, Neubeuern, für die Bereitstellung von Hardware.

6 Literaturverzeichnis

DARR, T.; GÖTZ, C.; TUTTAS, S.; BECKER, T.; STILLA, U.; 2013: Genauigkeit der inversen Navigation bei hoher Passpunktsichtbarkeit zur Absolutpositionierung eines geführten Messsystems. Tagungsband der 33. Wissenschaftlich-Technischen Jahrestagung der DGPF, Bd 22

FACCHINETTI, C.; TIÈCHE, F. & HÜGLI, H., 1995: Self-positioning robot navigation using ceiling image sequences. In: Proceeding ACCV'95 (Asian Conference on Computer Vision), Singapore, 3:, S. 814-818

FISCHER, M., WS 2008: Vorlesung Robotik, Kinematik V. FH München:, Fakultät 07

FRASER, C.S., 1984: State oft the art in industrial photogrammetry. Commission V, XVIth ISPRS Congress, Kyoto. S. 166-180

HANEL, A., 2012: Orientierungsgenauigkeit eines Kamerasystems bei Bewegungen in einer Fahrzeugkarosserie. München: Technische Universität, Fakultät für Bauingenieur- und Vermessungswesen, Bachelorarbeit

GÖTZ, C.; TUTTAS, S.; HOEGNER, L.; EDER, K. & STILLA, U., 2011: Accuracy evaluation for a precise indoor multi-camera pose estimation system. In: Stilla U et al. (Eds) PIA11. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (3/W22): S. 7-102

KRAUS, K., 1996: Photogrammetrie: Band 2. Verfeinerte Methoden und Anwendungen. Berlin: Gruyter, 3. Auflage.

LUHMANN, T., 2010: Nahbereichsphotogrammetrie: Grundlagen, Methoden und Anwendungen. Berlin: Wichmann, 3. Edition.

NIEMEIER, W., 2008: Ausgleichungsrechnung – Statistische Auswertemethoden. Berlin: Gruyter

SCHEWE, H. 1987: Automatic Photogrammetric Car-Body Measurement. Proc. 41st Photogrammetric Week, Stuttgart, 123-129.

SVS-VISTEK, 2012: SVCam-ECO Datasheet. Seefeld: SVS-VISTEK GmbH http://www.svs-vistek.com/_products/cameras/svcam/eco/eco.php, (05.08.1012).

VST, 2010: CCTV Lenses, SV-H series. Tokyo, Japan: VS Technology Corporation http://products.vst.co.jp/en/search/lineup.php?sid=210, (06.08.2012).

WROBEL, B.P., 2001: Minimum Solutions for Orientation. Volume 34 of Springer Series in Information Science .Berlin Heidelberg: Springer

WELLS, G., VENAILLE, C. & TORRAS, C., 1996: Vision-based robot positioning using neural networks. Image and Vision Computing. 14(10): 715-732.

Untersuchungen zum Genauigkeitspotential eines hybriden Kamerasystems aus RIM- und RGB-Kamera

HEIDI HASTEDT & THOMAS LUHMANN

Zusammenfassung: Die dreidimensionale Rekonstruktion von Objekten und Oberflächen ist ein wichtiger Aufgabenbereich der optischen 3D-Messtechnik. In der Praxis sind für die 3D-Rekonstruktion Systeme aus hochauflösenden Sensoren in Kombination mit leistungsstarken Algorithmen erforderlich. In diesem Beitrag soll ein Multisensorsystem bestehend aus einer Range-Imaging-Kamera (RIM) und einer hochauflösenden RGB-Kamera vorgestellt werden. Die 3D-Rekonstruktion erfolgt hierbei über die Ermittlung einer verdichteten Punktwolke und wird anhand eines Referenzobjektes verifiziert. Es soll untersucht werden, welches Genauigkeitspotential durch Anwendung des hybriden Kamerasystems zu erwarten ist. Anhand von best-fit Modellierungen als auch der Bestimmung absoluter Genauigkeitsmaße über Monoplotting wird das Genauigkeitspotential untersucht und quantifiziert.

1 Einleitung

Die dreidimensionale Rekonstruktion von Objekten und Oberflächen ist eine Kernaufgabe im Bereich der optischen 3D-Messtechnik. Zum einen werden hochauflösende, genaue und auch verifizierte 3D-Modelle benötigt, die vor allem im Bereich der präzisen messtechnischen Aufgabenfelder zum Einsatz kommen. Zum anderen sind hochauflösende, kolorierte und texturierte 3D-Modelle gefordert, wie sie für 3D-Stadtmodelle und 3D-Visualisierungen in kulturhistorischen Aufgaben benötigt werden. Hauptsächlich werden hierbei statische Objekte betrachtet, da viele Messsensoren nicht für dynamische Anwendungen geeignet sind.
In der Praxis sind für die 3D-Rekonstruktion Systeme aus hochauflösenden Sensoren in Kombination mit leistungsstarken Algorithmen erforderlich. Diese Systeme reichen von kostenintensiven, professionellen Lösungen wie z.B. einen Laserscanner oder Streifenlichtprojektionssystem bis zu low-cost Systemen wie dem DAVID Laserscanner oder freie Structure-from-Motion-Programme (z.B. Bundler) in Kombination mit einer Digitalkamera. Der Einsatz von Laserscannern für die 3D-Erfassung und 3D-Rekonstruktion von Objekten hat sich in der Praxis etabliert, bei Aufnahmeentfernungen von ca. 1000 m sind typische Genauigkeiten im Bereich mehrerer Millimeter für jeden Messpunkt erreichbar. Sie liefern direkt 3D-Punktwolken, zumeist durch eine integrierte Kamera bereits koloriert. Vollständige 3D-Rekonstruktionen aus Laserscanner-Punktwolken zeigt LUHMANN (2010). KERSTEN (2006) zeigt eine kombinierte Rekonstruktion komplexer Objekte aus Laserscanning und Photogrammetrie mit einem Genauigkeitsniveau von 1 – 2 cm. Verschiedene Untersuchungen beschäftigen sich mit Monoplotting-Techniken zur Kombination von Laserscanning-Punktwolken mit Bilddatensätzen unter Verwendung verschiedener Interpolationsmethoden.

1) Heidi Hastedt, Thomas Luhmann, Institut für Angewandte Photogrammetrie und Geoinformatik, Jade Hochschule Oldenburg, Ofener Straße 16/19, 26121 Oldenburg, Email: [heidi.hastedt; luhmann]@jade-hs.de

SCHWERMANN & EFFKEMANN (2002) berichten in diesem Zusammenhang über Implementierung und Anwendung im Programmsystem PHIDIAS (PHOCAD 2012), wo unter Berücksichtigung gleicher Objektkoordinatensyteme ein Monoplotting über orientierte Bilddatensätze und einer 3D-Punktwolke durchgeführt werden kann.
Klassische photogrammetrische Bildverbände in Verbindung mit geeigneten Algorithmen oder der Einsatz von Mehrkameramesssystemen erfordern in der Regel gut texturierte bzw. ausreichend Eigenstruktur aufweisende Objekte für eine 3D-Rekonstruktion auf Basis bildbasierter Matchingverfahren. Für eine schnelle und kostengünstige 3D-Modellierung sind Digitalkameras in Kombination mit Structure-from-Motion (SFM) Verfahren (z.B. BUNDLER 2012) einsetzbar. KERSTEN & LINDSTAEDT (2012) zeigen vergleichende 3D-Rekonstruktionen unter Verwendung verschiedener SFM-Verfahren. Als Referenzdaten wurden hier Laserscanning-Aufnahmen hinzugezogen. Mittlere Genauigkeiten resultieren im Bereich 1mm bis 35mm für Ergebnisse aus Bundler.
Die Verwendung von 3D-Range-Imaging Systemen (RIM), auch TOF-Kamera genannt, ermöglicht die direkte Erfassung von 3D-Informationen. Diese nach dem Time-of-Flight- bzw. Phasenmessprinzip arbeitenden optischen Systeme ermöglichen die zeitgleiche Erfassung von Intensitätsbildern und Punktentfernungen für jedes Pixel, so dass mit einem kalibrierten System ein Objekt direkt dreidimensional erfasst werden kann. Nachteilig wirken die grobe Auflösung von max. 200 x 200 Pixel, relativ schlechte Messgenauigkeiten durch systembedingte Messabweichungen und Rauschen sowie starke Einflüsse durch Licht und Materialeigenschaften. Verschiedene Kalibrier- und Korrekturmodelle wurden u.a. durch LINDNER (2010) und KAHLMANN (2007) vorgestellt. LINDNER et al. (2007) und REULKE (2006) kombinieren eine RIM-Kamera mit einer hochauflösenden RGB-Kamera als kalibriertes System zur 3D-Rekonstruktion von Objekten.
Der Aufbau eines hybriden Kamerasystems, bestehend aus einer RIM-Kamera und einer hochauflösenden RGB-Kamera, ermöglicht zum einen eine erhöhte Auflösung der 3D-Punktwolke durch Erfassung der gleichen Szene und Berechnung der verdichteten Punktwolke durch die Parameter der Systemkalibrierung via Monoplotting. Zum anderen können im RGB-Bild sichtbare Feinstrukturen als auch direkt kolorierte 2.5D / 3D-Szenen rekonstruiert werden. Die Bildung einer Oberflächenvermaschung bildet dabei Grundlage zum 3D-Monoplotting und gleichzeitig die Herausforderung in der Detailrekonstruktion der zu beobachtenden Szene. Es soll untersucht werden, welches Genauigkeitspotential durch Anwendung des hybriden Kamerasystems zu erwarten ist. Als Objektreferenz für eine flächenhafte Rekonstruktion dient hierbei zunächst ein Fußraummodell eines Fahrzeuges. Anhand von best-fit Modellierungen als auch der Bestimmung absoluter Genauigkeitsmaße über Differenzen zu Passpunkten wird das Genauigkeitspotential untersucht und quantifiziert.

2 Multisensorsystem

2.1 Systemaufbau und Kalibrierung

Das Multisensorsystem besteht aus einer RIM-Kamera, der PMD[vision]® CamCube 3.0, in Kombination mit einer hochauflösenden RGB-Kamera, der AVT Stingray F-125C. Beide Kameras sind in einem Laboraufbau auf einer Schiene fest zueinander montiert und kalibriert.

Abb. 1 zeigt den Aufbau des Multisensorsystems. In einer gemeinsamen Ausgleichung mit dem Programm AXIOS Ax.Ori wurden über eine Testfeldkalibrierung (Abb. 2) die Parameter der inneren Orientierung (Tab. 1) sowie eine relative Orientierung mit dem Ursprung in der RIM-Kamera mit bx_{rgb} = 112.0074 mm, by_{rgb} = -11.4682 mm, bz_{rgb} = -7.0384 mm, ω_{rgb} = 0.009455 rad, φ_{rgb} = 0.141715 rad und κ_{rgb} = 0.000202 rad bestimmt.



Abb. 1: Hybrides Kamerasystem

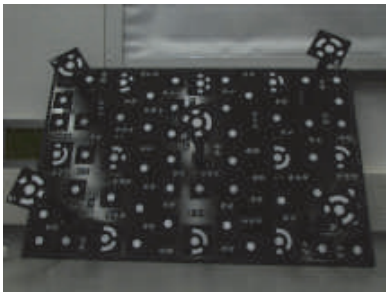


Abb. 2: Kalibriertestfeld

PMD[vision]® CamCube 3.0 200 x 200 Pixel, 45 µm Pixelgröße 12.8 mm Brennweite				AVT Stingray F-125C 1280 x 960 Pixel, 3.75 µm Pixelgröße 5 mm Brennweite			
c [mm]	12.4947 *±0.0019*	A_1	-0.0028 *±1.11e-5*	c [mm]	5.0019 *± 0.0006*	A_1	-0.0025 *±2.54e-5*
x'_0 [mm]	-0.1091 *±0.0027*	A_2	9.87e-6 *±7.94e-7*	x'_0 [mm]	-0.0221 *±0.0004*	A_2	2.65e-4 *±8.20e-6*
y'_0 [mm]	-0.0996 *±0.0027*	A_3	-3.32e-8 *±1.69e-8*	y'_0 [mm]	-0.0429 *±0.0006*	A_3	-8.32e-6 *±7.88e-7*
B_1	2.82e-5 *±2.27e-6*	C_1	-4.87e-5 *±2.13e-5*	B_1	1.37e-4 *±5.47e-6*	C_1	2.58e-4 *±1.48e-5*
B_2	-4.81e-5 *±2.87e-6*	C_2	-5.51e-4 *±2.02e-5*	B_2	-1.76e-4 *±7.44e-6*	C_2	-1.90e-4 *±1.18e-5*
R_0 [mm]	3.38			R_0 [mm]	1.80		

Tab. 1: Zusammenstellung der Systemkomponenten und Kalibrierung der IOR

Untersuchungen zeigen, dass je nach Reflexionseigenschaften des Objektes und der Entfernung zum Objekt stark unterschiedliche Entfernungsmessungen in Abhängigkeit zur Integrationszeit resultieren. Dies wurde an einer weißen, ebenen Wand bei einer Aufnahmeentfernung von <1 m und konstantem Aufbau getestet. Je nach gewählter Integrationszeit zwischen 11 ms - 9000 ms resultierten Streckenmessungen von 0.55 m - 0.8 m. Die für eine Auswertung zu verwendende Einstellung der Integrationszeit wird durch eine möglichst geringe Standardabweichung der ausgleichenden Ebene über alle 3D-Punkte ermittelt, in diesem Beispiel lagen Standardabweichungen zwischen 10 mm - 160 mm vor. Dieser Test lässt sich nicht auf beliebige zu messende Objekte übertragen, da in der Regel komplexe 3D-Objekte erfasst werden sollen. Für die richtige Auswahl der Integrationszeit sind verschiedene Messreihen erforderlich, um verschiedene Objektsituationen möglichst umfänglich abschätzen zu können. Zusätzlich sind

jeweils Aufwärmzeiten der Kamera von mindestens 30 Minuten für stabile Messungen zu berücksichtigen.

2.2 RIM/RGB-Prozessierung

Die mathematische Modellierung des Mehrkamerasystems basiert auf einer festen relativen Orientierung. Abb. 3 zeigt den funktionalen Zusammenhang für die RIM/RGB-Prozessierung nach HASTEDT & LUHMANN (2012).

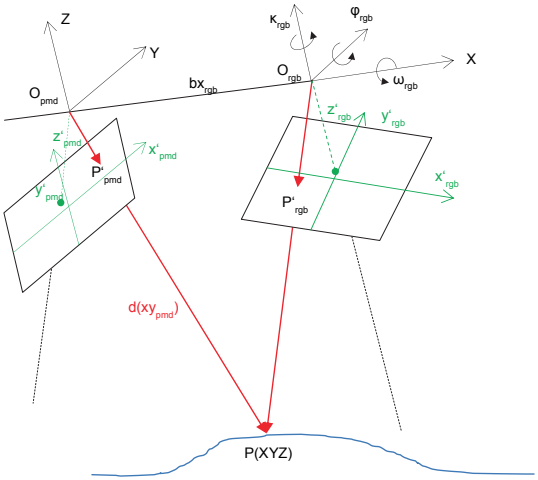


Abb. 3: Kamera- und Objektkoordinatensystem

Für die Prozessierung ist die Datenerfassung und -verarbeitung in einer C++-Routine implementiert. Die Generierung einer verdichteten Punktwolke erfolgt über die Berechnung des Schnittpunktes des Bildstrahls jedes Pixels der RGB-Kamera mit der RIM-Punktwolke, das Verfahren des 3D-Monoplottings folgt der Erläuterung nach HASTEDT & LUHMANN (2012). Hierbei wird die RIM-Punktwolke zuvor mit GEOMAGIC 2012 unter Berücksichtigung unterschiedlicher Prozessierungsschritte vermascht.

3 Praktische Untersuchungen

Zur Untersuchung der Leistungsfähigkeit des Multisensorsystems wurde ein Fußraummodell eines Kleinwagens als Testobjekt verwendet. Für dieses Testobjekt liegen Referenzdaten in Form eines Soll-Datensatzes für die Oberfläche als auch Passpunktkoordinaten vor. Abb. 4 zeigt das Fußraummodell im Versuchsaufbau mit dem Multisensorsystem. Zum einen wird das Ergebnis der RIM-Erfassung als Basisoberfläche für die RIM/RGB-Prozessierung näher betrachtet. Für die Datenverarbeitung soll weiterhin der Übergang einer 2.5D-Rekonstruktion durch

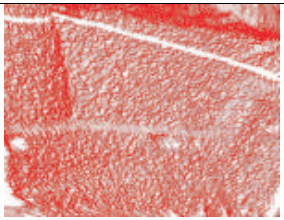
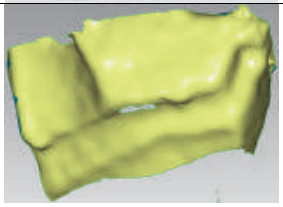
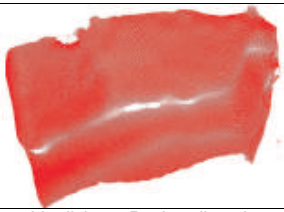
Monoplotting zu einer 3D-Rekonstruktion durch 3D-Monoplotting mittels unterschiedlich vermaschter Punktwolken aufgezeigt werden. Anhand von best-fit Modellierungen als auch der Bestimmung absoluter Genauigkeitsmaße über Differenzen zu Passpunkten wird das Genauigkeitspotential untersucht und quantifiziert.



Abb. 4: Versuchsaufbau am Fußraummodell

3.1 Basismodelle der RIM-Punktwolke für Analysen

Neben der Systemkalibrierung spielt die Verarbeitung der 3D-RIM-Punktwolke eine große Rolle. Bedingt durch das zu erwartende Rauschen und Vorliegen unterschiedlicher Fehlmessungen müssen verschiedene Datenfilterungen und ggf. Glättungen vorgenommen werden. Die Bildung einer Oberflächenvermaschung bildet dabei Grundlage für das Monoplotting und gleichzeitig die Herausforderung in der Detailrekonstruktion der zu beobachtenden Szene.

Basis ist die ungefilterte, jedoch beschnittene, Punktwolke der RIM-Kamera		
Basis ist die beschnittene, gefilterte und geglättete RIM-Punktwolke		
	Vermaschung der Original RIM-Punktwolke in Geomagic 2012	Verdichtete Punktwolke mit RIM/RGB-Prozessierung

Tab. 2: Vergleich der Eingangsdatensätze für Analysen

Tab. 2 zeigt die RIM-Punktwolke in der oberen Zeile als ungefilterte Punktewolke. Dargestellt ist in der linken Abbildung die mit GEOMAGIC 2012 vermaschte ungefilterte Punktwolke. Die Punktwolke ist hierbei beschnitten worden, da das Bildfeld der RIM-Kamera einen größeren Bereich abdeckt und hier grobe Fehlmessungen aufgrund der Objektbeschaffenheit nicht eliminiert werden können. Die vermaschte Oberfläche zeigt deutlich eine starke Unstetigkeit, die sich mit dem Rauschverhalten der RIM-Kamera erklären lässt. Die Einflüsse durch diese Unstetigkeiten wiederholen sich auch in der späteren Genauigkeitsbetrachtung. Die rechte Abbildung zeigt die nach der RIM/RGB-Prozessierung erzielte verdichtete Punktwolke. Tab. 2 zeigt in der unteren Zeile links eine beschnittene und zusätzlich stark gefilterte und geglättete Vermaschung (GEOMAGIC 2012) der RIM-Punktwolke, in der rechten Abbildung ist das Ergebnis der verdichteten Punktwolke dargestellt.

3.2 Ergebnisse 2.5D-Monoplotting in Phidias

Zur Ermittlung der Genauigkeit an diskreten Punkten bzw. Bereichen im 3D-Modell wurden Passpunkte über Monoplotting mit PHIDIAS (PHOCAD 2012) gemessen, die dafür verwendeten Punkte sind in Abb. 6 dargestellt. Für die Ergebnisbeurteilung wurden die gemessenen Punkte durch eine 7-Parameter-Transformation auf die Passpunkte transformiert. Tab. 3 zeigt die erzielten Ergebnisse. Die mittlere quadratische Abweichung liegt hier bei $RMSv_x \sim 4mm$, $RMSv_y = RMSv_z \sim 2mm$.

Ergebnisse des Monoplotting in Standard-RIM-Punktwolke			Ergebnisse des Monoplotting in verdichteter RIM/RGB-Punktwolke		
$RMSv_x$ [mm]	$RMSv_y$ [mm]	$RMSv_z$ [mm]	$RMSv_x$ [mm]	$RMSv_y$ [mm]	$RMSv_z$ [mm]
4.1	2.0	2.0	4.2	2.9	1.9

Tab. 3: Abweichungen der Monoplotting-Punktmessungen zu Passpunkten nach 7-Parameter-Ähnlichkeitstransformation

3.3 3D-Monoplotting

Für eine flächenhafte 3D-Rekonstruktion eines Fußraummodells ist ein 3D-Monoplotting eingesetzt worden. Abb. 5 zeigt die kolorierte, verdichtete Punktwolke auf Basis des gefilterten und geglätteten Modells der RIM-Punktwolke nach der RIM/RGB-Prozessierung. Zur Beurteilung des Genauigkeitsniveaus wurden für die in 3.1 vorgestellten Basismodelle der RIM-Punktwolken best-fit Anpassungen und 3D-Vergleiche zum Referenzmodell mit GEOMAGIC 2012 durchgeführt. Tab. 4 führt die Ergebnisse zusammen.

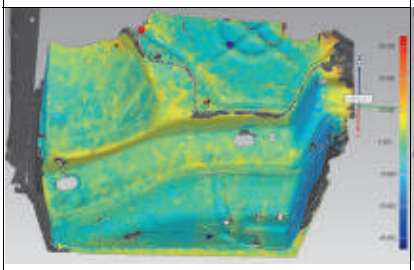
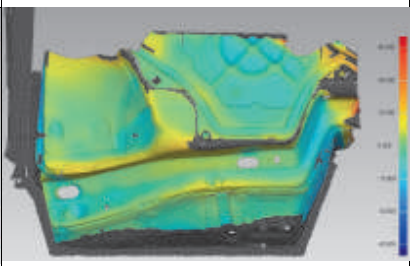
Nach einer best-fit Anpassung ergeben sich für die 3D-Abweichungen durchschnittliche 3D-Abweichungen zum Referenzmodell von +10.6 mm bis -6.8 mm bei einer Standardabweichung von 14.4 mm bei Verwendung der ungefilterten Punktwolke, sowie 3D-Abweichungen von +5.9 mm bis -5.4 mm bei einer Standardabweichung von 9.0 mm mit der gefilterten und geglätteten Punktwolke als Basismodell.



Abb. 5: Kolorierte Punktwolke der RIM/RGB-Prozessierung auf Basis des gefilterten Basismodells.



Abb. 6: Übersicht der Passpunkte für Monoplotting im Referenzmodell des Fußraums

	Vergleich verdichtete RIM/RGB-Punktwolke auf Basis des ungefilterten Modells der RIM-Punktwolke zur Referenz	Vergleich verdichtete RIM/RGB-Punktwolke auf Basis des gefilterten und geglätteten Modells der RIM-Punktwolke
		
std.	14.4 mm	9.0 mm
□	+10.6 mm bis -6.8 mm	+5.9 mm bis -5.4 mm

Tab. 4: 3D-Abweichung nach best-fit Anpassung in GEOMAGIC 2012 mit Respekt zum Referenzdatensatz.

4 Fazit und Ausblick

In dem vorliegenden Beitrag wurde ein Multisensorsystem bestehend aus einer RIM-Kamera und einer hochauflösenden RGB-Kamera vorgestellt. Die 3D-Rekonstruktion erfolgt hierbei über die Ermittlung einer verdichteten Punktwolke mittels RIM/RGB-Prozessierung. Das Genauigkeitspotential des abgeleiteten 3D-Modells wurde anhand eines Referenzobjektes, einem Fußraummodell eines Kleinwagens, getestet. Zum einem konnten Genauigkeiten durch Passpunktmessungen mittels 2.5D-Monoplotting im Bereich von $RMSv_x \sim 4$ mm sowie $RMSv_y = RMSv_z \sim 2$ mm ermittelt werden. Der Übergang zu 3D-Monoplotting erforderte die Betrachtung über eine best-fit Anpassung der Punktwolke zum Referenzmodell und des

anschließenden 3D-Vergleichs. Hierbei ergab sich unter Berücksichtigung eines gefilterten Basismodells für den 3D-Schnitt eine Standardabweichung der 3D-Abweichungen zum Referenzmodell von 9 mm. Die Ergebnisse zeigen, dass die Verwendung von Filteralgorithmen einen hohen Einfluss auf das Rekonstruktionsergebnis hat. Für weitere Untersuchungen sollen hierzu unterschiedliche Verfahren zur Vermaschung und Filterung der RIM-Punktwolke getestet und implementiert werden.

5 Literaturverzeichnis

BUNDLER, 2012. http://phototour.cs.washington.edu/bundler/ (19.12.2012)

GEOMAGIC, 2012. http://geomagic.com/de/ (19.12.2012)

HASTEDT, H., LUHMANN, TH., 2012: Investigations on a combined RGB / Time-of-flight approach for close range applications. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXIX-B5, pp. 333-338.

KAHLMANN, T. 2007. Range Imaging Metrology: Investigation, Calibration and Development. Dissertation ETH No. 17392.

KERSTEN, TH.P., LINDSTAEDT, M., 2012: Automatic 3D Object Reconstruction form multiple images for architectural, cultural heritage and archaeological applications using open-source software and web services. PFG 2012 / 6, pp. 727-740.

KERSTEN, TH.P., 2006. Combination and Comparison of Digital Photogrammetry and Terrestrial Laser Scanning for the Generation of Virtual Models in Cultural Heritage Applications. Ionnides, M., Arnold, D., Niccolucci, F., Mania, K. (eds.): 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage, VAST 2006: pp. 207-214

LINDNER, M., KOLB, A., HARTMANN, K. 2007. Data-Fusion of PMD-Based Distance-Information and High-Resolution RGB-Images. Signals, Circuits and Systems, 2007. ISSCS 2007. International Symposium on , vol.1, no., pp.1-4

LINDNER, M., 2010. Calibration and Real-Time Processing of Time-of-Flight Range Data. Dissertation Universität Siegen.

LUHMANN, T., 2010: Kombination von Laserscanning und Photogrammetrie – Potenziale und Grenzen. Terrestrisches Laserscanning 2010, Schriftenreihe des DVW, Nr. 64, pp. 3-20.

PHOCAD, 2012. http://www.phocad.de/ (19.12.2012)

REULKE, R. 2006. Combination of distance data with high resolution images. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVI, Part 5, Commission V Symposium, Dresden, Germany.

SCHWERMANN, R., EFFKEMANN, C. 2002. Kombiniertes Monoplotting in Laserscanner- und Bilddaten mit PHIDIAS. In: Photogrammetrie und Laserscanning – Anwendung für As-Built-Dokumentation und Facility Management, Wichmann, pp. 57-70

Epipolar Geometry of Heterogeneous Stereo Camera Systems

DOMINIK RUEß[1], ANDREAS LUBER[2] & RALF REULKE[1]

Zusammenfassung: In den letzten Jahrzehnten wuchs der Markt für Kameras mit nicht-perspektivischen Abbildungsverhalten beträchtlich. Insbesondere Kamerasysteme mit großem Öffnungswinkel finden zunehmend Verwendung. Analog zu klassischen, homogenen Stereosystemen, die aus zwei perspektivisch abbildenden Kameras bestehen, ist die Kombination verschiedenartiger Kamerasysteme denkbar.
In diesem Artikel untersuchen wir die Epipolargeometrie bei heterogenen Stereokamerasystemen mit generischen Abbildungsmodellen. Wir präsentieren eine analytische Lösung für die Transformation eines beliebigen Bildpunktes einer Kamera in seine Epipolarrepräsentation der zweiten Kamera. Zusätzlich haben wir einen Blockmatcher implementiert, um in diesen heterogenen Systemen eine Disparitätskarte zu erstellen. Die Ergebnisse zeigen, dass die so erzeugten Disparitätskarten trotz unterschiedlicher Abbildungseigenschaften der jeweiligen Kamera relativ genau werden.

1 Introduction

Recently, omnidirectional or wide-angle camera systems, e. g. fish-eyes or catadioptrics have become increasingly popular due to their extended fields of view. This minimizes the amount of cameras needed to cover a broad field of view. Like perspective camera systems, omnidirectional ones can be used for stereo camera systems. Disadvantages of these systems often include decreased resolution and a complex optical design. Particular models are required for such cameras, many of which were proposed in literature BAKER & NAYAR [1999]; BASU & LICARDIE [1995]; FLECK [1995]; GENNERY [2002]; GEYER & DANIILIDIS [2000]; KUMLER & BAUER [2000]; LUBER & REULKE [2010]; MIYAMOTO [1964]. Furthermore, many authors applied a general parametric model to describe the majority of commonly used camera systems including omnidirectional ones, FITZGIBBON [2001]. Independently from the camera class a generic model is applicable to most of the camera systems. This allows for a simple combination of two cameras possible. The modeling of epipolar geometry of a heterogeneous stereo camera system and its evaluation are the main topics of this work. Such systems combine the advantages of cameras, e.g. the wide field of view or the possibility to utilize epipolar geometry. The basic procedure of setting up the epipolar geometry with generic camera models is almost the same as it is with perspective cameras. The main difference is the appearance of the epipolar line that is used for stereo matching. Using generic camera models, these lines become curved. Along these

1) DLR, Institut für Robotik und Mechatronik, Rutherfordstr. 2, 12489 Berlin, Germany
2) DLR, Institut für Verkehrssystemtechnik, Rutherfordstr. 2, 12489 Berlin, Germany

curves homologous points have to be searched. The derivation and modeling of these epipolar curves will be shown in this work. Additionally, a stereo matcher that determines homologous points along the epipolar curves is implemented and used for deriving a dense disparity map of synthetic data.
This article is organized as follows. The next section gives an overview of related work concerning stereo vision with non-perspective camera systems. In the subsequent section an analytic solution for the epipolar geometry using generic camera models is presented. The proposed method is than applied to synthetic data. To prove its potential, the results are evaluated.

2 Related Work

OLLIS et al. [1999] have evaluated many different panoramic stereo configurations and discussed accuracy issues of these different setups. Here, an algebraic closed form solution for the stereo reconstruction can be achieved by using a particular two-parameter catadioptric camera model. ADORNI et al. [2003] uses a combination of catadioptric and perspective camera for mobile robot navigation. An implicit stereo algorithm is used for obstacle detection. ABRAHAM & FOERSTNER [2005] calibrate a stereo system consisting of two fish-eyes. Additionally they utilized rectification of the stereo images by projecting them onto cylinders, analogously to POLLEFEYS et al. [1999]. FIALA & BASU [2005] make use of the panoramic Hough model to calibrate catadioptric cameras. Straight homologous lines are detected and matched. In this way, a stereo reconstruction by applying epipolar geometry is possible. CARON et al. [2010] calibrate a homogenous catadioptric stereo system using the camera model for central catadioptrics proposed by BARRETO & ARAUJO [2001]. Here neither epipolar geometry nor stereo reconstructions were considered. SVOBODA & PAJDLA [2002] derived epipolar geometry of central catadioptrics. Their particular catadioptric model results in epipolar conics which serve as the basis for epipolar geometry instead of image space itself. MICUSIK & PAJDLA [2003] use epipolar geometry of omnidirectional camera for calibration. Besides calibrating the cameras, an essential matrix is derived that can be used for structure from motion. GEYER & DANIILIDIS [2002] lifted catadioptric image points to a three-dimensional unit sphere. A catadioptric fundamental matrix is derived and analyzed using these lifted coordinates. At the same time STURM [2002] also uses coordinate lifting and derive a 4x3 fundamental matrix where catadioptric and perspective cameras can be mixed into a heterogeneous stereo system. Later this idea is refined by STURM [2005] to a generalized epipolar constraint, applicable to non-central catadioptrics. BARRETO & DANIILIDIS [2005] adapted the idea and lift radial distorted camera images to a quadric 3D image space. Thus, the epipolar constraint can be linearized and yield a 4x4 radial fundamental matrix. LEI et al. [2009] uses a generalized essential matrix to set up epipolar geometry of arbitrary omnidirectional stereo camera systems. Images are projected onto cylinder and cuboids and unwrapped. This makes classical stereo methods, like stereo matching, possible. Finally STURM & BARRETO [2008] formally derived a general fundamental matrix of

two arbitrary central catadioptrics. The matrix is of dimension 15x15 and results in quadric projections of homologous points in stereo images. In RUEß et al. [2012] the effect of different projections characteristics in terms of radially distorted points in heterogeneous stereo camera system has been evaluated. It has been shown that differently distorted camera systems tend to increase the error in stereo reconstruction. LUBER et al. [2012] evaluated different generic camera models, proposed in literature. It was shown that these generic camera models can be used for calibration of perspective, fish-eye or catadioptric camera systems. Furthermore a combination of these different camera systems into a heterogeneous stereo camera system is possible and yield meaningful 3D stereo data.

2.1 Notation

In this article, we use $\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{X})$ to describe the complete imaging from object point $\boldsymbol{X}$ to image point $\boldsymbol{x}$. The image point $\boldsymbol{x}$ can be interpreted in sensor coordinates or in pixel coordinates – here this will be clear from the context, if not noted explicitly. We distinguish between first and second camera by using an apostrophe for the second camera, i.e. $\boldsymbol{x}$ and $\boldsymbol{f}(\boldsymbol{X})$ are used for the first camera and $\boldsymbol{x}'$ and $\boldsymbol{f}'(\boldsymbol{X})$ for the second camera. To display derivatives of a function $\dot{\boldsymbol{x}}(\alpha)$ is used and vectors of unit length are marked with a hat: $\|\hat{\boldsymbol{d}}\| = 1$.

3 Generic Epipolar Geometry

The concept of epipolar geometry does not have to result in epipolar lines for generic imaging geometry. The image of the epipolar plane is quite possibly a curved function. However, the most important feature of the epipolar geometry holds for generic camera models, as well: The camera centers and an object point form a plane. In standard pinhole geometry it was possible to determine a restricted search line $\boldsymbol{l}'$ from a point $\boldsymbol{x}$ by means of a fundamental matrix F: $\boldsymbol{l}' = \boldsymbol{F} \cdot \boldsymbol{x}$. It is not possible to completely rectify images from generic camera models, because close to 180 degree viewing angles, images would expand to infinity.
In this section, for a point $\boldsymbol{x}$ in the first camera, we present a way to compute an epipolar curve in a second camera. First of all, taking the inverse projection of the point $\boldsymbol{x}$ results in a an object point direction

$$\boldsymbol{d} = \boldsymbol{f}^{-1}(\boldsymbol{x}), \tag{3-1}$$

starting from the center of the first camera, $\boldsymbol{p}$. To obtain an parameterized epipolar curve $\boldsymbol{x}'_{\boldsymbol{p},\boldsymbol{d}}(\lambda)$ in the second camera, we just need to map the entire object ray to our camera:

$$\boldsymbol{x}'_{\boldsymbol{p},\boldsymbol{d}}(\lambda) = \boldsymbol{f}'(\boldsymbol{p} + \lambda \cdot \boldsymbol{d}). \tag{3-2}$$

This, however, is very impractical, since the epipolar curve can be closed within a camera image – we might have to go to $\lambda = \infty$, to close the gap.
We propose a new generic method: W.l.o.g we can assume that the second camera is positioned at the origin, $\boldsymbol{p}' = \mathbf{0}$. Let $\boldsymbol{e}_{p,d}(\lambda)$ be any smooth, cyclic, star shaped curve on the epipolar plane, with $\boldsymbol{e}_{p,d}(\lambda + t) = \boldsymbol{e}_{p,d}(\lambda)$, for some $t \in \mathbb{R}^+$and $\mathbf{0}$ being contained in the interior of this curve. Let further $\boldsymbol{R}'$ be the camera rotation matrix of the second camera. Now, a generic equation of epipolar curves in the second camera is given by:

$$\boldsymbol{x}'_{p,d}(\lambda) = \boldsymbol{f}'\left(\mathbb{R}' \cdot \boldsymbol{e}_{p,d}(\lambda)\right). \tag{3-3}$$

Observe that for pinhole cameras without distortion, this will result in two possibly identical epipolar lines, at the "front" and at the "back" of the camera. The latter one is invalid, obviously, since a pinhole camera can only be constructed with a field of view of less than 180 degree. For simplicity, we will refer to epipolar curves of equation (3-3 as $\boldsymbol{x}'(\lambda)$. This still means that $\boldsymbol{x}'(\lambda)$ and all similar functions depend on a fixed choice of $\boldsymbol{d}$ and $\boldsymbol{p}$.
To go from the general to the specific, we suggest using the unit sphere and unit circle for our purposes. Let S^2 be the unit sphere, which is centered at the second camera center, hence at $\mathbf{0}$. Let further $\hat{\boldsymbol{n}}_e$ be the normal vector of a given epipolar plane. It can be determined by:

$$\boldsymbol{n}_e = \boldsymbol{f}^{-1}(\boldsymbol{x}) \times (\boldsymbol{p} - \boldsymbol{p}') \tag{3-4}$$

If both vectors are collinear, or respectively the point $\boldsymbol{x}$ is on the imaged base-line, any plane containing both camera centers suffices. Now note different properties of the epipolar setup: The epipolar plane also contains $\mathbf{0}$ and hence, the section of any epipolar plane with normal $\hat{\boldsymbol{n}}_e$ and the unit sphere is a circle of radius 1. This also means that a simple sinus or cosines parameterization of the unit circle is favorable. Let

$$\boldsymbol{c}(\alpha) = \begin{pmatrix} \cos\alpha \\ \sin\alpha \\ 0 \end{pmatrix} \tag{3-5}$$

be a parameterization of the unit circle in the XY plane. This will be mapped into the epipolar plane by a rotation matrix $\tilde{\boldsymbol{R}}$. A rotation $\tilde{\boldsymbol{R}}$ of the XY plane to the epipolar plane can be determined by means of Gram-Schmidt orthonormalization. If $\hat{\boldsymbol{n}}_e$ is the third row of $\tilde{\boldsymbol{R}}$ and the latter is an orthonormal rotation, then $\tilde{\boldsymbol{R}}\boldsymbol{c}(\alpha)$ lies within the epipolar plane with normal $\hat{\boldsymbol{n}}_e$. Now, our specific suggestion for an equation of epipolar curves in the second camera is given by:

$$\boldsymbol{x}'(\alpha) = \boldsymbol{f}'\left(R'\tilde{R}\boldsymbol{c}(\alpha)\right) \tag{3-6}$$

In the second camera, to have $\alpha = 0$ represent the point $\boldsymbol{X} = \infty$, make sure the first column of rotation matrix corresponds to the same ray as the inverse object ray $\vec{\boldsymbol{d}}$ from the first camera. This ensures that the parameterization starts at infinite reconstruction depth, also refer to Figure 1 for an illustration. Observe that this is possible since $\langle \vec{\boldsymbol{d}}, \boldsymbol{n}_\varepsilon \rangle = 0$. To make sure the epipolar line is sampled towards the other camera, adapt the second row of the matrix by changing the sign (the second row is uniquely determined up to the direction, if allowing reflecting rotations).
Using equation (3-6, it is not clear how to determine a $\Delta\lambda$, such that it corresponds to a sampling step with a distance of approximately one pixel. One of the possible solutions, given an inverse object ray of the first camera, $\vec{\boldsymbol{d}}$, and a current point on that ray (which means a given distinct angle α_i), is: determine the absolute value of the local derivation of the epipolar curve in the second image. A $\Delta\lambda$ can be determined to approximately sample the curve at d_ε pixel distance of the last position:

$$\begin{aligned} \alpha_{i+1} &= \alpha_i + d_\varepsilon \cdot \Delta\alpha \\ &= \alpha_i + d_\varepsilon \cdot \frac{1}{\|\dot{\boldsymbol{x}}(\alpha)\|} \\ \Rightarrow d_\varepsilon &= \|\boldsymbol{x}'(\alpha_{i+1}) - \boldsymbol{x}'(\alpha_i)\|. \end{aligned} \tag{3-7}$$

The next method allows limiting the search space of α to certain depth boundary, i.e. α_{max} and α_{min}: Given an inverse object ray $\boldsymbol{d}$ of the first camera and a distance d, the corresponding α for the epipolar curve in the second image can be determined by means of two-dimensional trigonometry: W.l.o.g assume the epipolar plane is the $Z = 0$ plane. Further, $\tilde{\boldsymbol{p}}$ and $\tilde{\boldsymbol{p}}' = \boldsymbol{0}$ are the transformed camera centers within this plane. Now, α can be determined and hence, for d_{max} and d_{min} respective α_{max} and α_{min} can be finally established in the second camera, see Figure 1.

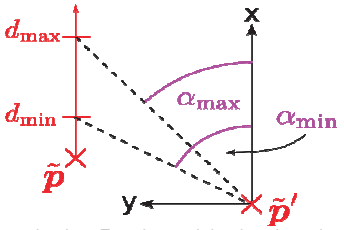


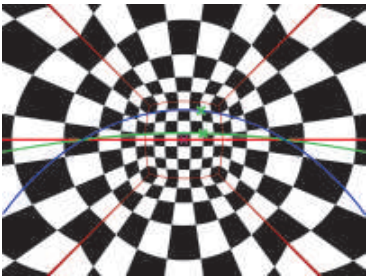
Figure 1: Search Angle Determination. For given minimal and maximal reconstruction depths, d_{max} and d_{min}, the respective α_{max} and α_{min} can be determined by means of trigonometry. Redly coloured variables, positions and directions are known entities. Note that $\vec{\boldsymbol{d}}$ has been mapped to the positive *X*-axis

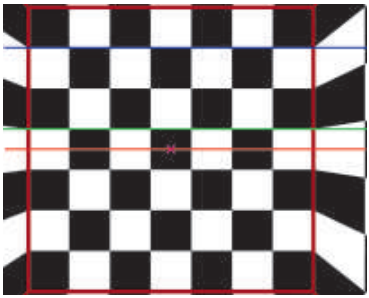
4 Reconstruction

To use the above tools of generic epipolar curves, we demonstrate a very basic dense reconstruction algorithm. We sample patches in both images and compare these patches by a metric. However, in contrast to the case of rectified images, it is not possible to use patches of the same size. Due to possibly completely different projections, a certain object surface may match to a different number of pixels and shape. To provide a proof of concept, we have implemented the following algorithm: For the first camera walk through all pixels. For each of these pixels $\boldsymbol{x}$ create a set of object points $\{\boldsymbol{X}\}$ in space. Every object point set $\{\boldsymbol{X}\}$ is a collection of points on the object ray $\boldsymbol{d_x}$ with distances to the first camera in between d_{min} and d_{max}. Next, sample a patch with a certain size (here, 9 pixels) centered at $\boldsymbol{x}$. The corners of the patch are back projected to space. For every $\boldsymbol{X}$, these corner rays are intersected with a plane, described by a normal $\boldsymbol{n}$. This surface plane $\boldsymbol{n}$ is assumed to best be viewed in two cameras by taking:

$$\boldsymbol{n} = \frac{1}{2}(\boldsymbol{d_x} + \boldsymbol{d'_X}), \tag{4-1}$$

where $\boldsymbol{d_x}$ and $\boldsymbol{d_X}$ are the directions to the current object point $\boldsymbol{X}$ from each respective camera, for a fixed $\boldsymbol{x}$. This results in 4 "surface" points which can be projected to the second camera. The resulting quadrangle can be transformed and resampled to a squared patch by means of a perspective transformation. This allows for direct patch comparison based on a metric and roughly resembles the same shape as in the first camera, for small patches. The best result per pixel is used to compute the distance.





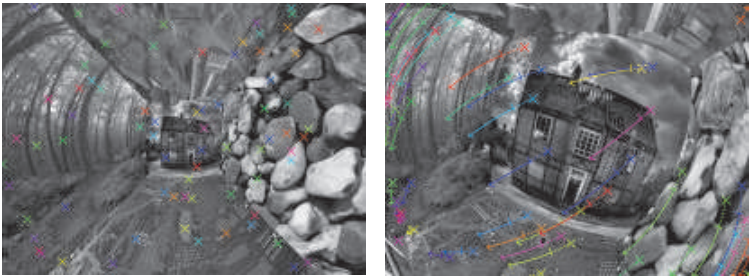


Figure 2: **Rendered Data Set**. Upper row: A cadadioptric, at $\boldsymbol{p} = \mathbf{0}$, and an ideal pinhole camera, at $\boldsymbol{p}' = (100,0,0)^T$, with corresponding epipolar lines. Lower row: A cadadioptric, at $\boldsymbol{p} = \mathbf{0}$ and a fisheye camera, at $\boldsymbol{p}' = (200,200,-200)^T$, with corresponding epipolar search lines and the cube faces differently textured.

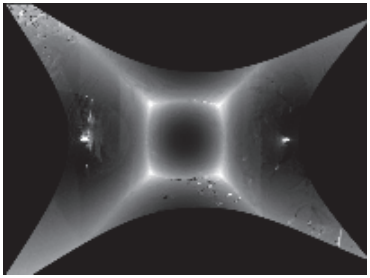


Figure 3: The depth map of the catadioptric camera at $\boldsymbol{p} = \mathbf{0}$ and the fisheye at $\boldsymbol{p}' = (100,0,0)^T$, in stereo normal position. Point reconstruction near the baseline is not possible. This can be observed around the large, white areas, where the second camera position has been imaged (white spot on the right hand side) as well as around the extension of the base line to space (white spot on the left hand side)

5 Results

For this paper, we present a toy example. Given a virtual cube, cameras are placed somewhere near the center of this cube. The side lengths of the cubes are virtual 2000mm. Each face is placed on one of the axis, perpendicular to it, at a distance of 1000mm to $\mathbf{0}$, each. The faces of the cube are rendered with individual textures per face. This setup gives accurate ground truth.
Three real camera parameter sets were used for this simulation example: A conventional pinhole model camera (ideal, distortion free) with a focal length of 2.5mm, a catadioptric camera with a

focal length of 1.13 mm and a fisheye lens with 1.78 mm focal length. The latter two cameras were real cameras, calibrated with LUBER & REULKE [2010]. These parameters were used to create an artificial view of the cube faces.
For reconstruction, the cameras are placed next to each other, on the *X*-axis and both have the same direction of view, parallel to the *Z*-axis. In the optimal case, the dense depth map represents the distance to the respective points of the cube faces.

5.1 Epipolar Geometry

The epipolar geometry has been computed for a different set of camera combinations and positions. In Figure 2 the first image pair row contains a two different camera classes in stereo normal position. The correct correspondence of the epipolar curves is particularly easy to see due to the use of a checkerboard pattern. The second row of this figure has a different camera system in a more general position.
Given a $d_{min} = 1000mm$ and a $d_{max} = 2000mm$, Figure 2 illustrates how the epipolar search space looks like. A cross indicates the corresponding position for the reconstructed point at infinite distance. The curve can then be sampled towards the closest distance.

5.2 Dense Block Matching

One result of the block matching algorithm is provided here. It shows that such a reconstruction is possible in general. We have used a patch of 9 by 9 pixels in the first or reference camera. Patches in the second image have varied in size, see text above. To compare the transformed second patch (transformed to size 9 × 9), the normed correlation coefficient has been used. Shows the result for one of the setups and it reveals a general problem in the proximity of the connecting base line. When viewing at each other, it is not possible to reconstruct a point in between, since there is no size information available on points. When compared to the ground truth, the root mean of squared distance error is 9.51 mm. This has only been measured in regions which could be observed by both cameras.

6 Discussion and Outlook

In this paper, we have presented an approach to transform the projective epipolar geometry to any kind of projection model. This particularly holds for generic projection models, as presented in LUBER & REULKE [2010]. Based on the geometry, it was possible to limit the search space on the epipolar curve, given some desired d_{min} and d_{max} extreme reconstruction distances. We have also presented a proof of concept for a block matcher, which is suitable for any kind of projection model. The block matching algorithm was able to capture the structure of the underlying toy example.
Some yet to solve issues have been revealed. These are different accuracy for different object point positions, problematic (inaccurate) reconstruction in the proximity of epipoles, differently

sized and shaped patches of two cameras. The problem of collecting ground truth data for real scenes has not been solved for this paper.
Future research will tackle these kinds of problems for dense reconstruction. Since the reconstruction does not seem to be possible in real-time, we may as well investigate in accurate but more expensive reconstruction approaches. A very interesting topic for future research is to provide accuracy estimation of reconstructions.

Acknowledgements

D. Rueß would like to acknowledge financial support of the Helmholtz Research School on Security Technologies.

7 References

GIOVANNI ADORNI, MONICA MORDONINI, STEFANO CAGNONI & ANTONIO SGORBISSA. Omnidirectional stereo systems for robot navigation. In CVPR, volume **7**, page 79, 2003.

SIMON BAKER & SHREE K. NAYAR. A theory of single-viewpoint catadioptric image formation. IJCV, 35(2):175–196, November 1999.

JOAO P. BARRETO & HELDER ARAUJO. Issues on the geometry of central catadioptric image formation. In CVPR, pages 422–427, 2001.

JOAO P. BARRETO & KOSTAS DANIILIDIS. Fundamental matrix for cameras with radial distortion. In ICCV, pages 625–632, 2005.

ANUP BASU & SERGIO LICARDIE. Alternative models for fish-eye lenses. Pattern Recognition, 16(4):433–441, 1995.

GUILLAUME CARON, EL MUSTAPHA MOUADDIB & E. MARCHAND. Single viewpoint stereoscopic sensor calibration. In Image/Video Communications over fixed and mobile networks, Rabat, Morocco, Morocco, 2010.

MARK FIALA & ANUP BASU. Panoramic stereo reconstruction using non-svp optics. Computer Vision and Image Understanding, 98:363–397, June 2005

ANDREW W. FITZGIBBON. Simultaneous linear estimation of multiple view geometry and lens distortion. In CVPR, volume 1, pages 125–132, 2001.

MARGARET M. FLECK. Perspective projection: The wrong imaging model. Technical report, University of Iowa, 1995.

DONALD B. GENNERY. Generalized camera calibration including fish-eye lenses. In IJCV, 68:239–266, 2002

CHRISTOPHER GEYER & KONSTANTINOS DANIILIDIS. Properties of the catadioptric fundamental matrix. In ECCV, pages 140–154, 2002. Springer-Verlag.

CHRISTOPHER GEYER & KOSTAS DANIILIDIS. A unifying theory for central panoramic systems and practical implications. In ECCV, pages 445–461, 2000.

JAMES J KUMLER & MARTIN L. BAUER. Fisheye lens designs and their relative performance. Proceedings of SPIE, 4093:360–369, 2000.

JIE LEI, XIN DU, YUN-FANG ZHU & JI-LIN LIU. Unwrapping and stereo rectification for omnidirectional images. Journal of Zhejiang University - Science A, 10:1125–1139, 2009.

ANDREAS LUBER & RALF REULKE. A unified calibration approach for generic cam- eras. In The international archives of ISPRS, pages 399 – 404, 2010.

ANDREAS LUBER, DOMINIK RUEß, KRISTIAN MANTHEY & RALF REULKE. Calibration and epipolar geometry of generic heterogeneous camera systems. In International Archives of ISPRS, volume XXXIX-B5, 2012.

BRANISLAV MICUSIK & TOMAS PAJDLA. Estimation of omnidirectional camera model from epipolar geometry. In CVPR, volume 1, pages 485–490, 2003.

KENRO MIYAMOTO. Fish eye lens. Journal Of The Optical Society Of America, 54(8):1060–1061, August 1964.

MARK OLLIS, HERMAN HERMAN & SANJIV SINGH. Analysis and design of panoramic stereo vision using equi-angular pixel cameras. Technical Report CMU-RI-TR-99-04, Robotics Institute, Pittsburgh, PA, January 1999.

MARC POLLEFEYS, REINHARD KOCH & LUC J. VAN GOOL. A simple and efficient rectification method for general motion. In ICCV, pages 496 – 501, 1999

DOMINIK RUEß, ANDREAS LUBER, KRISTIAN MANTHEY & RALF REULKE. Accuracy evaluation of stereo camera systems with generic camera models. In International Archives of ISPRS, volume XXXIX-B5, 2012.

PETER STURM. Mixing catadioptric and perspective cameras. In Proceedings of the Third Workshop OMNIVIS ’02, page 37, Washington, DC, USA, 2002.

IEEE Computer Society. PETER STURM & JOAO BARRETO. General imaging geometry for central catadioptric cameras. In ECCV, volume 4, pages 609–622. Springer, October 2008.

PETER F. STURM. Multi-view geometry for general camera models. In CVPR, pages 206–212, June 2005.

TOMAS SVOBODA & TOMAS PAJDLA. Epipolar geometry for central catadioptric cameras. Computational Vision, 49:23–37, August 2002.

Bewertung von inertialen Messsystemen mittels Laserscannern und bekannter Landmarken

ALEXANDER SCHLICHTING[1], CLAUS BRENNER[2] & STEFFEN SCHÖN[3]

Zusammenfassung: In diesem Beitrag wird ein Konzept zur Bewertung von INS/GPS-Sensoren in einem mit Laserscannern ausgestatteten Mobile-Mapping-System vorgestellt. Besonderes Augenmerk liegt hierbei auf der Genauigkeit der Position und Orientierung. Zu deren Beurteilung wurde ein Algorithmus entwickelt, mit welchem Landmarken aus einer Punktwolke segmentiert und anschließend Referenzlandmarken zugeordnet werden. Die Abweichungen zwischen den segmentierten Landmarken und den korrespondierenden Referenzen geben Aufschluss über die äußere Genauigkeit des Systems. Die Wiederholgenauigkeit wird anhand der Streuung der Abweichungen und der Differenzen der Abweichungen der segmentierten Landmarken zueinander bestimmt. Das entwickelte Bewertungskonzept wird anschließend auf drei verschiedene INS/GPS-Systeme angewendet.

1 Motivation

Mobiles Laserscanning erlaubt die Erfassung großer Datenmengen innerhalb von kurzer Zeit. Mit den Aufnahmen einer einzelnen Messfahrt können beispielsweise 3D-Stadtmodelle erstellt oder Objekte des Straßenraums kartographisch erfasst werden. Mobiles Laserscanning bietet somit im Vergleich zu herkömmlichen, statischen Messverfahren eine hohe Steigerung der Effizienz.

Um die während der Fahrt aufgenommenen Punkte direkt mit Koordinaten eines einheitlichen Koordinatensystems zu versehen, ist es notwendig, kontinuierlich die lokale Position und Orientierung (Pose) des Fahrzeuges zu bestimmen. Soll eine direkte Transformation der lokalen Koordinaten in ein globales Koordinatensystem vorgenommen werden, so wird in diesem Kontext auch die aktuelle Fahrzeugpose in diesem System benötigt. Zur Bestimmung der Pose werden in der Regel mit GPS gekoppelte Inertiale Navigationssysteme (INS) verwendet. Die Verwendung eines INS/GPS-Systems bietet den Vorteil, dass die Kurzzeitcharakteristik der Inertialeinheit mit der Langzeitstabilität des GPS kombiniert werden kann. Darüber hinaus liegen durch die GPS-Messungen globale Koordinaten vor.

Im Zuge der Sensorauswahl eines mit Laserscannern ausgestatteten mobilen Messsystems stellt sich die Frage, welche INS/GPS-Sensorik für bestimmte Anwendungen geeignet ist. Der vorliegende Artikel beschäftigt sich damit, die Eignung von inertialen Messsystemen für das Mobile Mapping mittels eines Testfelds festzustellen. Dies ist vergleichbar mit Untersuchungen zur Genauigkeit der direkten Georeferenzierung für luftgetragende Sensorsysteme, welche ebenfalls auf der Einrichtung von Testgebieten beruhen. Allerdings werden im vorliegenden Fall

1) Alexander Schlichting, Institut für Kartographie und Geoinformatik, Leibniz Universität Hannover, Appelstraße 9A, 30167 Hannover; E-Mail: alexander.schlichting@ikg.uni-hannover.de
2) apl. Prof. Dr.-Ing. Claus Brenner, Institut für Kartographie und Geoinformatik, Leibniz Universität Hannover, Appelstraße 9A, 30167 Hannover; E-Mail: claus.brenner@ikg.uni-hannover.de
3) Prof. Dr.-Ing. Steffen Schön, Institut für Erdmessung, Leibniz Universität Hannover, Schneiderberg 50, 30167 Hannover; E-Mail: schoen@ife.uni-hannover.de

vorhandene Merkmale des Straßenraums genutzt, wodurch keine Wartung des Testfelds anfällt. Dadurch entsteht ein sehr praktikables und kostengünstiges Verfahren.
Der Beitrag gliedert sich wie folgt. Kapitel 2 gibt einen kurzen Überblick über verwandte Arbeiten. In Kapitel 3 wird das Bewertungskonzept vorgestellt, wobei der Schwerpunkt auf die Genauigkeit der Positionsbestimmung und Orientierung des Systems gelegt wird. In Kapitel 4 wird beschrieben, wie dieses Konzept umgesetzt werden kann. Exemplarische Ergebnisse von Messfahrten werden in Kapitel 5 präsentiert, Kapitel 6 fasst schließlich die Ergebnisse zusammen.

2 Verwandte Arbeiten

Die Bewertung von INS/GPS-Sensoren findet zumeist durch die Analyse einer sich aus den Messungen des Positionierungssystems ergebenen Trajektorie statt. So wird in [DOROBANTU & ZEBHAUSER, 1999] die Trajektorie einer Low-Cost Inertialeinheit in Abhängigkeit der Updates der präzisen differentiellen GPS-Lösung verglichen. [NOURELDIN et al., 2009] wählen einen ähnlichen Ansatz. Sie erzeugen künstliche GPS-Ausfälle und betrachten die Drift des INS. [NIU et al., 2006] schlagen weiterhin eine Simulation der Navigationslösung vor. Hierzu wird das aus Labortests gewonnene Fehlerverhalten der Sensoren simuliert und den wahren Werten zugeführt. Anschließend kann die wahre Lage, Position, Geschwindigkeit und Beschleunigung mit den Werten verglichen werden, die der jeweilige Navigationsalgorithmus ausgibt. Nachteile dieses Verfahrens sind die Komplexität der verschiedenen Abweichungen sowie die Tatsache, dass hierbei auch das entsprechende GPS-Signal simuliert werden muss.
Die aufgezeigten Methoden betrachten lediglich die Kombination zwischen INS und GPS. Zur Bewertung eines solchen Sensorsystems in Kombination mit Laserscannern liegen bisher nur wenige Kriterien vor. Ein Konzept zur Bewertung der INS/GPS-Plattformen in einem mit Laserscannern ausgestattetem Mobile-Mapping-System wird im Folgenden vorgestellt.

3 Bewertungskonzept

Ziel des Bewertungskonzeptes ist es, Ergebnisse (Punktwolken, Fahrzeugpositionen) von Messfahrten eines beliebigen, kommerziellen Mobile-Mapping-Systems heranzuziehen, um möglichst einfach Aussagen über den Einfluss der verwendeten INS/GPS-Kombinationen zu erhalten. Dabei ist kein Zugang zur internen Verarbeitung der einzelnen Sensoren (GPS, INS, Laserscanner) oder deren Fusion notwendig.
Ein entscheidendes Kriterium zur Bewertung der INS/GPS-Sensoren ist die Genauigkeit des Systems. Für viele Anwendungsfälle ist vor allem die Richtigkeit der Pose relevant. Zu deren Beurteilung bedarf es einer Referenzlösung. Als Referenz dienen in unserer Arbeit Landmarken, welche vorab mit hoher Genauigkeit ($\sigma < 5$ cm) eingemessen wurden. Diese Landmarken werden in der sich aus den Messungen des Laserscanners ergebenden Punktwolke segmentiert. Anschließend wird die Abweichung zur Referenzposition untersucht. Die segmentierten Landmarken können gleichzeitig auch zur Bewertung der Wiederholgenauigkeit dienen, indem entweder die Streuung der Abweichungen zu mehreren Landmarken oder die Abweichungen zu derselben Landmarke zu unterschiedlichen Messzeitpunkten betrachtet werden.

Des Weiteren muss auch die Genauigkeit des Drehraten- und Beschleunigungssensors untersucht werden. Starke Beschleunigungen lassen sich in einem Fahrzeug durch Beschleunigungsvorgänge oder starke Bremsvorgänge erzeugen, hohe Drehraten im Gier-Winkel (Heading) durch enge Kurven. Werden parallel zu diesen Vorgängen Landmarken von den Laserscannern erfasst, so können anschließend wiederum die Abweichungen zu den Referenzlandmarken untersucht werden. Eine Möglichkeit zur Analyse starker Veränderungen im Roll- und Nick-Winkel ist das Fahren über eine Rampe oder Wippe. Hierbei können insbesondere die Auswirkungen auf die Form einer durch die Scanner erfassten Ebene betrachtet werden.
Weiter ist das Rauschen des Systems wichtig. Neben der direkten Betrachtung des Rauschens in der Position und der Orientierung des Fahrzeuges bietet es sich an, auch das Rauschen von Messpunkten auszuwerten, welche in einer Ebene liegen. Es ist jedoch zu beachten, dass somit auch das Rauschen der Laserscanner und die Unebenheit des betrachteten Objektes, beispielsweise einer Häuserwand, mit einbezogen werden.
Insbesondere in Stadtgebieten mit großen Abschattungen spielt auch die Drift der Inertialeinheit eine bedeutende Rolle. Ist keine GPS-Lösung verfügbar, so wächst der Fehler in der Positionsbestimmung mitunter sehr schnell an. Dieses Verhalten kann untersucht werden, indem das GPS-Signal aus den Messungen entkoppelt wird.

4 Umsetzung des entwickelten Bewertungskonzeptes

Zur Bestimmung der einzelnen in Kapitel 2 genannten Bewertungskriterien wurde ein Algorithmus entwickelt, welcher Landmarken aus der mit Laserscannern aufgezeichneten Punktwolke segmentiert. Diese Position wird anschließend mit der Referenzposition verglichen, welche aus einer terrestrischen Aufnahme der Landmarken mit übergeordneter Genauigkeit bekannt ist. Um zusätzlich die Genauigkeit der Orientierung im Roll- und Nick-Winkel zu untersuchen, werden die Auswirkungen auf eine Ebene bei auftretenden Winkelveränderungen untersucht. Diese Veränderungen werden durch eine Fahrt über eine Rampe erzeugt. Während der Fahrt mit einer Fahrzeugseite über diese Rampe wird eine zur Fahrtrichtung parallele Ebene durch die Laserscanner erfasst.

4.1 Wahl eines Referenzgebietes



Abb. 1: Gewähltes Referenzgebiet [GOOGLE, 2012]

Das in Abb. 1 dargestellte Referenzgebiet enthält sowohl Geradenabschnitte als auch Kurven. Zusätzlich bietet ein Kreisverkehr die Möglichkeit, Wiederholmessungen durchzuführen und die Genauigkeit der Orientierung zu untersuchen. Auf Grund der klaren und einfachen Zylinderform und der regelmäßigen Verteilung wurden Straßenlaternen als Landmarken gewählt. Abweichungen zwischen Referenzmessungen und den Messungen mit einem Mobile-Mapping-System erlauben so, die Genauigkeit der Positionsbestimmung abzuschätzen. In diesem Gebiet wurden 31 Straßenlaternen mit einer

Genauigkeit von unter fünf Zentimetern terrestrisch eingemessen. Diese Laternen dienen in der Auswertung als Referenzlandmarken.

4.2 Segmentierung der Landmarken

Zur Segmentierung der Landmarken ist es zunächst notwendig, potentielle stangenförmige Landmarken in der Punktmenge zu detektieren. Es muss validiert werden, ob die gefundenen Objekte den tatsächlichen Landmarken entsprechen, sodass anschließend die detektierten Punktmengen den Referenzlandmarken zugeordnet werden können. Um die Distanz der segmentierten Landmarke zu deren Referenz zu bestimmen, muss die Landmarke auf einen Punkt reduziert werden, beispielsweise durch eine Zylinderschätzung.

Der Algorithmus basiert im Wesentlichen darauf, dass für verschiedene Regionen in der Punktwolke die Anzahl der Fußpunkte in der Horizontalebene bestimmt wird. Regionen mit einer hohen Punktanzahl kommen als Landmarken in Frage. Ein ähnliches Konzept zur Stangendetektion findet sich auch in [BRENNER, 2009]. Die potentiellen Landmarken werden in der Folge auf weitere Kriterien überprüft und anhand dieser bewertet. Das Objekt mit dem höchsten Bewertungsfaktor innerhalb eines Suchradius wird einer Referenzlandmarke zugeordnet.

Im ersten Schritt wird der Bereich um eine bekannte Landmarke in einem *Octree* (räumliche Datenstruktur) gespeichert. Anschließend wird für jedes *Voxel*, das heißt für jeden Knoten in der untersten Ebene des Octrees, die Anzahl der Punkte innerhalb eines vorgegebenen Radius r in der *x/y*-Ebene bestimmt. Als Radius erwies sich der doppelte gemessene Radius der jeweiligen Referenzlandmarke als geeignet. Die Entscheidung, welches gefundene Objekt der Landmarke zugeordnet wird, erfolgt auf Basis eines Bewertungsfaktors, welcher sich aus der Anzahl der enthaltenen Punkte und der Distanz zur Landmarke ergibt. Ist die Anzahl der enthaltenen Punkte groß und die Distanz niedrig, so ergibt sich ein hoher Bewertungsfaktor. Konnte einer Landmarke eine segmentierte Stange zugeordnet werden, so werden aus den enthaltenden Punkten die Parameter eines bestangepassten Zylinders geschätzt. Da die Parameter nicht in jedem Fall bestimmt werden können, wird außerdem der Schwerpunkt der Punktmenge gespeichert. Die Höhe der Landmarken bezieht sich auf deren Bodenpunkt. Aus diesem Grund wird für den Schwerpunkt und den Zylindermittelpunkt die Höhe des niedrigsten Punktes der extrahierten Punktemenge betrachtet.

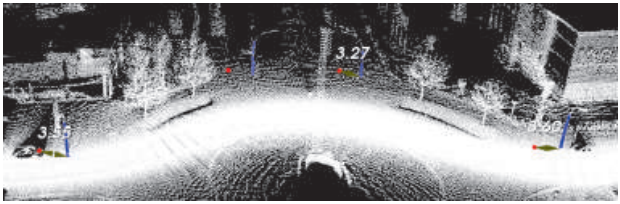


Abb. 2: Punktwolke mit segmentierten Landmarken (blau), deren Referenzpositionen (rot) sowie der jeweiligen Distanz.

Abb. 2 zeigt eine so durchgeführte Extraktion der Landmarken, wobei die Distanz zur Referenz hier nur für Stangen eingezeichnet ist, deren Zylinderparameter erfolgreich geschätzt werden konnten. Zu den rot eingefärbten Referenzpositionen werden die jeweiligen Landmarken in der Punktwolke gefunden und zugeordnet. Allerdings können auch fehlerhafte Zuordnungen auftreten, wie in Abb. 3 (links) zu erkennen ist. In diesem Fall befindet sich ein Baum im Suchradius der als Referenz dienenden Landmarke. Für dessen Baumstamm ergibt sich ein

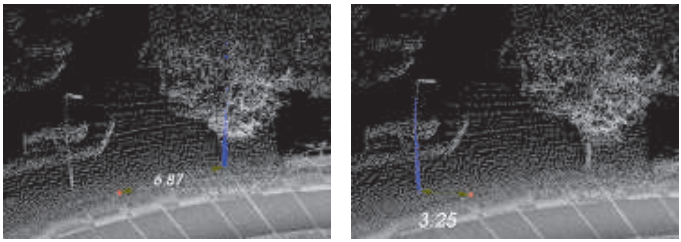


Abb. 3: Segmentierung ohne (links) und mit (rechts) Berücksichtigung der Anzahl der Punkte in der Umgebung sowie des maximalen Abstandes zweier aufeinanderfolgender Punkte in der Höhe.

höherer Bewertungsfaktor als für die Punktmenge der Straßenlaterne. Diese fehlerhafte Zuordnung soll unterbunden werden, indem für jedes gefundene Objekt überprüft wird, wie viele Punkte sich in einem erweiterten Radius in der Umgebung des Objektes befinden. Überschreitet die Anzahl der Punkte, die zusätzlich in dem erweiterten Radius auftreten einen Grenzwert, so wird das Objekt als mögliche Landmarke verworfen. Weiter treten bei Bäumen oder Büschen hohe Punktabstände in der Höhe auf, sodass sich hier über einen maximalen Abstand ein weiteres Filterkriterium definieren lässt. In unserer Arbeit wurde dieser Grenzwert zu einem Meter bestimmt. Es ist zu beachten, dass für diese Überprüfung die Punktmenge der Höhe nach geordnet sein muss.
Werden diese Überprüfungen vorgenommen, so erfolgt wie in Abb. 3 (rechts) ersichtlich, eine Verbesserung der Zuordnung.
Um darüber hinaus falsche Zuordnungen zu verhindern, wird der *Iterative-Closest-Point-Algorithmus* (ICP) [BESL & MCKAY, 1992] eingesetzt, welcher der Zuordnung mehrerer geometrischer Objekte zueinander bei unbekannten Korrespondenzen dient.

4.3 Ableitung von Qualitätsparametern

Zur Beurteilung eines INS/GPS-Sensors hinsichtlich dessen Genauigkeit müssen wie im Folgenden beschrieben die Werte der jeweiligen Kenngrößen bestimmt werden. In diesem Beitrag liegt der Fokus auf der Bestimmung der Richtigkeit und der Wiederholgenauigkeit der Pose sowie einer hiervon getrennten Untersuchung der Genauigkeit der Orientierung.

4.3.1 Richtigkeit der Pose

Um die Richtigkeit der Bestimmung der Pose zu untersuchen, werden die segmentierten Landmarken mit deren Referenzen bezüglich der Position verglichen werden. Hierbei werden die Abstände in alle drei Koordinatenrichtungen und die euklidische 2D- (x,y) und 3D-Distanz

(x,y,z) sowohl zu den geschätzten Zylindermittelpunkten als auch zu den Schwerpunkten der segmentierten Landmarken betrachtet. Anschließend wird aus den einzelnen Abweichungen der quadratische Mittelwert (RMS) berechnet. Über eine Visualisierung lassen sich zusätzlich Landmarken mit weiterhin fehlerhaften Zuordnungen oder sonstigen Fehlern bestimmen, sodass diese Landmarken anschließend aus der Berechnung des RMS entfernt werden können.

4.3.2 Wiederholgenauigkeit der Pose

Zur Beurteilung der Wiederholgenauigkeit können die Abweichungen zwischen segmentierten Stangen betrachtet werden, die denselben Referenzen zugeordneten wurden. Eine Visualisierung der Abweichungen ist in Abb. 4 zu sehen, wobei die blau und gelb eingefärbten Punkte die aus der Punktwolke der jeweiligen Vorbeifahrt die extrahierten Punkte darstellen. Auch hier werden die entsprechenden Standardabweichungen automatisch berechnet.

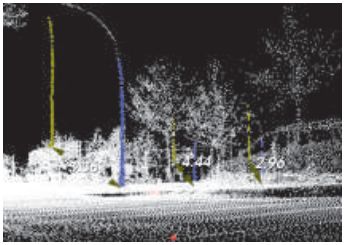


Abb. 4: Vergleich zweier Landmarken in verschiedenen Vorbeifahrten.

4.3.3 Genauigkeit der Orientierung

Die Genauigkeit der Orientierung wird in diesem Fall durch eine Analyse der Auswirkungen hoher Werte und starker Veränderungen im Roll- und Nick-Winkel während der Fahrt über eine Rampe untersucht. Der Algorithmus erkennt diese Veränderungen und teilt die Punktmenge in drei Abschnitte auf. Zur Segmentierung der Ebene aus der Punktwolke wird ein Region-Growing-Algorithmus verwendet. Abb. 5 zeigt die Aufteilung der Ebene. Die relativ starken Drehraten treten hierbei während der Aufnahme des Abschnitts 2 auf, während die Abschnitte 1 und 3 die Ebenen bezeichnen, in denen keine hohen Winkel und Drehraten auftreten. Anschließend werden aus den Punktmengen der einzelnen Abschnitte Ebenenparameter geschätzt und die Standardabweichung der jeweiligen Punktmenge zu dieser Ebene bestimmt. Diese Standardabweichung wird als Indikator zur Beurteilung der Genauigkeit der Orientierung genutzt. Ist die Abweichung der Ebene zum Abschnitt 2 signifikant höher als für die übrigen Abschnitte, so lässt dies auf einen verbleibenden Einfluss der hohen Drehraten und Orientierungswinkel schließen.

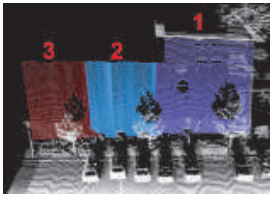


Abb. 5: Abschnitte der aufgenommenen Ebene.

Neben der Standardabweichung werden auch die Winkel $\varphi_{i,j}$ der Abschnitte zueinander analysiert. Diese ergeben sich exemplarisch für den Winkel zwischen Abschnitt 1 und 2 aus Gleichung (1).

$$\varphi_{1,2} = \arccos\left(\frac{n_1 \, n_2}{\quad}\right) \tag{1}$$

Hierbei bezeichnen n_1 und n_2 die Normalenvektoren der entsprechenden Ebenen. Es ist zu erwarten, dass die Veränderungen der Orientierungswinkel des Fahrzeuges durch die Messungen

der Drehratensensoren kompensiert werden. Je nachdem, wie hoch der Winkel zwischen den Abschnitten ist, desto schlechter ist diese Kompensation und damit die Genauigkeit des Sensors zu bewerten.

5 Anwendung der entwickelten Methoden



Abb. 6: Schematische Messkonfiguration

Das entwickelte Konzept wurde zur Bewertung der Genauigkeit dreier INS/GPS-Kombinationen unterschiedlicher Genauigkeitsklassen verwendet: Ein Sensor der Firma Oxford Technical Solutions (OXTS) mit einem Preis von etwa 30.000 Euro, ein weitaus günstigeren MEMS-Sensor der Marke XSens für etwa 4.000 Euro sowie einem Prototypen der Firma Northrop Grumman LITEF, zu dessen Preis noch keine Angaben gemacht werden können. Diese mit GPS gekoppelte Inertialsensoren waren in einem mit drei SICK-LMS511-Zeilenlaserscannern ausgestatteten VW Passat verbaut. Der systematische Fehler der Distanzmessung der Scanner beträgt bei einer Reichweite von 10-20 Metern 35 Millimeter, der zufällige Fehler acht Millimeter. Die Messkonfiguration ist in **Fehler! Verweisquelle konnte nicht gefunden werden.** schematisch abgebildet.

Zur Beurteilung der Sensoren wurden Fahrten bei unterschiedlicher Geschwindigkeit (15 km/h, 30 km/h) und in unterschiedlicher Richtung im Messgebiet durchgeführt, wobei stets alle drei Sensoren mitgeführt wurden. Die Messungen der Laserscanner erfolgten hierbei mit einer Frequenz von 100 Hz. Insgesamt konnten für drei Umfahrten zwischen 82 und 92 Landmarken (abhängig vom INS/GPS-System) aus den Punktwolken segmentiert und deren Referenzen zugeordnet werden. In diesem Beitrag konzentrieren wir uns auf die horizontalen Abweichungen der segmentierten Stangen. Diese werden für die jeweiligen Sensoren in Abb. 77 dargestellt.

Des Weiteren wurden zur Untersuchung der Genauigkeit der Orientierung drei Fahrten mit jeweils einer Fahrzeugseite über eine Rampe durchgeführt, während die Laserscanner eine zur Fahrtrichtung parallele Gebäudewand aufgenommen haben. Der Abstand des Fahrzeuges zur Ebene blieb hierbei konstant.

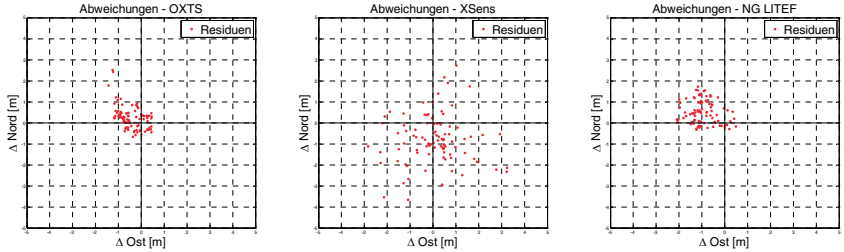


Abb. 7: Plot der Abweichungen der segmentierten Landmarken zu deren Referenzpositionen.

5.1 Beurteilung der Richtigkeit

Die sich aus den Messungen ergebenden RMS-Werte der Zylindermittelpunkte sind in Tab. 1 aufgelistet. Hierbei beschreibt *PP* die verfügbare Postprocessing-Lösung der mit dem NG-LITEF-Empfänger gemessenen Trajektorie.

RMS [m]	OXTS	XSens	NG LITEF	NG LITEF PP
2D	**0,90**	1,84	1,34	1,33
Nord	**0,61**	1,40	0,74	0,74
Ost	**0,66**	1,20	1,12	1,11

Tab. 1: RMS der Abstände der segmentierten Landmarken zu deren Referenzpositionen. Die niedrigsten Abweichungen sind hervorgehoben.

Es ist ersichtlich, dass die Abweichungen des OXTS-Sensors am geringsten sind. Hier beträgt der RMS-Wert zur Referenzposition einer Landmarke in der Lage 0,90 Meter. Die Abweichungen der XSens- und LITEF-Lösungen sind mit 1,84 bzw. 1,34 Meter deutlich höher.

5.2 Beurteilung der Wiederholgenauigkeit

Zur Untersuchung der Wiederholgenauigkeit des Systems sind verschiedene Ansätze möglich. An dieser Stelle wurden die Streuung der Abweichungen aus Abb. 7, die Abweichungen zwischen denselben segmentierten Landmarken bei Wiederholmessungen sowie die Veränderungen der Abweichungen in aufeinanderfolgenden Landmarken untersucht.

1 σ [m]	OXTS	XSens	NG LITEF	NG LITEF PP
2D	**0,59**	0,94	0,62	0,62
Nord	**0,56**	1,16	0,53	0,53
Ost	**0,50**	1,20	0,60	0,60

Tab. 2: Standardabweichung der Differenzen zu den Referenzlandmarken. Die niedrigsten Abweichungen sind hervorgehoben.

Betrachtet man die Streuung zwischen den gemessenen und den Referenzpositionen der Landmarken, so ist ersichtlich, dass die Abweichungen des XSens-Sensors in etwa um den Faktor zwei höher sind, als die der übrigen Lösungen (siehe Tab. 2).
Die mittleren Standardabweichungen zwischen den jeweils selben Landmarken in drei unterschiedlichen Messfahrten sind in Tab. 3 aufgelistet.

1 σ [m]	OXTS	XSens	NG LITEF	NG LITEF PP
2D	**1,14**	1,42	1,36	1,18
Nord	**0,73**	0,98	0,87	0,77
Ost	**0,82**	1,03	1,04	0,89

Tab. 3: Mittlere Standardabweichungen der Abweichungen der segmentierten Landmarken in drei Vorbeifahrten. Die niedrigsten Abweichungen sind hervorgehoben.

Tab. 4 zeigt die quadratischen Mittel der Veränderung der Abweichungen zweier aufeinander folgenden Landmarken L_i und L_{i+1} zur Referenzposition. Diese Veränderung ergibt sich aus

$$\Delta L_i L_{i+1} = \Delta L_i - \Delta L_{i+1}, \text{ mit } \Delta L_i = L_{seg} - L_{ref} . \quad (2)$$

L_{seg} bezeichnet hierbei die aus der Punktwolke segmentierte Landmarke, L_{ref} dessen Referenz. Die mittlere Distanz der betrachteten Landmarken beträgt 52,05 Meter.
Die INS/GPS-Plattform der Firma XSens liefert auch hier die höchsten Abweichungen, während wiederum die OXTS- und LITEF-Plattformen vergleichbare Ergebnisse liefern.

RMS [m]	OXTS	XSens	NG LITEF	NG LITEF PP
2D	**0,21**	0,71	0,24	0,24
Nord	**0,19**	0,83	0,24	0,24
Ost	**0,20**	0,90	0,24	0,25

Tab. 4: Quadratisches Mittel zwischen zwei aufeinanderfolgenden Landmarken.

5.3 Beurteilung der Genauigkeit der Orientierung

Tab. 5 gibt das Rauschen der segmentierten Ebenen sowie die Winkel zwischen den Ebenen an. Es ist zu erkennen, dass weder das Rauschen unter Einfluss der Drehraten (Abschnitt 2) höher ist, noch die Winkel zwischen den Abschnitten 1 und 2 sowie 2 und 3 signifikant höher sind. Es ist anzumerken, dass sich auch die Beschaffenheit der Gebäudewandabschnitte auf die Ergebnisse auswirkt. Dennoch kann kein Einfluss der Drehraten und hohen Orientierungswinkel auf die Messgenauigkeit des Gesamtsystems festgestellt werden.

		OXTS	XSens	NG LITEF	NG LITEF PP
Rauschen	1 (1 σ) [m]	0,046	0,043	0,041	0,039
	2 (1 σ) [m]	0,021	0,032	0,018	0,015
	3 (1 σ) [m]	0,017	0,046	0,020	0,016
Winkel	1-2 [°]	0,48	0,41	0,69	0,64
	1-3 [°]	0,96	0,72	1,11	0,16
	2-3 [°]	0,67	0,77	0,21	1,02

Tab. 5: Standardabweichungen der Gebäudewandabschnitte und Winkel zwischen den entsprechenden Ebenen für die einzelnen Sensoren.

6 Zusammenfassung

In diesem Beitrag wurde ein Konzept vorgestellt, um die Genauigkeit von INS/GPS-Systemen in mit Laserscannern ausgestatteten Mobile-Mapping-Systemen zu bewerten. Hierzu wurden insbesondere die Richtigkeit und die Wiederholgenauigkeit der Pose sowie die Genauigkeit der Orientierung als wesentliche Bewertungskriterien festgelegt. Es wurden Methoden entwickelt, um aus den Messungen der Laserscanner die entsprechenden Kenngrößen zu bestimmen. In diesem Zusammenhang wurden in einem Referenzgebiet Landmarken mit übergeordneter Genauigkeit eingemessen. Durch einen entsprechenden Algorithmus ließen sich diese Landmarken in Form von Straßenlaternen aus den in Messfahrten aufgezeichneten Punktwolken segmentieren und die Position der segmentierten Objekte mit den Referenzen vergleichen. Kern

dieses Algorithmus ist die Projektion der Punktwolke auf die *x*/*y*-Eben. Zur Segmentierung einer Landmarke werden Gebiete mit einer hohen Punktdichte gewählt und die zugehörigen dreidimensionalen Punkte extrahiert. Anhand der Anzahl der Punkte und der Distanz zur Referenz wird anschließend entschieden, welche Punktmenge der jeweiligen Landmarke zugeordnet wird. Fehlerhafte Zuordnungen werden durch die Betrachtung der Anzahl der Punkte in einem erweiterten Radius und der maximalen Distanz aufeinanderfolgender Punkte verhindert. Mit diesem Algorithmus kann eine automatische Segmentierung und Zuordnung der Landmarken aus einer aufgenommenen Punktwolke erfolgen.
Aus den Abweichungen der segmentierten Landmarken zu deren Referenzen lässt sich auf die äußere Genauigkeit (Richtigkeit) des verwendeten INS/GPS-Sensors schließen, aus der Streuung der Abweichungen auf die Wiederholgenauigkeit (Präzision).
Die Genauigkeit der Orientierung wurde durch eine Fahrt über eine Rampe und anschließender Analyse der parallel von den Laserscannern aufgenommenen Ebene untersucht.
Das entwickelte Konzept wurde auf drei verschiedene Systeme angewendet. Hierbei zeigte sich, dass die Kenngrößen zur Bewertung der Genauigkeit bestimmt werden konnten. Für die Richtigkeit und die Wiederholgenauigkeit ergaben sich für die Sensoren der Firma OXTS und NG LITEF vergleichbare Ergebnisse, für den deutlich preiswerteren XSens-Sensor wie erwartet geringere Genauigkeiten. Bezüglich der Orientierungswinkel und Drehraten konnte kein signifikanter Einfluss auf die Genauigkeit des Gesamtsystems ermittelt werden.
Für eine gesamte Bewertung eines Sensors gilt es weitere Bewertungskriterien zu definieren und Methoden zu entwickeln, diese automatisch aus Messungen des Mobile-Mapping-Systems zu ermitteln.
Insgesamt lässt sich feststellen, dass eine Methode entwickelt werden konnte, um eine automatische Bewertung der Genauigkeit von INS/GPS-Sensoren mit Hilfe von Laserscannern und bekannter Landmarken in einem Mobile-Mapping-System erfolgreich durchzuführen.

7 Danksagung

An dieser Stelle möchten wir uns bei Dr.-Ing. Kay Fuerstenberg, Dr.-Ing. Roland Krzikalla sowie Dipl.-Ing. Marc Sowen für die Durchführung der Messungen und die Unterstützung während der Auswertung der Ergebnisse bedanken.

8 Literaturverzeichnis

BESL, P. & MCKAY, N., 1992: A method for registration of 3-D shapes. IEEE Trans. Pattern Analysis and Machine Intelligence, **14** (2), S. 239-256.

BRENNER, C., 2009: Extraction of features from mobile laser scanning data for future driver assistance systems. Advances in GISScience, Lecture Notes in Geoinformation and Cartography, Springer Berlin Heidelberg, S. 25-42.

DOROBANTU R, & ZEBHAUSER, B. R, 1999: Field Evaluation of a Low-Cost Strapdown IMU by means GPS. Ortung und Navigation, 1/1999, S. 51-65.

GOOGLE INC., 2012: Google Earth. 03.12.2012.

NIU, X.; GOODALL, C.; NASSAR, S. & EL-SHEIMY, N., 2006: An efficient method for evaluating the performance of MEMS IMUs. IEEE/ION Position Location and Navigation Symposium, PLANS 2006, San Diego, USA, S. 766-771.

NOURELDIN, A.; KARAMAT, T.B.; EBERTS, M.D. & EL-SHAFIE, A., 2009: Performance enhancement of MEMS-based INS/GPS integration for low-cost navigation Applications. IEEE Transactions on vehicular Technology, **58** (3), S. 1077-1096.

3D GDI zur Unterstützung der Standortbewertung von Windenergieanlagen (WEA)

INES DÖRING

1 Einleitung

Zur Erreichung der Ziele für die Energiewende zeigt der Vortrag, welche Rolle Geoinformationen anhand des Beispiels der Standortplanung von Windenergieanlagen spielen.
Windkraft ist nicht nur die kostengünstige Variante der regenerativen Energien, sie besitzt auch das höchste Potenzial. Laut einer Untersuchung des Fraunhofer-Instituts für Windenergie und Energiesystemtechnik (IWES) könnten zwei Prozent der Fläche des Binnenlandes in Deutschland 65 Prozent des deutschen Stromverbrauchs durch Windenergieanlagen (WEA) abdecken. Doch es gibt Hindernisse, vor allem durch aufwendige Planungs- und Genehmigungsverfahren. An diesem Punkt wird die Energiewende langsam und kompliziert.
Im Zuge von Genehmigungsverfahren ist es gerade hier wichtig Transparenz aufzuzeigen. Für Ihre Planungen können Sie Standortdiskussionen objektivieren und Ihren Bürgen damit einen frühen Einblick erlauben.
Anhand unseres Software-Know-hows möchten wir unsere Möglichkeiten zur Standortbewertung von Windenergieanlagen vorstellen. Mit diesem Thema adressieren wir neben den Auswertemöglichkeiten für Experten auf Basis von ArcGIS auch nicht GIS-Anwender innerhalb und außerhalb der Verwaltung. Der Bürger selbst kann mit einfachsten Interaktionen die Einflüsse von Planungen für Windenergieanlagen ermitteln und kommunizieren. Hier haben wir einen vollständigen über das Internet bedienbaren Prozess entwickelt, der es erlaubt, dem System eine Standortplanung zuzuführen und entsprechend für diese Planung einen Bewertungsbericht auf Basis von ArcGIS oder als einfaches 3D PDF-Dokument zu erstellen. Die ingenieurtechnische Bewertung der Standorte wird hierbei nicht ersetzt, sondern

M.O.S.S. Computer Grafik Systeme GmbH · Hohenbrunner Weg 13 · 82024 Taufkirchen
E-Mail: idoering@moss.de

sie erhalten eine einfache Möglichkeit in der Anfangsphase der Standortbewertung mit einigen wenigen Parametern Ihre Planung transparent und allgemeinverständlich zu präsentieren. Dazu wählen sie in der einfachsten Form die Art der WEA sowie den Standort aus und bekommen per Mail ein 3D-PDF zugesandt.
Durch die Nutzung von Internettechnologie kann die Funktionalität in bereits existierende Webauftritte integriert werden.

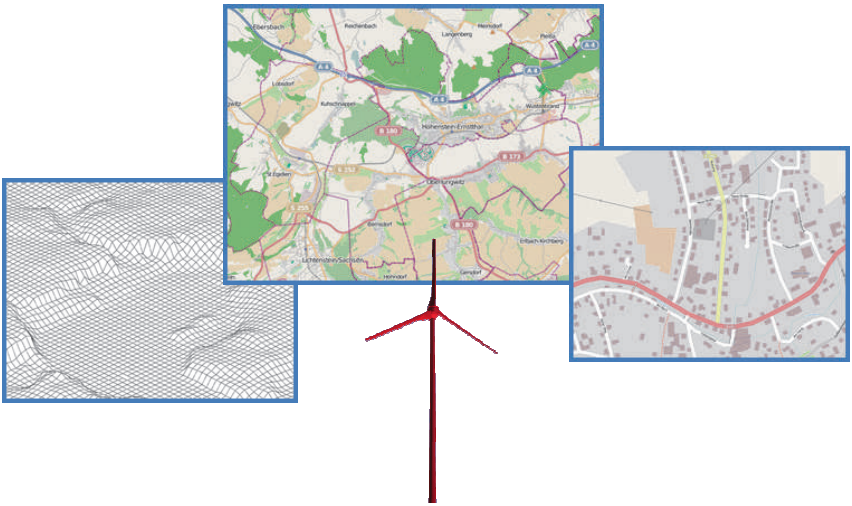
2 3D-GDI zur Unterstützung

Durch die Unterstützung von 3D-Gedateninfrastrukturen wird aufgezeigt, welche Rolle hierbei 3D-Daten und damit verbunden eine Geodateninfrastruktur zur Datenhaltung und -verteilung spielt.

Als 3D-Umgebung können bereits vorliegende 3D Stadt- und Landschaftsmodelle, aber auch dynamisch von der Plattform z.B. aus Open-Streetmap erzeugte Modelle, genutzt werden. Verwendet wird hier unser Produkt novaFACTORY. Mittlerweile bei fast allen deutschen Landesvermessungen dient es für die Verwaltung, Prozessierung und Bereitstellung der geotopographischen Daten. Im Bereich 3D-GDI wird die blattschnittfreie und datenbankbasierte Haltung von 3D-Stadtmodellen auf Basis von CityGML unterstützt. Die Daten können via Web wahlweise in den unterschiedlichsten Formaten abgegeben werden, wobei sich hier je nach Wunsch die verschiedensten Parameter auswählen lassen.



Basierend auf novaFACTORY wurde die Plattform myWINDRADL.de für die Standortplanung von Windenergieanlagen entwickelt. Mit Unterstützung von envVision, das für Datenmanagement, Umweltanalysen- und -planungen zum Einsatz kommt.



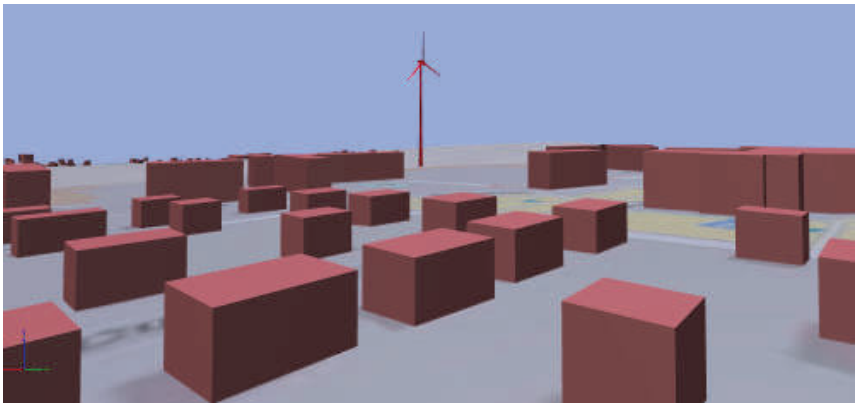
3 Sichtbarkeitsanalyse

Mit der aktuell kostenlos zur Verfügung stehenden Sichtbarkeitsanaylse können sie sich anhand eines 3D-Modells einen Überblick über die Standortwahl einer Windenergieanlage verschaffen. Dabei helfen ihnen Sichtbarrieren einen möglichst realitätsnahen Eindruck der Sichtbarkeit zu vermitteln. Als Sichtbarrieren werden alle Objekte bezeichnet, die die Sicht auf ein anderes Objekt (in diesem Fall die Windenergieanlage) verhindern.



Am Anfang der Planung von WEA steht immer die Standortanalyse. Für die Analyse werden das digitale Geländemodell mit einer Auflösung von 20 m,

bereits existierende Windenergieanlagen sowie OpenStreetMap Daten mit der Datengrundlage für die Sichtbarrieren Gebäude und Waldflächen modelliert. Die vorhandenen Gebäude und Waldflächen werden von 2D nach 3D extrudiert. Ausgegangen wird hierbei bei den Gebäuden von der Grundfläche. Je nach Größe der Grundfläche wird eine Gebäudehöhe zugewiesen. Für Waldflächen wird zunächst standardmäßig eine Höhe von 25 m verwendet.



Für die Weboberfläche der myWINDRADL-Plattform wird der WEGA 2011 Client hier verwendet. Auf der Deutschlandkarte Zoomen Sie mit den Werkzeugen links im Bild in die Region, in der Sie Ihre Berechnung durchführen möchten. Anschließend wird durch einen Klick auf die entsprechend geplante Windenergieanlage die Anlage per Mausklick platziert. Mit der Eingabe Ihrer Emailadresse und dem Klick auf den Button Start wird aus den hinterlegten Ausgangsdaten eine 3D-PDF Datei erzeugt, die Ihnen ein Gebiet von 2x2km ausgehend von dem gesetzten Windrad darstellt.

Der Funktionsumfang der Lösung kann individuell erweitert werden und umfasst die vordefinierten Analysen:

- Die Planungsdarstellung in einer interaktiven 3D Szene, die eine Gesamtübersicht gibt.
- Die Abstandsanalyse, die die gesetzlich einzuhaltenden Schutz- und Abstandsflächen im Umfeld der geplanten Anlage ausweist.
- Die Analyse des Schattenwurfes der Planung und der Beeinträchtigung durch Schatten zu den Externsonnenständen zur Sommer- und Wintersonnenwende.

- Für die Schallschutzanalyse erfolgt die Simulation, der sich für den Generator ergebenden Schallemissionen, mittels einer Lärmsimulationssoftware.
- Die Analyse von Landnutzungskonflikten erfolgt durch Verschneiden der Planungen mit anderen Landnutzungsformen oder ausgewiesenen Vorranggebieten.

4 Weitere Analysen

Bei der **Abstandsanalyse** kennzeichnen farbige Kreise die einzuhaltenden Abstandsflächen für Windkraftanlagen.

- Rot (innerer Bereich): Abstandsfläche
- Orange (mittlerer Bereich): Projektion der Anlagenhöhe
- Gelb (äußerer Bereich): Gewährleistung der Standsicherheit

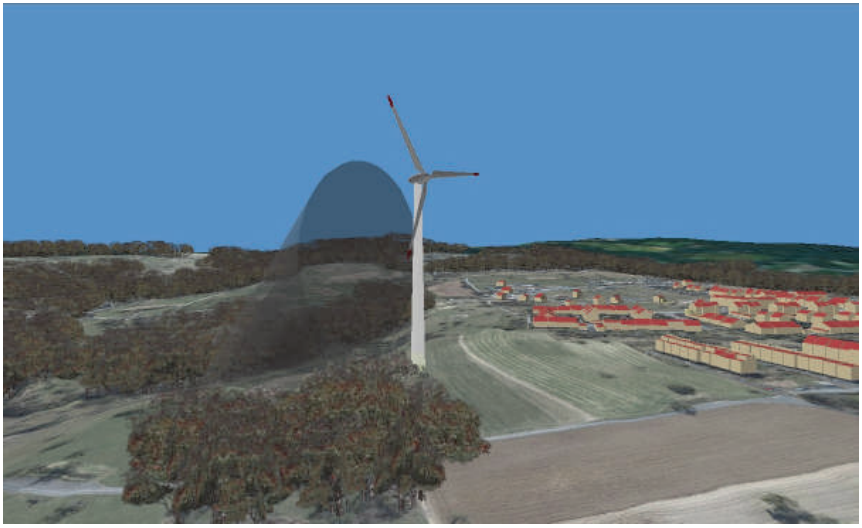


Die einzelnen Bundesländer haben Richtlinien zur Planung entwickelt, bei der solche Abstandsgrößen geregelt werden. Im Rahmen der Genehmigungsverfahren ergeben sich aber unterschiedliche Abstandserfordernisse und Restriktionen, die in der Regel einzelfallbezogen zu klären sind.

In der **Schattenwurfanalyse** wird der Schatteneinfluss der geplanten Windkraftanlage für die Extremtage eines Jahres berechnet:

- Zwei für die Jahreszeitenextreme (Sommer, Winter)
- Eine für die Äquinoktien (Frühling/Herbst)





Das Angebot ist je nach Nutzungsintensität und Anforderung gestaffelt. Zielgruppe der Anwendungen sind Kommunen, Planer, Windparkbetreiber oder auch Herstellerfirmen.

5 Ausblick

Im erweiterten Angebot von myWINDRADL.de werden exemplarisch die hinsichtlich des Sonnenstandes markanten Tage herangezogen: Die beiden Tagundnachtgleichen im Frühjahr und Herbst sowie die Winter- und Som-

mersonnenwende. Der Schattenwurf auf Wohnhäuser zum Beispiel darf theoretisch nicht mehr als 30 Stunden pro Jahr und 30 Minuten pro Tag betragen. Für solche Fragenstellungen liefert myWINDRADL die Antworten mit einem schnellen einfach zu bedienenden Internetservice.

Die Planungsparameter einzelner Vorhaben können gespeichert werden und sind so wieder abrufbar oder können veröffentlicht werden. Erzeugt wird eine Transparenz mit der die Nachteile und Vorteile, für die Planung einer Windenergieanlage, gegenübergestellt werden.

3DRoadTV-Service – Georeferenzierter 3D-Bilddatendienst für die Infrastrukturpflege und -verwaltung

HANNES EUGSTER[1,3], FRANK GOTTSMANN[2], CHRISTOPH KÄSER[2], STEPHAN NEBIKER[1] & RAINER KOCH[4]

Zusammenfassung: Eine hochaufgelöste georeferenzierte 3D-Bilddatenbasis unserer Strassen oder Schienenkorridore eignet sich zur Schaffung einer umfassenden Geoinformationsplattform für virtuelle Feldbegehungen, systematische Veränderungsanalysen, sowie die Kartierung und Überlagerung beliebiger vektorieller Geodaten. Entscheidend für den Mehrwert der Bilddaten ist ihre breite und verteilte, intuitive Nutzung bei Infrastrukturbetreibern, der öffentlichen Verwaltung sowie bei Ingenieurunternehmungen.

1 Einleitung

1.1 Motivation und Ausgangslage

Mobile Stereobild-basierte Messsysteme ermöglichen die effiziente und einfache hochaufgelöste Erfassung von Strassen- und Schienenkorridoren. Die bisherige konventionelle Vermessung im Feld mittels GNSS oder Tachymetrie ist in vielen Fällen zu aufwändig oder gefährlich, womit eine mobile Datenerfassung ins Zentrum des Interesses gerückt ist. Die Stereobilder können mit gleichzeitig aufgezeichneten Navigationsdaten und mit Hilfe von Dense Image Matching Methoden zu einer präzise georeferenzierten 3D-Bilddatenbasis aufbereitet werden, die anschliessend für verschiedene ausgestaltete 3DCityTV-Services zur Verfügung stehen (NEBIKER & EUGSTER, 2013). Der 3DCityTV-Service kann in der Cloud betrieben, komfortabel in bestehende (Geo-) Informationssysteme integriert und verteilt auf unterschiedlichsten Arbeitsplätzen und Endgeräten genutzt werden. Mögliche Anwendungen sind die visuelle Gewinnung von Informationen in den Strassen- oder Schienenkorridoren, das zentimetergenaue Messen in drei Dimensionen sowie das Überlagern und automatisierte oder interaktive Kartieren von Geoobjekten. Abbilddung 1 zeigt anhand einer einfachen Strassen-Szenerie die hohe Informationsfülle eines georeferenzierten 3D-Bildes.

Erste experimentelle mobile Bildaufnahmen im Straßenraum für Navigations- und Vermessungszwecke wurden bereits zu Beginn der 80er-Jahre durchgeführt (LIPPMAN, 1980). Ebenfalls bereits 20 Jahre zurück liegt die Entwicklung erster Stereobild-basierter Mobile Mapping Prototypsysteme (NOVAK, 1993; SCHWARZ et al., 1993). Heute kommen Bildbasierte, Laserscanningbasierte, Querebenheits-, Längsebenheits- und Griffigkeitssensoren auf mobilen

1) Hannes Eugster, Stephan Nebiker, Fachhochschule Nordwestschweiz, Institut Vermessung und Geoinformation, 4132 Muttenz, Schweiz; E-Mail: hannes.eugster@fhnw.ch, stephan.nebiker@fhnw.ch
2) Frank Gottsmann, Christoph Käser, Bundesamt für Strassen ASTRA, 3003 Ittigen, Schweiz; E-Mail: frank.gottsmann@astra.admin.ch, christoph.kaeser@astra.admin.ch
3) Hannes Eugster, iNovitas AG – Mobile Mapping Solutions, 4132 Muttenz, Schweiz; E-Mail: hannes.eugster@inovitas.ch
4) Rainer Koch, Rosenthaler + Partner AG, Management und Informatik, 4132 Muttenz, Schweiz; E-Mail: kor@rpag.ch

Trägerplattformen für die Erfassung unterschiedlicher Strasseninformationsparameter zum Einsatz.



Abb. 1: Hohe Informationsdichte von 3D-Bilddaten im Straßenraum

1.2 Gliederung des Beitrags

Der erste Teil des Beitrags beschreibt die entwickelte Technologie für die Umsetzung eines 3DCityTV-Services. Darin werden im Detail: a) das mobile Stereobild-basierte Erfassungssystem zur Digitalisierung von Strassen- und Schienenkorridoren, b) der verwendete Datenverarbeitungs-Workflow zur Aufbereitung einer georeferenzierten 3D-Bilddatenbasis und c) die Systemarchitektur für die einfache und verteilte Nutzung der 3D-Bilddatenbasis über das Web, vorgestellt. Im zweiten Teil wird die spezifische Umsetzung und geplante Integration eines 3D-Bilddatendienstes in die Systemumgebung MISTRA (Management Informationssystem Strasse der Schweiz) des Bundesamtes für Strassen diskutiert. Am Beispiel der prototypischen Umsetzung eines 3DRoadTV-Service, werden unterschiedliche Integrations- und Nutzungsmöglichkeiten sowie spezifische Funktionalitäten aufgezeigt, welche mit einem 3D-Bilddatendienst realisiert werden können. Eine kurzes Fazit sowie ein Ausblick über die weiterführenden Entwicklungen schliessen den Beitrag ab.

2 Stereobild-basiertes Mobile Mapping

Die Realisierung eines 3DCityTV-Services erfordert als Datenbasis eine hochaufgelöste georeferenzierte 3D-Bilddatenbasis der geforderten Strassen- oder Schienenkorridore. Für die Realisierung dieser Datenbasis wurde nachfolgendes Stereobild-basiertes mobiles Erfassungssystem und die dazugehörende Datenverarbeitungsprozesse entwickelt. Abbildung 2 zeigt im dafür geschaffenen Workflow die entwickelten Komponenten von der Datenerfassung bis zur Cloud-basierten Nutzung der 3D-Bildatenbasis.

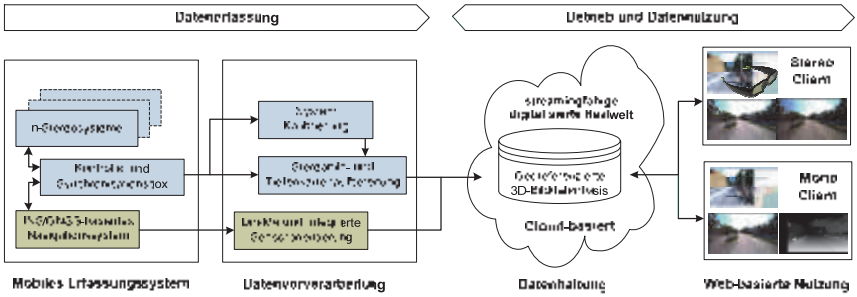


Abb. 2: Komponenten des Datenverarbeitungs-Workflows von der Datenerfassung bis zur verteilten Web-basierten Nutzung

2.1 Mobiles Erfassungssystem

Das entwickelte Stereobild-basierte Erfassungssystem (siehe Abb. 3) besteht aus aus dem Kontroller, der Synchronisationsbox, mit bis zu fünf konfigurierbaren Stereosystemen sowie einem INS/GNSS-Navigationssystem. Die Stereosysteme sind mit Bildsensoren von 2 bis 11 MP bestückt und können mit Bildraten bis 30 bzw. 5 Frames pro Sekunde betrieben werden, was zu sehr dichten Stereobildsequenzen führt – aber auch zu Bilddatenmengen in der Grössenordnung von 1 TB pro Stunde, abhängig von der Aufnahmefrequenz und dem Betriebsmodus. Der Kontroller ermöglicht über die Erfassungssoftware die Konfiguration der Bilddatenaufnahme sowie deren visuelle Prüfung während der Messfahrt. Im Weiteren werden die erfassten Rohbilder mit 12Bit radiometrischer Auflösung parallel auf mehrere Disks aufgezeichnet. Die Stereosysteme können an die Synchronisationsbox angeschlossen und mittels Hardware-Trigger ausgelöst und die Aufnahmen mit dem Navigationssystem synchronisiert werden. Die Komponenten des Erfassungssystems können sehr einfach auf unterschiedliche Fahrzeug- wie auch Schienen-gestützte Trägerplattformen integriert werden.

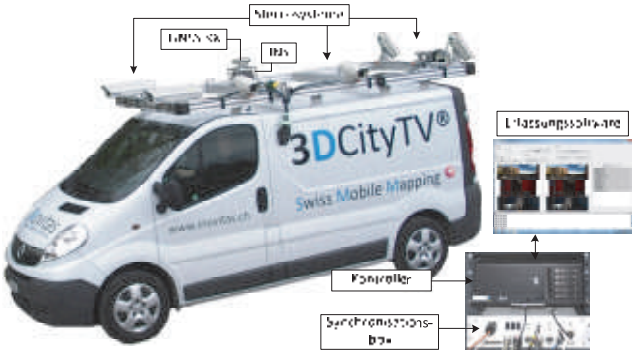


Abb. 3: Stereobild-basiertes Mobile Mapping System mit flexibel konfigurierbaren Stereosystemen

2.2 Datenvorverarbeitung

Die Datenvorverarbeitung transformiert die aufgezeichneten Rohstereobildsequenzen und Navigationsdaten zur verteilt nutzbaren georeferenzierten 3D-Bilddatenbasis. Voraussetzung für die Datenaufbereitung ist die vorgängige Kalibrierung der inneren Orientierungsparameter jeder Stereokamera, der relativen Orientierung jedes Stereosystems, der Orientierungsparameter zwischen den einzelnen Stereosystemen sowie die Bestimmung der entsprechenden Hebelarm- und Fehlausrichtungsparameter zwischen dem Referenzstereosystem und dem Navigationssystem (BURKHARD et. al., 2012). Der entwickelte und im Detail in NEBIKER & EUGSTER (2013) beschriebene Bildverarbeitungsprozess beinhaltet die Generierung von radiometrisch korrigierten, verzeichnungsfreien und normalisierten Stereobildpaaren. Im Anschluss wird mittels Dense Image Matching (OpenCV, 2012) für jedes Stereobildpaar einen zusätzlichen Tiefenkanal berechnet, welcher die Stereobilder in sogenannte 3D-Bilder transformiert. Die Stereobildsequenzen bzw. 3D-Bilder werden mittels Direkter Sensororientierung (DSO) georeferenziert und können bei Bedarf mittels integriertem Sensororientierungsverfahren (ISO) gestützt und verbessert werden (EUGSTER et, al., 2012). Damit die digitalisierten Strassen- und Schienenkorridore verteilt genutzt werden können, werden die aufbereiteten georeferenzierten Bildsequenzen in eine streambare 3D-Bilddatenbasis überführt.

2.3 Verteilte Datennutzung – 3DCityTV-WebSDK

Die 3D-Bilddatenbasis kann wie in Abbildung 2 dargestellt wahlweise über eine interaktive Stereoanwendung sowie im Monobildmodus betrachtet und ausgewertet werden. Im Monobildmodus wird der zusätzliche Tiefenkanal mittels Monoplotting-Verfahren für die geometrische Auswertung der 3D-Bilder verwendet. Über beide Verfahren können Objekte in 3D kartiert, beliebige Masse extrahiert und bestehende Geoobjekte lagekorrekt überblendet werden. Verschiedene Untersuchungen zeigen, dass mit dem System absolute 3D-Messgenauigkeiten im Referenzkoordinatensystem von ca. 3-4 cm und relative Messgenauigkeiten im Bereich von 1 cm erreicht werden (BURKHARD et. al., 2012). Dabei hängt die erreichbare absolute 3D-Positionsgenauigkeit im Wesentlichen von der mittels GNSS bestimmten absoluten Lage der Aufnahmetrajektorie ab. Die 3D-Bilddatenbasis kann sehr einfach über den Software-Werkzeugkasten 3DCityTV-WebSDK über das Web verteilt und flexibel genutzt werden. Die JavaScript basierte und HTML5/WebGL voraussetzende 3DCityTV-WebSDK bietet verschiedenste Methoden für den Zugriff auf die 3D-Bilder, die Auswertung, die Navigation sowie die Integration bestehender Geoobjekte. Wie Abbildung 4 zeigt, abstrahiert die WebSDK die Nutzung der 3D-Bilddatenbasis und ermöglicht damit die einfache Realisierung unterschiedlicher Web- oder Desktop-Applikationen. Sehr einfach lässt sich damit der 3DCityTV-Service in bestehende (Geo-) Informationssysteme einbinden, womit Objekte direkt aus dem bestehenden Informationssystems im visuellen Kontext der Strassen- oder Schienenkorridore angezeigt oder digitalisiert werden können. Weiterführend eignet sich die aufbereitete 3D-Bilddatenbasis für automatische Informationsextraktionsprozesse wie CAVEGN & NEBIKER (2012) anhand der Verkehrssignale zeigt.

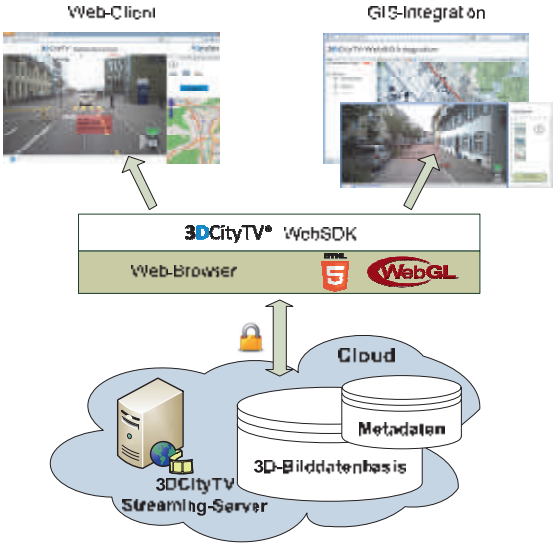


Abb. 4: Nutzung des 3D-Bilddatendienstes über die HTML5/WebGL basierte 3DCityTV-WebSDK

3 3D-Bilddatendienst für Infrastrukturpflege und -verwaltung aus Sicht des Bundesamts für Strassen (ASTRA)

3.1 Die Situation heute

Das Bundesamt für Strassen der Schweiz (ASTRA) ist zuständig für die Planung, den Bau sowie den Betrieb der Strasseninfrastruktur im Eigentum des Bundes (Nationalstrassen). Insgesamt werden ab 2014 rund 2'200km Strasse durch das ASTRA bewirtschaftet.

Zur Durchführung der Prozesse für Projektierung, Bau und Betrieb werden heute eine Vielzahl unterschiedliche Grundlagendaten und Instrumente benötigt. Zur Unterstützung wurde am ASTRA das Managementinformationssystem Strasse und Strassenverkehr eingeführt (MISTRA). Das System steht neben der Bundesverwaltung auch den Kantonen zur Verfügung.

Mit seinen mittlerweile 14 Fachapplikationen, werden in MISTRA eine Vielzahl an Infrastrukturanlagen rund um die Nationalstrassen verwaltet. Weil gute Entscheidungen nur auf Grundlage einer aktuellen und möglichst vollständigen Datenbasis getroffen werden können, stellen die Erfassung und Nachführung solcher umfassenden Datensammlungen eine grosse, aber unverzichtbare organisatorische und finanzielle Herausforderung für alle Beteiligten dar. Nur so kann das Wissen über die Strasseninfrastruktur langfristig erhalten und weitergegeben werden.

Als Raumbezug für die auf Strassen bezogenen Daten wird im MISTRA das auf den Strassenachsen basierte lineare Räumliche Basis-Bezugssystem (RBBS) verwendet. Alle in diesem System lokalisierten Daten können miteinander kombiniert und ausgewertet werden. Zusätzlich wird für die meisten dieser Daten auch eine Lokalisierung im Landeskoor-

dinatensystem ermittelt. Damit können die gleichen Daten auch mit klassischen GIS-Werkzeugen analysiert und ausgewertet werden. Als effizientes Mittel für die Datenerhebung wie auch für die Situationsanalyse werden seit einigen Jahren vermehrt Befahrungsbilder meist in Form von Monovideos eingesetzt. Die Befahrungsbilder werden durch spezialisierte Firmen erfasst und dem ASTRA in aufbereiteter Form zur Verfügung gestellt. Die Befahrungsbilder sind mittels GNSS verortet und ermöglichen meist einzelne Objekte grob zu vermessen (Länge, Breite, Höhe). Sehr oft kann auf diese Daten nicht verteilt zugegriffen und die Sachdaten können nicht im RBBS referenziert werden. Neben den Befahrungsbildern wird oft für die Erstellung von Projektierungsgrundlagen mobiles Laserscanning eingesetzt. Die Nutzung der Bild- und Videodaten ist heute allerdings noch nicht offizieller Teil von standardisierten Arbeitsprozessen im ASTRA. Bei ihrer Verwendung gehen die Nutzer sehr individuell vor und lassen sich von kommerziellen (Heim-) Anwendungen wie Google-Streetview inspirieren.

3.2 Nutzen eines 3D-Bilddatendienstes für die Nationalstrassen

3.2.1 Projektierung, Bau, Unterhalt und Betrieb

Der vermutlich größte Kundenkreis für Befahrungsbilder findet sich vor Ort in den Filialen, den im Auftrag des ASTRA arbeitenden Gebietseinheiten, sowie in der Abteilung Infrastruktur. Fotos und Befahrungsbilder dienen als vielseitiges und allgemeines Informations- und Kommunikationsinstrument innerhalb der Organisation, wie aber auch gegenüber Dritten.
Insbesondere im Zuge einer Projektidee oder Projektierung sind aktuelle, qualitativ gute und hinsichtlich der Fragestellung vollständige Bilddaten zentral, da gerade erste Entscheide bereits im Büro getroffen und auf eine Begehung vor Ort verzichtet werden kann. Begehungen insbesondere auf Hochleistungsstrassen sind oft mit aufwändigen Sicherungsmassnahmen verbunden, was neben dem Zeitaufwand auch hohe Kosten verursacht.
Darüber hinaus dienen sie der Situationsbeurteilung oder als Basis für die Kartierung von Objekten entlang des Trassees, von Kunstbauten, der Verkehrssignalisation, von Verkehrs-rückhaltesystemen, Lärmschutzwänden sowie diverser Strasseinventarobjekte.

3.2.2 Erfassung und Nachführung von Geodaten

Die Verwaltung von Infrastrukturen mittels GIS-unterstützer Informationssysteme ist mittlerweile Standard bei grossen Infrastrukturbetreibern. Dabei sind gewaltige GIS-Datensammlungen entstanden, die gepflegt werden müssen. Erfassung und Nachführung der Daten stellen häufig den teuersten Kostenblock in einem IT-Projekt und damit grosse organisatorische und finanzielle Risiken dar.
Häufig werden die Erfassungs- und Nachführungsarbeiten heute noch manuell und nach individuellen Methoden durchgeführt, mit entsprechenden Investitionen in Ausbildung (Digitalisierung), Hardware (spezielle Erfassungsgeräte) und Sicherheit. Jeder Ansatz diese Aufgabe effizienter, risikoloser und damit beherrschbarer zu machen, wird daher begrüsst.
Zur Verbesserung der Situation wird ein System benötigt, das mindestens semi-automatisierte 3D-Objekterfassung aus Bilddaten unterstützt und die Daten einfach in die vorhandenen MISTRA-Datenerfassungs- und Bearbeitungsumgebung überführt werden kann (ESRI ArcGIS). Die erfassten bzw. korrigierten Objekte sollen eine akzeptable Qualität aufweisen (Korrektheit

und Vollständigkeit) und Nacherfassungen und Kontrollen sollten sich auf ein notwendiges Mindestmass reduzieren lassen.
Unterstützung bieten soll das System, neben der Datenersterfassung, hauptsächlich bei laufenden (Lage- und Attribut-) Korrekturen an Achsen und (Infrastruktur) Inventarobjekten (Ausschilderung, Kanaldeckel, Leitplanken, Beleuchtung etc.)

3.3 Konzeption und Anforderungen

Aus den skizzierten Anwendungsfeldern ergeben sich an einen zeitgemässen 3D-Geobilddatendienst folgende Basisanforderungen:

- Hochaufgelöstes digitales Abbild in Form von georeferenzierten 3D-Bildsequenzen der Fahrbahnkorridore und deren verteilte, performante Nutzung über das Web im Kontext unterschiedlicher Applikationen und Anwendergruppen.
- Möglichkeit zur lagekorrekten Überlagerung von Geodaten vorliegend im RBBS sowie in Landeskoordinaten in die Bilddaten.
- 3D-Kartierung beliebiger Geoobjekte entlang des Fahrbahnkorridors dual im RBBS wie auch in Landeskoordinaten und Extraktion beliebiger relativer Masse wie Höhenunterschiede, Distanzen und Flächen im Genauigkeitsbereich von 5-10 cm.
- Navigation in den Bildsequenzen über ein bestehendes Karteninterface oder im erweiterten Modus über definierte Strassenachsen.
- Zur Verfügung stellen geeigneter Informationen an den Nutzenden über das User-Interface über Qualität der extrahierten 3D-Information, Zuverlässigkeit und Aktualität der Befahrungsdaten.
- Software-Werkzeugkasten für die Nutzung des 3D-Bilddienstes in MISTRA-Fachapplikationen oder weiteren Anwendungen auf unterschiedlichen Endgeräten.

Weiterführend soll die Funktionalität generell für beliebige manuelle oder automatisierte Nachführungs- und Auswertungsprozesse eingesetzt werden können. Insbesondere auf Basis mehrerer verfügbarer Befahrungszeitstände sind zudem Veränderungsanalysen denkbar.

3.4 Prototyp 3DRoadTV-Service

Anhand mehrerer über die Schweiz verteilt befahrener Pilotstreckenabschnitte erfolgte eine prototypische Realisierung eines ASTRA 3D-Videos Dienstes bezeichnet als 3DRoadTV-Service. Ziel war die Validierung der erarbeiteten Konzepte und die Überprüfung der technischen Machbarkeit eines solchen Bilddatendienstes. Nachfolgende Abbildung 5 zeigt vereinfacht die gewählt Systemarchitektur und die realisierten Komponenten. Die aufbereitete 3D-Bilddatenbasis der Pilotstreckenabschnitte wird mit der 3DCityTV-Server Komponente in der Cloud betrieben und steht für die Nutzung über unterschiedliche Anwendungen zur Verfügung. Die 3D-Bilddatenbasis besteht aus streamingfähigen 3D-Bildern und der zusätzlich benötigten 3DCityTV-Metadatenbank. Der Zugriff auf diese Datenbasis erfolgt über die in Abschnitt 2.3 eingeführte 3DCityTV-WebSDK. Für die Umsetzung dieses Prototyps mit Hilfe der 3DCityTV-Serverkomponenten und WebSDK wurden zusätzliche Datenstrukturen und Methoden implementiert, welche insbesondere das duale Arbeiten mit Geoobjekten im

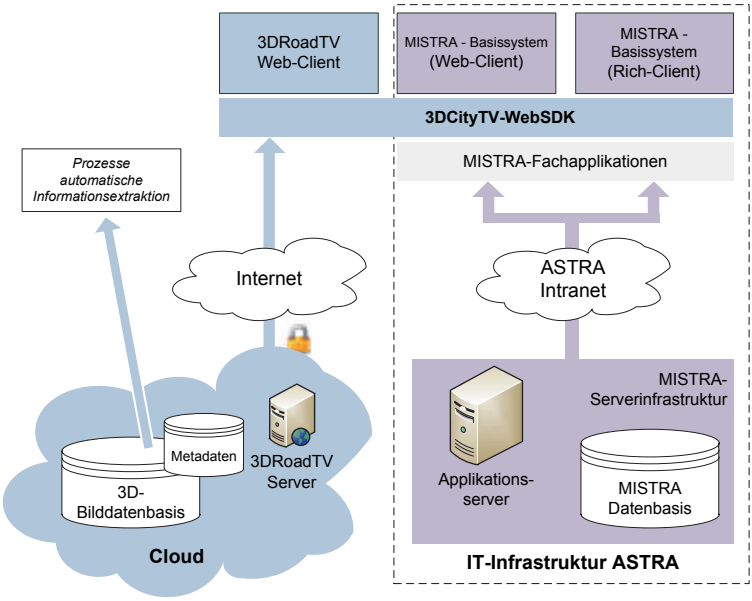


Abb. 5: Architektur 3DRoadTV-Service

linearen- (RBBS) und planaren Bezugssystem (Landeskoordinaten) ermöglichen. Basierend auf dieser SDK wurde als erstes der Prototyp 3DRoadTV Web-Client implementiert mit welchem die einzelnen Funktionen und Einsatzmöglichkeiten der SDK aufgezeigt und getestet werden können. Abbildung 6 zeigt die Benutzeroberfläche dieser Applikation mit welcher kartiert, bestehende Geodaten lagekorrekt überlagert oder diverse Masse im Fahrbahnbereich extrahiert werden können. Über die Karte kann in der 3D-Bilddatenbasis navigiert und auf die unterschiedlichen Stereosysteme gewechselt werden. Weiter wird der 3DRoadTV-Service mit Hilfe der WebSDK direkt ins MISTRA-Basissystem integriert. Die Integration hat das Ziel direkt Fachdaten über die WebSDK im Videofenster darzustellen bzw. kartierte Fachdaten direkt im Basissystem zu erfassen. Im Rahmen weiterer Untersuchungen wurden unterschiedliche Informationsextraktionsprozesse auf Basis der zur Verfügung stehenden 3D-Bilddatenbasis getestet, um die Funktionstüchtigkeit der automatisierten Fachdatenkartierung und -nachführung auf Basis der Befahrungsdaten zu verifizieren.



Abb. 6: Benutzeroberfläche Prototyp 3DRoadTV Web-Client

4 Fazit und Ausblick

Die vorgestellte Stereobild-basierte mobile Erfassungstechnologie liefert sehr dichte, georeferenzierte Stereobildsequenzen mit einer geometrischen Genauigkeit im Sub dm-Bereich. Die erfassten Bilddaten können mittels Stereobild-basierter Prozessierungskette in eine navigierbare, streamingfähige 3D-Bilddatenbasis verwandelt und in der Cloud betrieben werden. Mit der 3DCityTV-WebSDK lässt sich diese 3D-Bilddatenbasis vollständig webbasiert nutzen und in Dritt-Informationssysteme integrieren. Dabei erlauben die hoch aufgelösten Messbilder mit ihren Pixelgenau co-registrierten Tiefenkarten via 3D-Monoplotting eine unvergleichbar bedienerfreundliche und gleichzeitig präzise 3D-Messung.

Erste Erfahrungen mit dem prototypischen Aufbau eines 3D-Bilddatendienstes verlaufen vielversprechend. Neben passenden Auswertungs-Werkzeugen sind den Nutzern vor allem eine unkomplizierte Bedienung und eine flüssige virtuelle Fahrt wichtig. Die Nutzer müssen in der Lage sein, die Vorteile der 3D-Bilder für ihre Arbeiten selbstständig und barrierefrei entdecken und damit deren Mehrwert für das ASTRA (weiter-) entwickeln zu können.

Da der angestrebte 3D-Bilddatendienst als breit einzubindender Webservice vielen Fachapplikationen gleichermassen zur Verfügung stehen soll, ist es wichtig, dass seine technische Integration einfach und flexibel passieren kann.

Für Erfolg und Akzeptanz mitentscheidend ist ein überzeugendes Wartungs- und Pflegekonzept des Services einerseits sowie der Befahrungsdaten anderseits. Aufgrund vielversprechender Kosten-Nutzenschätzungen evaluiert das Projekt MISTRA in diesem Zusammenhang eine Cloud-Lösung. Das ASTRA erwartet sich durch die Nutzung des geplanten 3D-Bilddatendienstes mit ihren neuen intelligenten Auswertemöglichkeiten einen spürbaren Effizienzsprung für viele Bereiche. Je aktueller, präziser und besser beschrieben (vertrauenswürdiger) die

Videoinformationen zukünftig zur Verfügung gestellt werden können, desto öfter werden sich teure Vorort-Inaugenscheinnahmen vermeiden lassen (z.B. im Rahmen von Vorabklärungen) und desto schneller können Informationen daraus gezogen werden, die zu besseren Entscheidungen führen werden.

5 Literaturverzeichnis

BURKHARD, J. et al., 2012: Stereovision Mobile Mapping: System Design And Performance Evaluation. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXIX-B5, pp.453–458.

CAVEGN, S. & NEBIKER, S., 2012: Automatisierte Verkehrszeichenkartierung aus mobil erfassten Stereobilddaten unter Verwendung der Tiefeninformation aus Dense-Stereo-Matching. Photogrammetrie - Fernerkundung - Geoinformation, 2012(5), pp.631–645.

EUGSTER, H. et al., 2012: Integrated georeferencing of Stereo Image Sequences Captured with a Stereovision Mobile Mapping System – Approaches and Practical Results. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXIX-B1, pp.309–314.

NEBIKER, S. & EUGSTER, H., 2013: 3DCityTV – Cloud-basierte Stereovision Mobile Mapping Services, 17. Internationale Geodätische Woche, 17.-23. Feb 2013, Obergurgel (A)

LIPPMAN, A., 1980: Movie-maps: An application of the optical videodisc to computer graphics. ACM SIGGRAPH Computer Graphics, 14(3), pp.32–42.

NOVAK, K., 1993: Real-Time Mapping Technology. In: International Archives of Photogrammetry and Remote Sensing. pp. 569–575.

OPENCV 2012: OpenCV V2.1 Documentation. Camera Calibration and 3d Reconstruction. StereoSGBM. – http://opencv.willowgarage.com/documentation/cpp/ camera_calibration_and_3d_reconstruction.html#stereosgbm (22.11.2012).

SCHARSTEIN, D. & SZELISKI, R., 2012: Middlebury Stereo Evaluation – Version 2. – http://vision.middlebury.edu/stereo/eval/ (22.11.2012).

SCHWARZ, K.P. et al., 1993: VIASAT - A mobile highway survey system of high accuracy. In Vehicle Navigation and Information Systems Conference, 1993. Proceedings of the IEEE-IEE. pp. 476–481.

Extraktion von senkrechten Fassadenebenen aus 3D-Punktwolken von Schrägluftbildern

MAGDALENA LINKIEWICZ[1]

Zusammenfassung: In dieser Arbeit wird eine Methode zur automatischen Extraktion von Fassadenebenen von Gebäuden aus einer hochaufgelösten 3D-Punktwolke von Schrägluftbildern vorgestellt. Dabei werden Methoden der lokalen Regression in der 2D-Ebene verwendet. Aus der lokalen Regression werden die lokale Richtung der Punktwolke sowie ein Maß für die Linearität der lokalen Punktwolke bestimmt. Anhand dieser Daten wird die 3D-Punktwolke nach Fassadenzugehörigkeit zu segmentiert. Weitere Algorithmen erzeugen aus der segmentierten Punktwolke den Gebäudegrundriss als geschlossenen Polygonzug. Die verwendeten Algorithmen wurden in MatLab implementiert und für Gebäude mit unterschiedlicher Komplexität angewendet.

1 Einleitung

In der Abteilung Sensorkonzepte und Anwendungen in der Einrichtung Optische Informationssysteme des Instituts für Robotik und Mechatronik am Deutschen Zentrum für Luft- und Raumfahrt e. V. (DLR) wird seit dem Jahr 2009 das Modulare Luftbild-Kamerasystem MACS (Modular Airborne Camera System) entwickelt (LEHMANN et al. 2011) (Abb 1), welches in einer Ausführung Obliqueaufnahmen (Schrägbildaufnahmen) bereitstellt. Im Gegensatz zum herkömmlichen Ansatz von senkrecht montierten Kameras haben Obliqueaufnahmen den Vorteil, dass Fassaden gut sichtbar sind.

Das Ergebnis der Prozessierung der Obliqueaufnahmen des MACS-Systems ist unter anderem eine hochaufgelöste 3D-Punktwolke (50 Punkte/m^2 auf der Fassade). Diese Punktwolke soll als Grundlage für die Entwicklung von Methoden zur Erstellung von photogrammetrischen Produkten, wie zum Beispiel einem 3D-Stadmodell, das dem Bedarf der Anwender auf dem Markt entspricht und die aktuellen Standards von LoD3 erfüllt, dienen.

In der hier entwickelten Methode werden Algorithmen zur automatischen Extraktion der senkrechten Fassadenebenen von Gebäuden aus der 3D-Punktwolke von Schrägluftbildern entwickelt. Ziel ist:

- eine hohe geometrischen Genauigkeit von unter 0,5 m ,
- die Bestimmung der Lage und Höhe einzelner Fassaden,
- die Zusammenfassung von Fassaden eines Gebäudes zu dem semantischen Objekt Gebäude.

[1] Magdalena Linkiewicz, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Abteilung Sensorkonzepte und Anwendungen, Rutherfordstr. 2, 12489 Berlin.

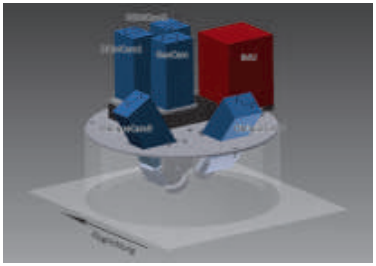


Abb 1: Aufbau des Kamerasystems MACS-Oblique. Quelle: DLR, Bearbeitung (KONECNY 2011).

2 Datengrundlage

Die Befliegung wurde alternierend mit einer Längsüberlappung von ca. 90% und einer Querüberlappung von ca. 50% durchgeführt. Diese Aufnahmekonstellation hat die Aufnahme von Luftbilddaten aus vier Perspektiven gewährleistet. Die Aerotriangulationen werden im Modul MATCH-AT der Software Inpho von Trimbel durchgeführt. Ein Oberflächenmodell wurde mit Hilfe des SGM-Algorithmus (Semi-Global Matching) erstellt (HIRSCHMÜLLER 2008, S. 328ff.).

3 Richtung und Maß für Linearität durch lokale Regression

In dieser Arbeit sollen die Fassadenebenen durch lineare Regression bestimmt werden. Die Überlegungen gehen dahin, die Punktwolke in kleine räumliche Bereiche aufzuteilen. Dieser Ansatz ist in der Literatur bekannt. Eine solche Methode der lokalen Regression haben beispielsweise Hoppe (HOPPE et al. 1992) und Mitra (MITRA et al. 2004) entwickelt. Bei Hoppe wird für jeden Punkt x_i aus einer Punktwolke eine Anzahl von k-nächstliegenden Punkten, als k-Nachbarschaft definiert. Für diese Punkte wird die Regression nach Pearson durchgeführt. Mitra entwickelt die Methode von Hoppe weiter, indem er optimale Werte für k findet. Die Methode zur Bestimmung der Fassadenebene soll nicht davon abhängen, in welchem Winkel eine Fassade im übergeordneten Koordinatensystem liegt. Wünschenswert ist eine Methode zur Bestimmung der Ausgleichsgerade, die vom Koordinatensystem unabhängig ist.
Zur Detektion der Fassadenrichtung werden die 3D-Punkte auf die zweidimensionale xy-Ebene projiziert. Die Ausgleichsgerade nach Pearson (PEARSON 1901) ergibt sich wie folgt.
Die Kovarianzmatrix für eine 2D-Punktwolke $p = \{p_i = (x_i, y_i) \in \mathbb{R}^2, i = 1, \ldots, n\}$ ist eine 2x2-Matrix:

$$M_{cov}(p) := \begin{pmatrix} \mathrm{cov}(x,x) & \mathrm{cov}(x,y) \\ \mathrm{cov}(x,y) & \mathrm{cov}(y,y) \end{pmatrix}.$$

Der zum größten Eigenwert zugehörige Eigenvektor gibt die Richtung der Ausgleichsgerade an. Diese verläuft durch den Schwerpunkt der Punktwolke.

Die Quadratwurzeln der Eigenwerte entsprechen der Standardabweichung in Richtung des jeweiligen Eigenvektors. Das Verhältnis der kleineren zur größeren Standardabweichung ist ein Maß für die Linearität der Punktwolke:

$$Q = \frac{\sigma_{min}}{\sigma_{max}}.$$

Es ist bekannt, dass das Rauschen einen negativen Einfluss auf die lineare Regression hat. Für die Extraktion von Fassadenrichtungen haben die Dach- und Bodenpunkte einen ähnlichen Einfluss wie Rauschen, da sie außerhalb der Fassade liegen.

Das Ziel der nachfolgend vorgestellten Filterung besteht darin, die Boden- und Dachpunkte, sowie alle weiteren unerwünschten Objekte aus der 3D-Punktwolke zu entfernen (Abb. 2), so dass nur die Fassadenpunkte übrigbleiben. Dazu wird die Punktwolke in ein quadratisches Gitter in der xy-Ebene aufgeteilt. Für jede Zelle des Gitters werden diese mit den 8 benachbarten Zellen (3x3 Umgebung) zu einem sogenannten Fenster zusammengefasst. Einfache Überlegungen zeigen (LINKIEWICZ 2012), dass für Zellen, die Fassadenpunkte enthalten, ab einer bestimmten Fenstergröße, im Histogramm der z-Werte ein typisches Muster mit zwei deutlichen Maxima am unteren bzw. oberen Rand zu erwarten ist. Diese Maxima werden von den Boden bzw. Dachpunkten verursacht. Die Schwellenwerte zum lokalen Entfernen der Boden- und Dachpunkte ergeben sich durch den Wert der Histogrammklassen, die zu den Maxima benachbart sind. Weiterhin wird gefordert, dass der Abstand der Schwellenwerte einen Mindestbetrag überschreitet, um beispielsweise 3D-Punkte auf Vegetation zu entfernen. Diese Überlegungen gelten für Flachdächer, wie sie im Untersuchungsgebiet auftreten.

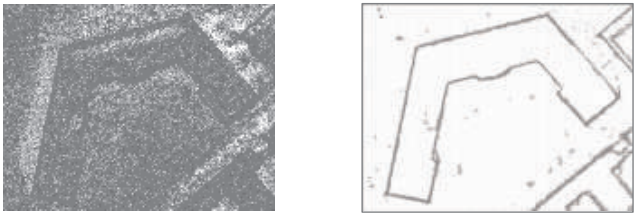


Abb. 2: links: ursprüngliche 3D-Punktwolke; rechts: die 3D-Punktwolke nach dem z-Filter. Quelle: links: DLR.

Für die Segmentierung nach Fassadenzugehörigkeit wird die gefilterte Punktwolke durch ein quadratisches Gitter in der xy-Ebene aufgeteilt. Für jede Zelle des Gitters ist das Regressionsfenster durch die 3x3-Umgebung der benachbarten Zellen definiert. Für jedes Regressionsfenster wird die Regression nach Pearson lokal durchgeführt und so die lokale Fassadenrichtung und das Maß für die Linearität Q bestimmt.

Dies wird in Abb. 3 an einem Ausschnitt aus dem Gebäude B4 dargestellt. Dieser zeichnet sich durch eine kurze Wand aus. Die Breite des Regressionsfensters beträgt hier 4,5 m. Es werden die Ergebnisse der Regression mit nicht überlappenden Regressionsfenstern (links) und mit Verwendung der 3x3-Umgebung dargestellt (Mitte). Bei der Regression unter Verwendung einer 3x3-Umgebung überlappen sich die einzelnen Regressionsfenster und bedecken die Punktwolke dichter, als bei der Regression ohne überlappende Regressionsfenster. Die Überlappung der Regressionsfenster hat den Vorteil, dass Punkte, die in einem Regressionsfenster ungünstig

liegen, beispielsweise wenn sich die Fassade nur zu einem kleinen Teil im Regressionsfenster befindet, in einem benachbarten Regressionsfenster günstig liegen können. Dargestellt werden die Richtungen der Ausgleichsgerade, in Abhängigkeit vom Wert des Koeffizienten ▫ im Intervall [(0, 0.3) – rot, [0.3, 0.5] – violett, (0.5, 1] – schwarz. Weiterhin wird mit dünnen grauen Linien die Zuordnung der lokalen Ausgleichsgerade zur Zelle dargestellt. Im rechten Bild wurde die Regression mit 3x3-Umgebung durchgeführt, aber nur für Zellen mit einer Mindestanzahl von Punkten. Damit wird erreicht, dass die Regression nur berechnet wird, wenn die Punktwolke in etwa durch die Mitte der Zelle führt.

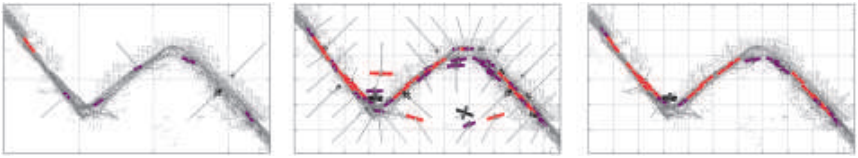


Abb. 3: Einfluss der Verwendung der Umgebung auf die lokale Regression. links) nicht überlappende Regressionsfenster, Mitte) überlappende Regressionsfenster, rechts) reduzierte Anzahl von überlappenden Regressionsfenstern.

Unter der Bedingung einer Mindestanzahl von Punkten werden die besten Ergebnisse erzielt. Am hier dargestellten Beispiel ist ersichtlich, dass das Regressionsfenster nicht zu groß gewählt werden darf, da sonst kürzere Fassaden nicht extrahiert werden. Es darf auch nicht zu klein gewählt werden, da sonst das Verhältnis der minimalen und maximalen Standardabweichung zu groß wird und die Fassadenrichtung nicht zuverlässig detektiert wird.

Die lokale Regression wird für im Folgenden unter Verwendung der 3x3-Umgebung als Regressionsfenster für alle Zellen mit einer Mindestanzahl von Punkten durchgeführt. Zellen bei denen der Quotient aus minimaler und maximaler Standardabweichung den Schwellenwert 0,3 unterschreitet, werden als Fassadenzellen bezeichnet. Zellen, die keine Fassadenzellen sind, sowie deren Punkte, werden für den weiteren Verlauf der Arbeit ignoriert. In Abb. 4 sind lokalen Ausgleichsgeraden für das Gebäude B2 farbig dargestellt. Rot dargestellte Ausgleichsgeraden entsprechen Fassadenzellen. Es wird deutlich, dass fast der gesamte Gebäudeumriss von Fassadenzellen abgedeckt ist. Nur im Bereich der kurzen Fassade ist der Koeffizient ρ größer als 0,3.

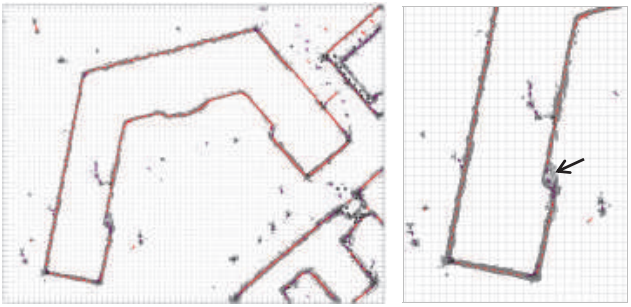


Abb. 4: Ergebnisse der lokalen Regression für Gebäude B2. Pfeil markiert nicht detektierte Fassade.

4 Weitere Algorithmen

In diesem Kapitel wird ein mehrstufiger Ansatz beschrieben, um Fassaden als geschlossenen Polygonzug mit zusätzlichen Informationen über die untere und obere Fassadenkante zu extrahieren. Die Abfolge aller verwendeten Algorithmen werden am Beispiel des Gebäudes B2 illustriert (Abb. 5). Der mehrstufige Ansatz umfasst folgende Schritte:

- Algorithmus A: Initialisierung von Gruppen von Fassadenzellen. Dabei wird eine Fassadenzelle einer Gruppe von Fassadenzellen hinzugefügt, wenn die Richtung der lokalen Ausgleichsgerade ähnlich der Richtung der Gruppe von Fassadenzellen liegt. Die Richtung der Gruppe von Fassadenzellen wird als Mittelwert der zuerst hinzugefügten Zellen dieser Gruppe bestimmt. Wenn es keine benachbarte Fassadenzelle gibt, wird eine neue Gruppe von Fassadenzellen erstellt.
- Algorithmus B: Zusammenfassen von Gruppen von Fassadenzellen. Wenn zwei Gruppen von Fassadenzellen benachbarte Zellen enthalten und die Richtungen der Gruppen von Fassadenzellen ähnlich sind, werden die Fassadenzellen der beiden Gruppen vereinigt.
- Algorithmus C: Regression pro Gruppe von Fassadenzellen.
- Algorithmus D: Bestimmung von Fassadenstrecken durch Berechnung von Anfangs- und Endpunkten der Ausgleichsgeraden pro Fassade. Dazu werden die Punkte der entsprechenden Fassadenzellen auf die Ausgleichsgerade projiziert. Die beiden äußersten Punkte definieren die Fassadenstrecke.
- Algorithmus E: Bestimmung von Nachbarschaftsbeziehungen zwischen Fassadenstrecken. Die Nachbarschaftsrelation verbindet das Ende einer Fassadenstrecke mit dem Ende einer anderen Fassadenstrecke. Dabei wird der minimale Abstand zu allen anderen Enden von Fassadenstrecken verwendet.
- Algorithmus F: Vereinigung von Fassadenstrecken. Benachbarte Fassadenstrecken mit ähnlicher Richtung werden vereinigt. Dabei bleiben die äußeren Enden der Fassadenstrecken erhalten. Die inneren werden gelöscht. Die Vereinigung wird nur durchgeführt, wenn der parallel gemessene Abstand der Geraden klein genug ist.
- Algorithmus G: Berechnung von Fassadenecken. Benachbarte Fassadenstrecken werden bis zu ihrem gemeinsamen Schnittpunkt verlängert, wenn ihr Schnittwinkel groß genug ist.
- Algorithmus H: Zusammenfassen von benachbarten Fassadenstrecken zu einem Gebäude. Durch die Nachbarschaftsrelation kann eine Schleife über alle benachbarten Fassadenstrecken durchgeführt werden. Ist die Schleife geschlossen, dann wird das Gebäude durch einen geschlossenen Polygonzug dargestellt. Ansonsten wird das Gebäude fehlerhaft durch einen offenen Polygonzug dargestellt.
- Algorithmus I: Berechnung der unteren und oberen Fassadenkanten. Dabei werden die z-Werte der Punkte der vollen 3D-Punktwolke mit einem maximalen Abstand von der Fassadenstrecke analysiert. Nach Abzug einer Quantile von z.B. 1% ergibt sich die Höhe der unteren und oberen Fassadenkanten. Die Höhe des Gebäudes wird als Median dieser Werte bestimmt.

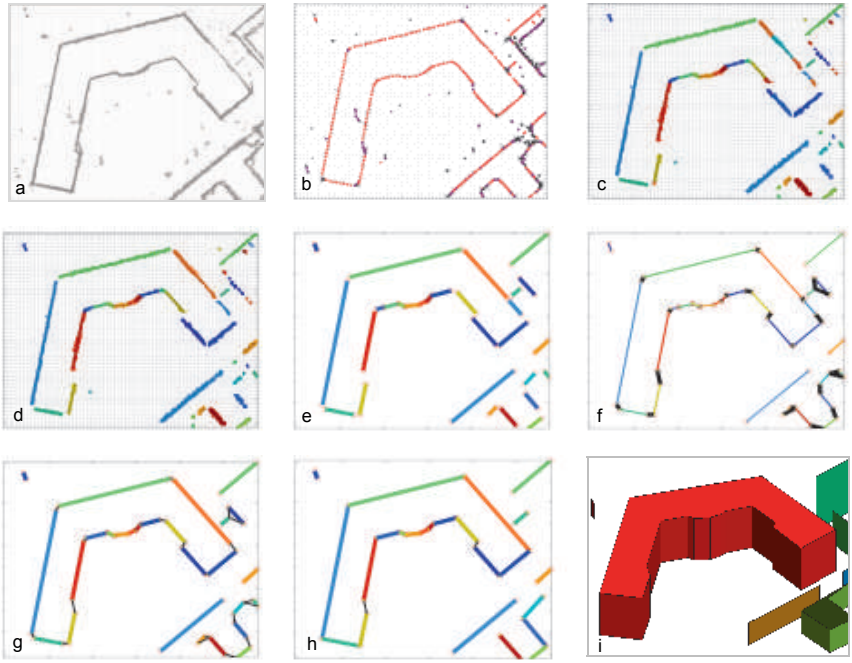


Abb. 5: Folge von Algorithmen zur Fassadenextraktion, a) gefilterte Punktwolke, b) Richtungen aus lokaler Regression, c) Initialisierung von Gruppen von Fassadenzellen (Alg. A), d) Zusammenfassen von Gruppen von Fassadenzellen (Alg. B), e) Fassadenstrecken durch Regression (Alg. C,D), f) Nachbarschaftsbeziehungen (Alg. E), g) Vereinigung von Fassadenstrecken (Alg. F), h) Fassadenecken (Alg. G), i) Gebäude(Alg. H,I).

5 Ergebnisse

In den oben beschriebenen Algorithmen werden verschiedene Parameter verwendet. Deren Werte werden auf Grund der Untersuchungen an den Gebäuden B1, B2, B3 und B4 gewählt. Bei einer festen Kombination von voreingestellten Parametern werden bis auf fünf Fassaden alle Fassaden der vier getesteten Gebäude erkannt. Besonders gute Ergebnisse werden für Gebäude mit hohen Fassaden erzielt, die nicht durch Vegetation oder Anbauten verschattet werden. Problematisch sind kurzen Fassaden oder Fassaden, bei denen die Punktwolke durch Anbauten eine starke Abweichung von der Fassadenebene hat. Mit individuell angepassten Einstellungen konnten die fehlenden Fassaden (bis auf eine) erkannt werden. Die extrahierten Fassadenstrecken wurden anhand ihrer Koordinaten in ArcMap 10 digitalisiert. Ein Vergleich des Nadir-TrueOrthobildes mit den extrahierten Fassadenstrecken zeigt eine gute Übereinstimmung.

Im Gebäude B2 werden fast alle auf die xy-Ebene projizierten Fassadenebenen extrahiert (Abb. 6). Eine Ausnahme stellen die zwei kurzen Fassaden dar, die als schwarze Strecken dargestellt werden. Die horizontale Länge dieser Fassaden beträgt 3 m. Sie werden nicht extrahiert, da das

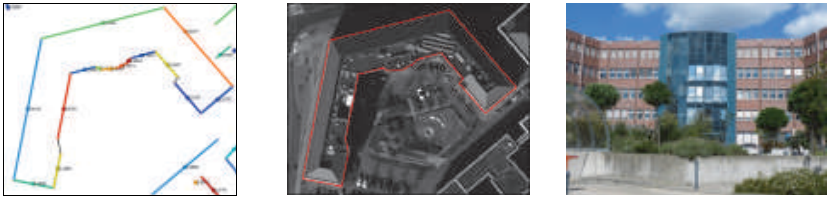


Abb. 6: Ergebnisse für Gebäude B2.

Regressionsfenster eine Breite von 4,8 m hat, was die Extraktion von Fassadenstrecken, die kürzer als 4,8 m sind, verhindert. Trotzdem ist es hier, dank der Herstellung von Nachbarschaftsbeziehungen zwischen den Fassadenstrecken gelungen, einen geschlossenen Polygonzug von Fassadenstrecken zu erstellen. Die Standardabweichung der Abstände der gefilterten Punkte zu den Fassadenstrecken betragen 0.19 – 0.31 m.
Bei Gebäude B1 wurden alle Fassaden detektiert (Abb. 7). Die Breite der Fassade auf dem Nadir-TrueOrthobild beträgt 5 Pixel. Bei einer Bodenauflösung 0,15 m und Berücksichtigung der diagonalen Lage ergibt sich eine absolute Breite von 0,85 m. Die relative Lagegenauigkeit der extrahierten Fassade für das Gebäude B1 kann also mit mindestens 0,425 m angegeben werden.

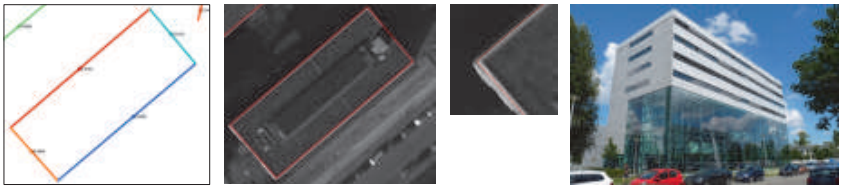


Abb. 7: Ergebnisse für Gebäude B1.

Bei dem Gebäude B4 handelt es sich um einen Komplex von einem niedrigen und einem hohen Gebäude mit einem Verbindungsgang (Abb. 8). Nicht extrahiert wurden niedrige und verdeckte Fassaden, sowie Fassaden mit Anbauten.
Zur Bewertung der Anwendbarkeit der entwickelten Algorithmen für praktische Berechnungen soll die Rechenzeit für ein Beispiel analysiert werden. Für das Gebäude B2 aus ca. 2.200.000 Punkten beträgt die Berechnungszeit 78 Sekunden. Die gefilterte Punktwolke enthält dabei ca. 330.000 Fassadenpunkte. Eine Analyse der Verteilung der Rechenzeit zeigt, dass diese

hauptsächlich durch das Hinzufügen von Elementen zu Arrays verursacht werden. Die vergleichsweise hohe Rechenzeit beim Erweitern von Arrays ist eine Eigenschaft von MatLab.
Die Berechnung der lokalen Regression und die Bestimmung der Fassadenstrecken als geschlossenes Polygon benötigen mit 5,3 Sekunden ca. 6% der Rechenzeit auf einer CPU.

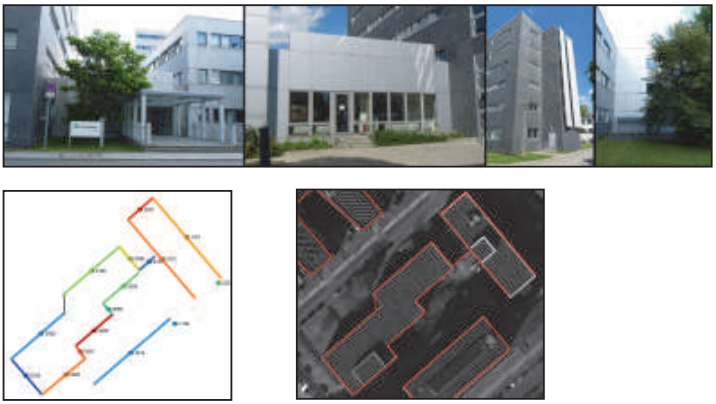


Abb. 8: Ergebnisse für Gebäude B4. Nicht alle Fassaden wurden extrahiert.

6 Fazit

Die Ergebnisse dieser Arbeit zeigen, dass eine effiziente, vollautomatische Extraktion von Fassadenebenen durch lokale Regression für hochaufgelöste 3D-Punktwolken möglich ist. Allerdings hängt deren Qualität von der Komplexität der Gebäudegrundrisse und der Qualität der Punktwolke ab. Eine fehlerfreie Extraktion ist mit der in dieser Arbeit entwickelten Methodik nicht zu erwarten. Die entwickelte Methode wurde auf frei stehende Gebäude mit Flachdächern angewendet. Eine Anwendung in eng bebauten städtischen Gebieten steht aus.
Mit der entwickelten Methode werden sowohl planare senkrechte Fassaden, als auch gekrümmte senkrechte Fassaden, automatisch extrahiert. Letzte werden durch mehrere Fassadenebenen approximiert. Alle extrahierten Fassadenstrecken werden sortiert und als geschlossener Polygonzug exportiert. Als Ergebnis wird der Grundriss der Gebäude im zweidimensionalen Koordinatensystem erzeugt. Außerdem werden die absoluten Höhen der unteren und der oberen Fassadenkante eines jeden Gebäudes im dreidimensionalen Koordinatensystem bestimmt. Besonders gute Ergebnisse werden für Gebäude mit hohen Fassaden erzielt, die nicht durch Vegetation oder Anbauten verschattet werden. Problematisch sind kurze Fassaden oder Fassaden, bei denen die Punktwolke durch Anbauten eine starke Abweichung von der Fassadenebene hat.

7 Literaturverzeichnis

HIRSCHMÜLLER, H., 2008: Stereo Processing by Semi-Global Matching and Mutual Information: IEEE Transactions on Pattern Analysis and Machine Intelligence, Band 30, Nr. 2, 2008, S. 328-341.

HOPPE, H., DEROSE, T., DUCHAMP, T., MCDONALD, J. & STUETZLE, W., 1992: Surface Reconstruction from Unorganized Points: ACM SIGGRAPH 1992 Proceedings, S. 71-78, Online: http://research.microsoft.com/en-us/um/ people/hoppe/recon.pdf [Stand: 29.08.2012].

KONECNY, K., 2011: Evaluierung der Nutzbarkeit von Punktwolken aus Schrägluftbildern zur Kartierung von Grundrisselementen: Masterarbeit, Beuth Hochschule für Technik Berlin, Berlin.

LEHMANN, F., BERGER, R., BRAUCHLE, J., HEIN, D., MEISSNER, H., PLESS, S., STRACKENBROCK, B. & WIEDEN, A., 2011: MACS – Modular Airborne Camera System for generating photogrammetrichigh-resolution products: Photogrammetrie, Fernerkundung, Geoinformation. Nummer 6/2011, E. Schweizerbartsche Verlagsbuchhandlug, Stuttgart, S. 423-434.

LINKIEWICZ, M., 2012: Extraktion von senkrechten Fassadenebenen aus 3D-Punktwolken von Schrägluftbildern.- Masterarbeit, Beuth Hochschule für Technik Berlin, Berlin.

MITRA, N. & NGUYEN, A., 2004: Estimating Surface Normals in Noisy Point Cloud Data.- Special issue of Int. J. Computational Geometry and its Applications, Band 14 (4-5), S. 261-276, Online: http://graphics.stanford.edu/projects/lgl/papers/mn-esnnpcd-03/mn-esnnpcd-03.pdf [Stand: 08.09.2012].

PEARSON, K.,1901: On lines and planes of closest fit to systems of points in space: Philosophical Magazine, S. 559-572, Online: http://stat.smmu.edu.cn/history/pearson1901.pdf [Stand 21.08.2012].

Rekonstruktion von Fenstern aus Schrägsicht-ALS Punktwolken zur Anreicherung von Gebäudemodellen

SEBASTIAN TUTTAS[1] & UWE STILLA[1]

Zusammenfassung: In der vorgestellten Arbeit werden Schrägsicht-ALS (Airborne Laserscanning) Punktwolken dazu verwendet, ein Gebäudemodell der TU München, in dem nur die Dachstrukturen modelliert sind, mit rechteckigen Fenster anzureichern. Die Registrierung des Modells mit der Punktwolke erfolgt über korrespondierende Dachebenen. Initiale Fensterpositionen werden durch die Detektion von Punkten, die hinter der Hauptfassade liegen, bestimmt. Die innenliegenden Punkte werden anschließend gelöscht und ein Rechteck in den leeren Bereich um die initialen Fensterpositionen eingepasst. Die Unsicherheit der Rechteckkanten wird modelliert um daraus spalten- und zeilenweise die Fensterumrisse zu schätzen. Für 11 unterschiedliche Fassaden, mit einer Gesamtzahl von 321 Fenstern, wurde eine Detektionsrate von 65% und eine Falschalarmrate von 14% erreicht. Die Unsicherheit der Fenstergrößen bei der Rekonstruktion liegt im Dezimeterbereich.

1 Einleitung

Für viele Städte existieren großflächig Stadtmodelle als Geometriemodell mit Dachstrukturen und ebenen Fassadenflächen. Um eine realistischere Darstellung zu erreichen, ist auch eine Modellierung der Fassadenelemente notwendig. Als wesentliches Element spielen dabei Fenster die wichtigste Rolle. Schrägsicht-ALS (Airborne Laserscanning) Punktwolken erlauben im Gegensatz zu Punktwolken, die in klassischer Nadirsicht aufgenommen wurden, die Interpretation von Fassaden. Gleichzeitig kann bei mehrfachen Überflügen die Szene nahezu komplett in 3D erfasst werden. Deshalb werden in der hier vorgestellten Arbeit diese Punktwolken verwendet um ein Gebäudemodell der TU München (TUM) mit Fenstern zu versehen.

Ansätze zur Fensterdetektion und -rekonstruktion aus terrestrischen Laserscanning-Daten basieren in den meisten Fällen darauf, dass Fenster als Löcher in den Punktwolken erscheinen. In PU & VOSSELMAN (2009) und BOULAASSAL et al. (2011) werden die Fensterkantenpunkte anhand von langen Dreieckskanten in einer Vermaschung der Punktwolke detektiert und Linien an diese Punkte angepasst. In WANG et al. (2012) werden die Kantenpunkte anhand von lokalen Nachbarschaften der Punkte bestimmt. Anschließend werden mit einem plane-sweeping-Verfahren ein vertikales und ein horizontales Profil der Kantenpunkte erstellt, in denen die Maxima die Fensterkanten darstellen. Dabei wird implizit von einer gitterartigen Anordnung der Fenster ausgegangen. Die Annahme, dass die Fenster einer Fassade eine oder mehrere zweidimensionale periodische Strukturen formen wird von MESOLONGITIS & STAMOS (2012) genutzt. Sie versuchen eine Gitterstruktur in ein Binärbild einzupassen, das aus Löchern in den Punktwolken generiert wurde.

1) Photogrammetrie & Fernerkundung, Technische Universität München, Arcisstraße 21, 80333 München, http://www.pf.bv.tum.de

Der Ansatz von SCHMITTWILKEN & PLÜMER (2010) unterscheidet sich von den vorhergenannten Arbeiten, da dieser auf Lage- und Formparameter für Fassadenelementen aus Trainingsdaten basiert, die zur Klassifikation und Rekonstruktion verwendet werden. Auch der bildbasierte Ansatz von REZNIK & MAYER (2008) beruht auf Trainingsdaten für die Detektion von Fenstern. Es wird hier ebenfalls die Annahme genutzt, dass Fenster i.d.R. periodische Strukturen formen. Die Arbeit von LEE & NEVATIA (2004) zielt auf die Integration von Fenstern in bestehende 3D-Gebäudemodelle ab, wofür in diesem Fall Texturen verwendet werden. Aus diesen werden Fensterkanten aus vertikalen und horizontalen Histogrammen von Kantenpixeln abgeleitet. Bei der Fassadenrekonstruktion werden auch formale Grammatiken eingesetzt, wobei im Ansatz von BECKER (2009) abgeleitete Fensterstrukturen synthetisch in nicht durch Daten abgedeckte Gebäudebereiche eines Gebäudemodells eingepasst werden. Bei RIPPERDA & BRENNER (2009) werden die Grammatiken in einem rjMCMC (reversible jump Markov Chain Monte Carlo) Ansatz verwendet um Fassaden zu rekonstruieren.
Die Evaluierung von Ergebnissen bei der Fassadenmodellierung aus terrestrischen Laserscanning-Daten, wird von LANDES et al. (2012) behandelt.

2 Methodik

Die in Kapitel 1 genannten Ansätze für Punktwolken sind auf hohe Punktdichten, z.B. für das Anpassen der Fensterkanten, angewiesen, wohingegen die hier vorgestellte Methode mit einer Punktdichte im Bereich von 10 Punkte/m^2 auskommen muss, da es auf Schrägsicht-ALS Punktwolken angewendet wird. Ein wesentlicher Unterschied zu den anderen Ansätzen besteht darin, dass die Fensterdetektion (Abschnitt 2.2) auf die im Gebäudeinneren liegenden Punkten beruht, die ein einfaches und universelles Merkmal für Fenster darstellen. Die Fensterrekonstruktion wird in Abschnitt 2.3 beschrieben. Auch dieser Ansatz ist auf eine geringe Punktdichte hin angepasst. Vor diesen beiden Punkten steht die Koregistrierung von Gebäudemodell und Punktwolke, welche in Abschnitt 2.1 dargestellt wird. Dabei wird auf die Schwierigkeiten eingegangen, die aufgrund der unterschiedlichen Erfassungsarten entstehen. Die Punktwolkendaten werden zu einem Modell registriert, dass manuell aus einem Stereoluftbildpaar generiert wurde.

2.1 Registrierung von Gebäudemodell und Punktwolke

Für die Korigistrierung von Modell und Punkwolke wird ein ebenenbasiertes Verfahren nach HEBEL & STILLA (2009) verwendet. Dabei werden Transformationsparameter (3 Translationen, 3 Rotationen) für eine Best-Fit-Transformation berechnet, die auf zwei Bedingungen basiert (1: Der Vektor vom Mittelpunkt der Punktwolkenebene zum Mittelpunkt der Modelleben soll senkrecht zum Normalenvektor der Modellebene stehen; 2: Die Normalenvektoren beider Ebenen sollen parallel sein). Um korrespondierende Ebenen zu finden, müssen zunächst zu den Modellebenen Punktwolken-Punkte zugeordnet werden. Dabei wird davon ausgegangen, dass die räumliche Lage der beiden Datensätze auf wenige Meter genau übereinstimmt. Der Unterschied in der Vororientierung beruht auf folgenden Fehlerquellen:

- Georeferenzierung Laserscanning (UTM) und Luftbilddaten (GK)
- Erfassung der Gebäudeumrisse bei der Erstellung der Gebäudemodelle
- Transformationsparameter (GK zu UTM)

Aufgrund der Charakteristik der Erfassungsweise der beiden Datensätze ist eine eindeutige Transformation von Punktwolke zu Modell nicht möglich. Da die Erfassung des Gebäudemodells aufgrund von Dachformen erfolgte, ist hier die Gebäudegröße durch die Größe des Dachbereichs festgelegt. In der Schrägsicht-Laserpunktwolke werden jedoch die Fassadenflächen an der wahren Position abgebildet, die je nach Gebäude mehr oder weniger weit von der Dachkante zurückgesetzt ist. Außerdem sind aufgrund der manuellen Erstellung des Gebäudemodells auch Abweichungen der Dachflächenneigung von mehreren Grad möglich. Aus diesen Gründen wurde die in Abb. 1 dargestellte Vorgehensweise entwickelt, um korrespondierende Punkte bzw. Ebenen zu detektieren und daraus die Transformation zu berechnen.

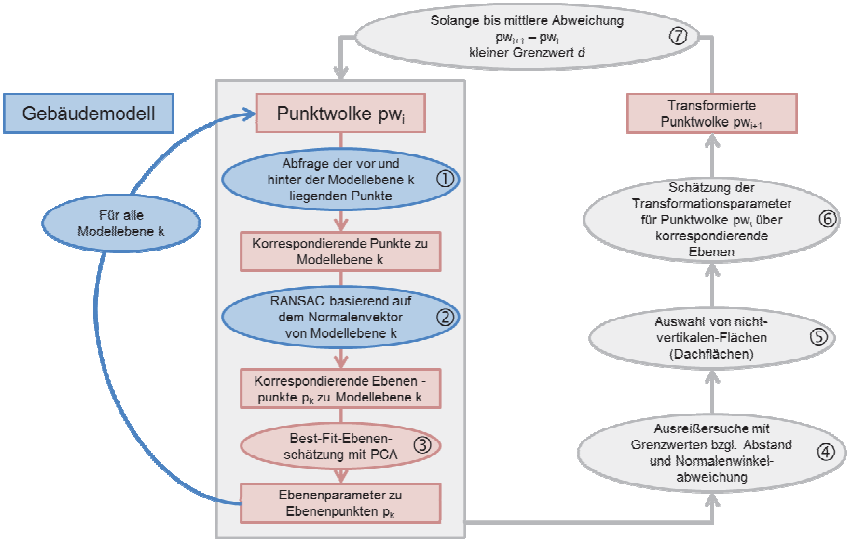


Abb. 1: Ablaufdiagramm zur Registrierung von Punktwolke und Gebäudemodell

Zunächst werden für jede Modellebene die Punkte ausgewählt, die in einem bestimmten Abstand a vor und hinter der Fassade liegen (1). Anschließend wird in diesen Punkten mit RANSAC die Ebene gesucht, die der Modellebene entspricht (2). Um hierbei in jedem Fall die zur Modellebene gehörende Ebene zu erhalten werden nur bestimmte Punkte als Samples (3 Punkte als Minimalkonfiguration einer Ebene) akzeptiert. Ist der Raumwinkel zwischen der Normalenrichtung aus der Ebene des Samples und der Normalenrichtung der Modellebene größer als ein Grenzwert γ (hier 15°) wird das Sample verworfen. Aus den resultierenden

Punkten (Inliers) werden mit Hilfe einer PCA (Hauptkomponentenanalyse) die Ebenenparameter geschätzt (3). Um bestmögliche Transformationsparameter zu erhalten, werden Ebenenpaare mit zu großen Abweichungen im Abstand der Schwerpunkte oder der Normalenrichtungen aussortiert (4). Um das oben beschriebene Problem bzgl. der Lage der Fassadenflächen zu umgehen werden nur Dachflächen berücksichtigt, d.h. vertikale Flächen werden ausgeschlossen (5). Die Schätzung der Transformationsparameter erfolgt dann über alle verbleibenden Ebenenpaare (6). Da die Punktwolke zu Beginn noch etwas zum Modell verschoben ist, kann nicht garantiert werden, dass alle zu einer Ebene gehörenden Punkte ausgewählt werden, wenn die Punkte vor und hinter der Modellebene verwendet werden. Aus diesem Grund wird der in Abb. 1 gezeigte Ablauf iterativ durchgeführt, solange bis die mittlere Abweichung zwischen den Koordinaten der Punktwolken vor und nach der Transformation einen Grenzwert d unterschreitet (7).

2.2 Fensterdetektion

Die Fensterdetektion beruht auf der Identifikation von im Gebäude liegenden Punkten (TUTTAS & STILLA, 2012). In dieser Arbeit wird die Detektion jedoch bezüglich einer vorgelagerten Ebenensegmentierung durchgeführt. Dabei ergab sich das Problem, dass in einem Ebenensegment Fassadenteile mit kleinem Versatz zusammengefügt sein können und somit leicht zurückliegende Teile komplett als innen liegende Punkte detektiert werden können. Da hier jedoch auf ein Gebäudemodell zurückgegriffen werden kann, sind alle Fassadenteile eindeutig getrennt. Deshalb beruht hier die Detektion der innen liegende Punkte nicht auf einem Schwellwert auf der Verteilung der Punkte, sondern es werden alle hinter der geschätzten Hauptebene liegenden Punkte verwendet. Die Hauptebene wird über eine Anpassung einer Ebene mit RANSAC bestimmt. Hieraus lässt sich auch erkennen, dass der Ansatz auf Gebäude bezogen ist, bei denen die Wandfläche gegenüber der Fensterfläche überwiegt. Im Falle einer Glasfassade ist es beispielsweise nicht möglich eine Hauptebene zu schätzen.
Eine weitere Neuerung bezüglich der oben genannten Veröffentlichung besteht darin, dass nun die innen liegenden Punkten nicht mit dem genäherten Einfallswinkel von 45° auf die Hauptebene projiziert werden, sondern der Strahl von Reflexionspunkt zu Sensorposition mit der Hauptebene geschnitten wird, um die Fensterposition auf dieser zu bestimmen.

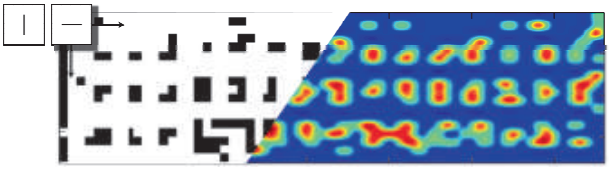


Abb. 2: Binärbild und Bild der Korrelationssumme aus vertikalem Linienmerkmal und Binärbild

Aus den detektierten Punkten wird ein Binärbild mit 1 m Auflösung erstellt, bei dem jedes Pixel den Eintrag 1 erhält wenn ein innen liegender Punkt vorliegt. Das Bild wird auf 10 cm Auflösung gesampelt und anschließend die Korrelationssumme, jeweils mit einem horizontalen und vertikalen Merkmal, gebildet (siehe Abb. 2). Dies dient der Glättung der Fensterinformation,

da je Fenster nur sehr wenige Punkte (ca. 1-10) detektiert werden. Die Länge der Linienmerkmale beträgt 2 m, der umgebende Bereich ist 4x4 m groß. Die Werte aus dem Bild der Korrelationssumme werden zeilen- und spaltenweise summiert und die Kombination aller Maxima als initiale Fensterpositionen verwendet.

2.3 Fensterrekonstruktion

Für die Fensterrekonstruktion wird der in TUTTAS & STILLA (2012) beschriebene Algorithmus verwendet. Die im vorherigen Abschnitt detektierten initialen Fensterpositionen werden als Startpunkte für die Schätzung eines Rechteckes verwendet. Aufgrund der Punktdichte von nur ca. 10 Punkte/m^2 ergeben sich bei diesem Vorgehen stark variierende Fenstergrößen, da die Fensterkanten durch die wenigen Punkte nicht eindeutig definiert sind. Deshalb wird die Unsicherheit in der Position jeder Fensterkante mit Hilfe einer Rayleigh-Verteilung modelliert, woraus dann eine endgültige Position geschätzt wird. Dazu werden die Funktionen zeilen- und spaltenweise aufsummiert. Das nächstliegende Maximum der resultierenden Funktion zur ursprünglichen Kante definiert dann die finalen Fensterumrisse. Hierbei soll genutzt werden, dass mehrere gleiche Fenster die Hypothese einer bestimmten Kante unterstützen, andersartigen Fenstern jedoch eine eigene Kante zugewiesen werden kann.
Da die Rekonstruktion auf den geschätzten Hauptfassaden aus der Punktwolke beruht, liegen die Fenster auch auf dieser Fassadenebene. Im letzten Schritt werden deshalb die Fenster auf die Ebene des Gebäudemodells projiziert. Dies geschieht entlang der Normalenrichtung der Ebene aus der Punktwolke.

3 Ergebnisse und Evaluierung

3.1 Koregistrierung

Die Eingangspunktwolke wurde mit dem Verfahren von HEBEL & STILLA (2009) erzeugt. Sie wurde aus Punktwolken von vier Überflügen, unter Verwendung des unter 2.1 genannten ebenenbasierten Verfahrens, zusammengesetzt.
Das Ergebnis der Koregistrierung ist in Abb. 3 für alle Punkte, die zu einer Modellebene zugeordnet worden sind, zu sehen. Die Koregistrierung erfolgte, leicht variierend je nach Iterationsschritt, jeweils über ca. 150 Dachflächen. Der Grenzwert d für den Abbruch der Iteration wurde auf 10 cm gesetzt, was zu 9 Durchläufen führte. Für den Abstand a für die Punktsuche vor und hinter der Modellebene wurde 5 m gewählt. Die Grenzwerte für die in Abb. 1 (Schritt 4) aufgeführte Ausreißersuche betragen 5 m für den Abstand und 25° für den Winkel zwischen den Normalen.
In Abb. 4 sind die Abstände des Schwerpunkts der Modellebene zur geschätzten Hauptfassade in der Punktwolke sowie die Winkel zwischen deren Normalenvektoren aufgezeigt. Die Abweichungen sind bei 90% der Fassaden unter 1 m bzw. 1°. Diese Werte spiegeln die Größenordnung wieder, wie gut Modell und Punktwolke übereinstimmen. Die sehr großen Werte bei den Schwerpunkten zeigen falsch erfasste Fassaden im Modell auf. Die Abweichungen machen es notwendig, die oben genannte Projektion der Fensterumrisse von Punktwolken- auf Modellebene durchzuführen.

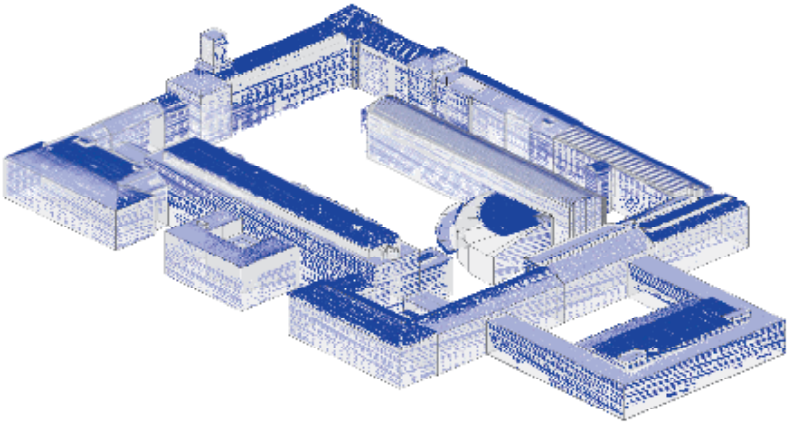
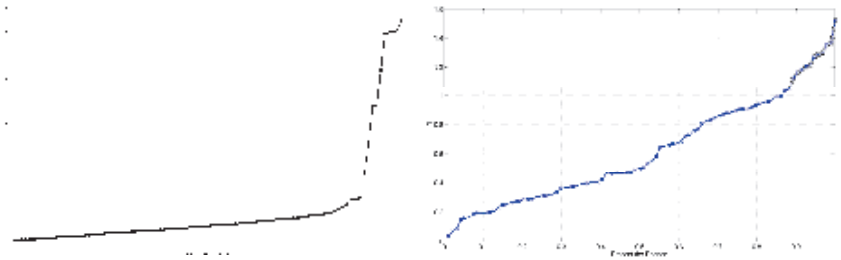


Abb. 3: Gebäudemodell und Punktwolke (nur den jeweiligen Modellebenen zugeordnete Punkte) nach der Koregistrierung

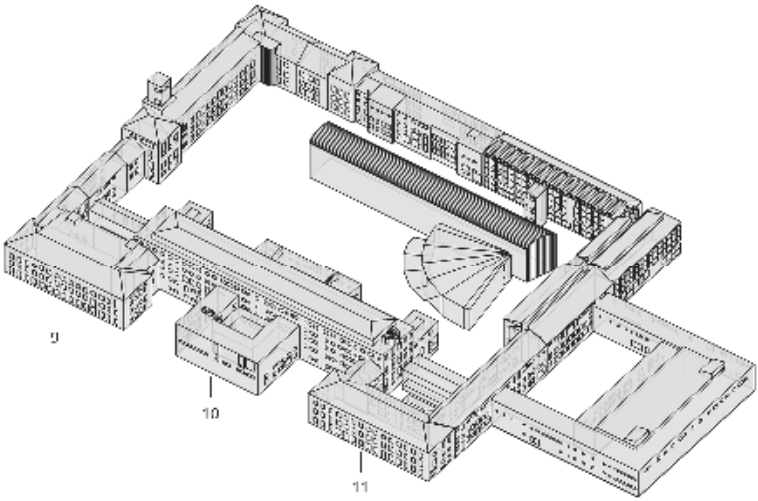


a) b)

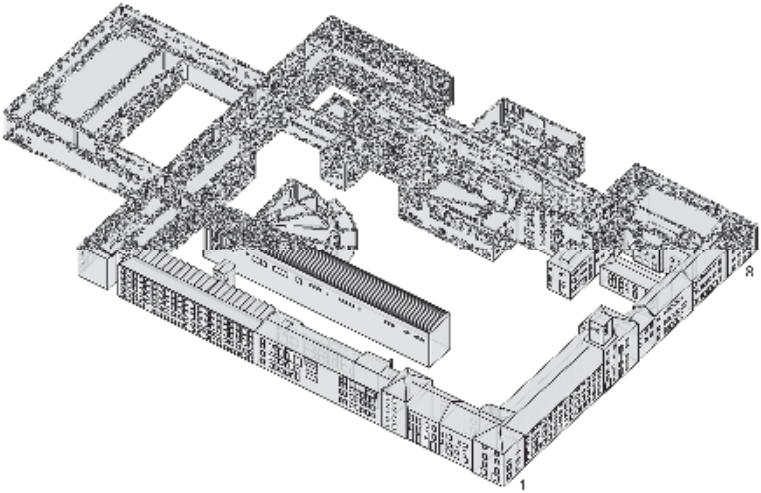
Abb. 4: Abstand des Schwerpunkts der Modellebene von der geschätzten Punktwolken-Ebene (a) sowie Winkel zwischen den Normalenvektoren beider Ebenen (b) von den 94 Ebenen für die Fenster berechnet wurden.

3.2 Fensterdetektion und -rekonstruktion

In Abb. 5 ist das Ergebnis der Fensterrekonstruktion für das gesamte Gebäudemodell zu sehen. Für die Bestimmung der Hauptfassadenebene wurde im RANSAC-Algorithmus ein Schwellwert von 10 cm für den Abstand eines Punkts von der Ebene verwendet.



a)



b)

Abb. 5: Darstellung des Gebäudemodells mit Fenstern, Blick von Nordosten (a) und Südwesten (b)

Eine quantitative Evaluierung wurde für die Außenfassaden im Süden und Osten der TUM durchgeführt. In Abb. 6 und Abb. 7 sind die rekonstruierten Fenster zusammen mit Texturen für die genannten Fassaden gezeigt. In Tab. 1 sind für die Fassadenteile die Detektionsraten angegeben. Für alle Fassaden konnte eine Detektionsrate von 65% bei 14% Fehldetektionen erreicht werden. Da die Fassade 10 deutlich von der Charakteristik der anderen Fassaden abweicht und die große Zahl an dicht beieinanderliegenden Fenstern die Statistik verzerren, wird auch die Detektionsrate von 78% bei 9% Fehldetektionen angegeben, die sich ohne Fassade 10 ergibt. In allen Fällen werden jedoch die sehr kleinen Fenster in der obersten Reihe bei Fassaden 8, 9 und 11 miteinbezogen, die die Detektionsrate von 75% erklären, wenn auch zunächst alle Fenster detektiert erscheinen.



Abb. 6: Texturen und rekonstruierte Fenster für Fassadenteile 1 bis 8 (von links nach rechts)



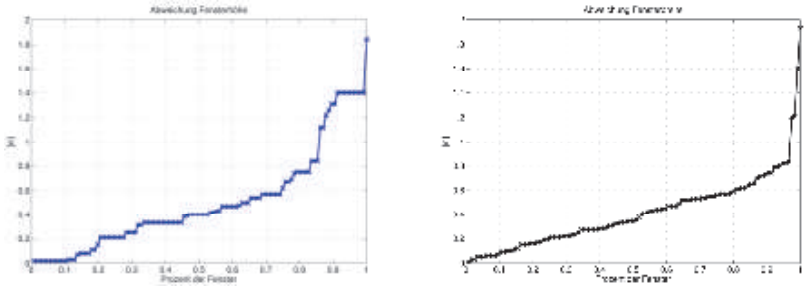
Abb. 7: Texturen und rekonstruierte Fenster für Fassadenteile 9 bis 11 (von links nach rechts)

Tab. 1: Detektionsraten (DR) und Falschalarmrate (FR) für die in Abb. 6 und Abb. 7 gezeigten Fassaden (GT= Ground Truth, D = Detektionen, TP = True Positives, FP = False Positives, FN = False Negatives)

												Σ	**ohne 10**
Nummer	1	2	3	4	5	6	7	8	9	10	11		
GT	14	42	9	9	9	24	21	16	60	57	60	**321**	**264**
D	6	56	10	6	6	22	18	12	45	15	45	**241**	**226**
TP	2	42	9	6	6	21	18	12	45	2	45	**208**	**206**
FP	4	14	1	0	0	1	0	0	0	13	0	**33**	**20**
FN	12	0	0	3	3	3	3	4	15	55	15	**113**	**58**
DR	0.14	1.00	1.00	0.67	0.67	0.88	0.86	0.75	0.75	0.04	0.75	**0.65**	**0.78**
FR	0.67	0.25	0.10	0.00	0.00	0.05	0.00	0.00	0.00	0.87	0.00	**0.14**	**0.09**

Für die Fassaden aus Abb. 6 wurden die Referenzgrößen der Fenster aus photogrammetrischer Messung ermittelt. Dazu wurde ein Bündelblock aus Fassadenbildern gerechnet und anschließend Orthofotos der jeweiligen Fassaden berechnet. Die Messung der Fensterhöhe und

Fensterbreite erfolgte dann manuell in den Orthofotos. Zur Analyse wurden alle 116 korrekt detektierten Fenster der Fassaden 1 bis 8 herangezogen. Es werden nur die Abweichungen in der Fenstergröße und nicht die Positionsgenauigkeit betrachtet, da keine Koregistrierung von Textur und Punktwolke erfolgt ist. Der Betrag der Abweichung über alle Fenster ist in Abb. 8 für Höhe (a) und Breite (b) dargestellt. Der Mittelwert der Abweichungen beträgt 0.50 m für die Höhe und 0.41 m für die Breite.



a) b)

Abb. 8: Betrag der Abweichungen der Fensterhöhe (a) und der Fensterbreite für die in Abb. 6 gezeigten Fenster

4 Fazit und Ausblick

Der Artikel beschreibt einen Ansatz zur Anreicherung eines mit einer Schrägsicht-ALS Punktwolke koregistrierten Gebäudemodells mit Fenstern. Für eine Auswahl der Fassaden ist eine ausführliche Qualitätsanalyse durchgeführt worden. Wird eine Fassade mit stark irregulärer Anordnung nicht berücksichtigt erhält man eine Detektionsrate von 78%. Die Größenabweichung der Fenster zur Referenz variiert stark, von wenigen [cm] bis zu einem [m], was von der Anzahl der Fenster eines gleichen Typs sowie der lokalen Punktdichte beeinflusst wird.

Als zukünftige Aufgabe wird die Erkennung von gleichgearteten Fenstern angesehen, um diesen dann eine einheitliche Fenstergröße zuordnen zu können. Dies kann regelbasiert oder mit Hilfe von Texturinformationen erfolgen.

5 Danksagung

Wir bedanken uns bei Marcus Hebel (Fraunhofer IOSB) für die Bereitstellung der koregistrierten Punktwolke.

Die Arbeit wurde teilweise durch das DFG-Projekt STI 545/6-1 „Entwicklung eines automatisierten Verfahrens zur Baufortschrittskontrolle auf Basis der Integration von Punktwolkeninterpretation und 4D-Bauwerksmodellierung“ gefördert.

6 Literaturverzeichnis

BECKER, S., 2009: Generation and application of rules for quality dependent façade reconstruction. ISPRS Journal of Photogrammetry and Remote Sensing, **64** (6), S. 640-653.

BOULAASSAL, H.; LANDES, T. & GRUSSENMEYER, P., 2011: 3D modelling of façade features on large sites aquired by vehicle based laser scanning. Archives of Photogrammetry, Cartography and Remote Sensing edited by Polish Society for Photogrammetry and Remote Sensing. Vol. 22: S. 215-226.

LANDES, T.; BOULAASSAL, H. & GRUSSENMEYER, P., 2012: Quality Assessment Of Geometric Façade Models Reconstructed From TLS Data. The Photogrammetric Record **27** (138): S. 137-154.

LEE, S.C. & NEVATIA, R., 2004: Extraction and integration of window in a 3D building model from ground view images. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2004), Washington, DC, USA, Vol. 2, S. 113-120.

HEBEL, M. & STILLA U., 2009: Automatische Koregistrierung von ALS-Daten aus mehreren Schrägansichten städtischer Quartiere. PFG Photogrammetrie, Fernerkundung, Geoinformation, 2009 (3), S. 261-275.

MESOLONGITIS, A. & STAMOS, I., 2012: Detection of windows in point clouds of urban scenes. Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on, S. 17-24.

PU, S. & VOSSELMAN, G., 2009: Knowledge based reconstruction of building models from terrestrial laser scanning data. ISPRS Journal of Photogrammetry and Remote Sensing, **64** (6), S. 575-584.

REZNIK S. & MAYER H., 2008: Implicit shape models, self-diagnosis, and model selection for 3D façade interpretation. PFG Photogrammetrie, Fernerkundung, Geoinformation, 2008 (3), S. 187-196.

RIPPERDA, N. & BRENNER, C., 2009: Application of a Formal Grammar to Facade Reconstruction in Semiautomatic and Automatic Environments. Proceedings of 12th AGILE Conference on GIScience, Hannover, Germany, 2009.

SCHMITTWILKEN, J. & PLÜMER, L., 2010: Model-based reconstruction and classification of façade parts in 3D point clouds. In: Int. Archives Photogramm. Remote Sens. Spatial Inf. Sci., Sainte-Mandé, France, Vol. XXXVIII, Part 3A, S. 269-274.

TUTTAS, S. & STILLA, U., 2012: Reconstruction of rectangular windows in multi-looking oblique view ALS data. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., I-3, S. 317-322.

WANG, R.; FERRIE, F. & MACFARLANE J., 2012:. A Method for Detecting Windows from Mobile LiDAR Data. Photogrammetric Engineering and Remote Sensing, **78** (11), S. 1129-1140.

True- Oblique- Orthomosaik aus Schrägluftbildern zur Extraktion von 3D-Geoinformationen

ALEXANDER WIEDEN & MAGDALENA LINKIEWICZ

Im Zuge dessen wurde 2009 das modulare Luftbild-Kamerasystem MACS (Modular Airborne Camera System) entwickelt, welches in einer Ausführung Obliqueaufnahmen bereitstellt. Weiter wird ein Auswertekonzept erarbeitet, welches auf klassische Analysemethoden aus dem Bereich der digitalen Bildverarbeitung und der Geoinformatik basiert, um neue Datenprodukte aus Obliquebefliegungen semiautomatisch zu generieren. Das Potential dieser Methode für die Extraktion von Gebäudegeometrie und die Erfassung von semantischen Geoinformationen wird in einem Fallbeispiel erläutert.

1 Einleitung

Die Anforderungen an photogrammetrische Systeme sind in den letzten Jahren gewachsen. Das Bedürfnis, nicht nur einen 2D/ -2,5 -Blick auf die Welt zu haben, nimmt zu. Die Abteilung Sensorkonzepte und Anwendungen im Deutschen Zentrum für Luft- und Raumfahrt betrachtet heute photogrammetrische Sensorkonzepte aus der Perspektive der Anwender und Nutzer. Dabei zeigten sich Schwerpunkte im Bereich der Flexibilität und der Effektivität zwischen Anforderungen und Kosten der Systeme. Im Zuge dieser Erkenntnis entwickelte das Deutsche Zentrum Für Luft- und Raumfahrt das Modular-Airborn-Kamerasystem (MACS). Dieses Kamerasystem zeichnet sich durch ein modulares und flexibles Konzept aus, welches an spezifische Aufgabenstellungen und Anforderungen angepasst und für sie erweitert werden kann.Die Flexibilität dieses Systems sollte kosteneffizient sein und somit für Nutzer erweiterte Möglichkeiten in speziellen Anwendungsfeldern bieten.

Die Aufnahme von Schrägluftbildern unterscheidet sich wesentlich von klassischen Aufnahmen von Luftbildern. Der photogrammetrische Normalfall bezieht sich auf die automatisierte photogrammetrische Auswertung von Senkrechtbildern. Für diesen Fall existieren hoch automatisierte Algorithmen zur Ableitung von photogrammetrischen Produkten wie dem digitalen Oberflächenmodell und dem Orthophoto. Ebenso existieren für diese Datenprodukte robuste und effektive Algorithmen im Bereich der Klassifizierung und Segmentierung

Photogrammetrische Obliqueaufnahmen bilden im Gegensatz zu bisherigen photogrammetrischen Aufnahmen die Welt aus einer Oblique-Perspektive ab. Der Vorteil dieser Aufnahmen liegt in der Erweiterung der Perspektive. Diese Aufnahmen erfassen in urbanen Gebieten Fassaden von Gebäuden und in geomorphologisch erhabenen Bereichen Hänge und Steilwände. Gerade in den letzten Jahren wurden hierfür einige photogrammetrische Kamerasysteme entwickelt. Von rotierenden 3-Kopf-Systemen wie AOS (WIEDEMANN, 2011), über das 4-Kopfsystem DigiCam Oblique bis zu 5-Kopfsystemen wie DiGiCam Penta (PETRIE, 2008) und dem ersten System von Pictometry (Pictometry, 2007).

1) Alexander Wieden, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Einrichtung Optische Informationssysteme, Sensorkonzepte und Anwendungen, Rutherfordstraße 2, 12489 Berlin, Email: alexander.wieden@dlr.de
2) Magdalena Linkiewicz, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Einrichtung Optische Informationssysteme Sensorkonzepte und Anwendungen, Rutherfordstraße 2, 12489 Berlin, Email: magdalena.linkiewicz@dlr.de

Vorwiegend wurden diese Systeme für die Bereiche der Fassadentexturierung, Denkmaldokumentation oder auch Monitoring eingesetzt.

Photogrammetrische Methoden, diese Aufnahmen vorwiegend automatisch auszuwerten, stellen bis heute ein ungelöstes Problem dar. Zum einen liegt dies an der besonderen Abbildung und zum anderen an der Komplexität der Datenprodukte. Obliquedaten stellen den Anspruch dreidimensionale Daten zu erzeugen. Dies bedeutet, dass die herkömmlichen Datenprodukte aus Luftbildern, wie das digitale Oberflächenmodell und das digitale Orthomosaik aufgrund Ihrer dimensionalen Beschränkung nicht mehr ausreichen.

Eine Lösungsmöglichkeit bietet die dreidimensionale Punktewolke. Diese kann aufgrund Ihrer Eigenschaften für eine X, Y- Koordinate mehrere Z Werte Abbilden. Ein Nachteil dieser dreidimensionalen Punktewolken ist die Verfügbarkeit von methodischen Ansätzen. Die Herausforderung ist der Weg von generischen 3D-Daten zu semantischen und topologischen 3D-Daten. Die Komplexität, generische 3D-Informationen in topologische und hierarchische Beziehungen zu setzen, scheint unlösbar.

Die Generierung von 3D-Punktwolken ist die komfortabelste Lösung, birgt jedoch keine robusten Lösungsmöglichkeiten für die Ableitung von geometrischen und semantischen Objekten. Ein Grund hierfür ist die Komplexität von 3D-Punktwolken. Da die geometrische Information neben der, wenn vorhanden, radiometrischen Information der Nachbarpunkte einziger topologischer Indikator ist müssen im dreidimensionalen Raum bis zu 26 Parameter geprüft werden. Im Gegensatz dazu werden in zweidimensionale Beziehungen maximal 8 Nachbarinformationen und in eindimensionalen Beziehungen maximal zwei Nachbarinformationen benötigt (siehe Abbildung 1)

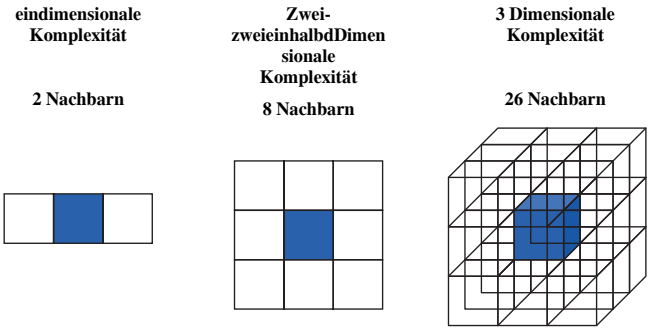


Abbildung 1: Dimensionale Steigerung der Komplexität

Im Bereich der Photogrammetrie und Fernerkundung wird der Begriff „3D- Rekonstruktion“ im Einklang mit digitalen Oberflächenmodellen oder digitalen Geländemodellen genannt. Der Ausdruck 3D ist erst mit der Funktion $Z = f(X,Y,Z)$ erfüllt. Diese Funktion lässt für einen X-

und Y-Wert mehrere Informationen in der Z-Ebene zu (LUHMANN, 2006, Nahbereichsphotogrammetrie, S. 74).

Bei Betrachtung der heutigen Datenprodukte im Bereich der Fernerkundung und Geowissenschaften wird klar, dass die methodische Bearbeitung von zweidimensionalen Rasterdaten, wie Bildmosaike und Oberflächenmodelle. für viele Fragestellungen verschiedene Lösungsmöglichkeiten bereithält. Klassische Verfahren wie Segmentierung, Vektorisierung und Klassifizierung werden erfolgreich im Bereich der Geowissenschaften eingesetzt.

2 Methodischer Ansatz

Um Methoden der Digitalen Bildverarbeitung im Zusammenhang mit Obliqueaufnahmen nutzen zu können ist es notwendig Datenprodukte zu erzeugen welche sich methodisch auswerten lassen und Ihren Anspruch der Dreidimensionalität nicht verliert.

Einen möglichen Ansatz stellt die Minimierung der Dimensionalität durch Mehrfachabbildungen dar. Hierzu wird eine dreidimensionale Fragestellung durch mehrere zweidimensionale Abbildungen gelöst (siehe Abbildung 2)..

Im Bereich der Obliqueaufnahmen wird der dreidimensionale Objektraum durch mehrere zweidimensionale Projektionsebenen Abgebildet (siehe Abbildung 3). Das bedeutet, dass von einem Objekt aus unterschiedlichen Ansichten metrische zweidimensionale Abbildungen berechnet werden, welche den klassischen Datenprodukten, dem digitalen Oberflächenmodell und dem digitalen Orthomosaik sehr ähnlich sind. Hierfür wird nicht für jedes Bild eine eigene Tiefenkarte erzeugt, sondern eine lokaler Raum definiert, welcher durch eine Anzahl von Aufnahmen, einem Cluster, optimal abgebildet wird.

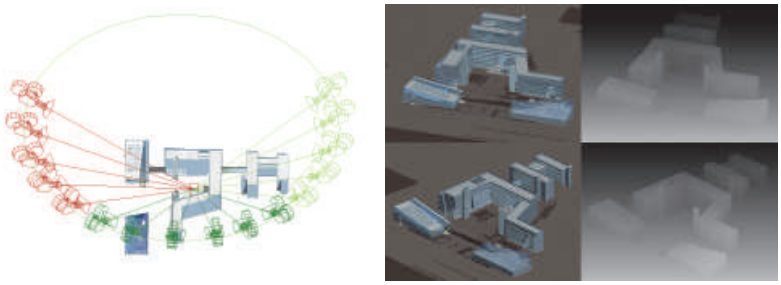


Abbildung 2: Zwei lokale Datenprodukte aus rot und grünen Clustern

Ein Vorteil dieser lokalen Räume ist die die Abbildung von dreidimensionalen Fragestellungen durch eine geringe Anzahl von Datensätzen. Die optimierte Definition der lokalen Räume hängt von den Eigenschaften der Matching-Algorithmen ab.

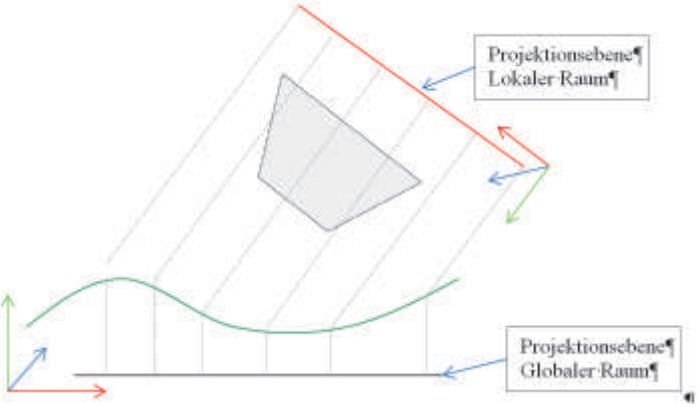


Abbildung 3: 2D- Orthoprojektion und Lokale Orthoprojektion

Um ein digitales Oberflächenmodell aus einem Obliquecluster abzuleiten wird die Hauptausrichtung und der Schwerpunkt des Clusters bestimmt. Die Hauptausrichtung des Schwerpunkts eines Clusters ähnlicher Aufnahmegeometrien bestimmt die Ausrichtung des Lokalen Raums. Um die Auswirkungen von Matching-Algorithmen auf die inhärenten Bildinhalte zu evaluieren, wurden paarweise die Bildinformationen zwischen zwei Bildern nach Merkmalen untersucht und verglichen. Als ähnliche Aufnahmegeometrien werden Bilder mit einer maximalen Rotationsänderung von 20° bezeichnet. Ist das Mittel aller Beobachtungen bestimmt, werden die Rotationswinkel in die Rotationsmatrix eingesetzt, um die Hauptrotationsmatrix (norm) zu definieren.

$$R_{\omega\varphi\kappa} = \begin{pmatrix} \cos\varphi\cdot\cos\kappa & -\cos\varphi\cdot\sin\kappa & \sin\varphi \\ \cos\varphi\cdot\sin\kappa+\sin\omega\cdot\sin\varphi\cdot\cos\kappa & \cos\omega\cdot\cos\kappa-\sin\omega\cdot\sin\varphi\cdot\sin\kappa & -\sin\omega\cdot\cos\varphi \\ \sin\varphi\cdot\sin\kappa-\cos\omega\cdot\sin\varphi\cdot\cos\kappa & \sin\omega\cdot\cos\kappa+\cos\omega\cdot\sin\varphi\cdot\sin\kappa & \cos\omega\cdot\cos\varphi \end{pmatrix}$$

Die Hauptrotationsmatrix ermöglicht eine Transformation homologer Koordinaten von einem globalen Raum in einen lokalen Raum.

$$A = R^{t}_{(norm)} \cdot R_{(org)}$$

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R^{t}_{(norm)} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Die Normierte Rotationmatrix wird mit den Näherungswerten der originalen Bildrotation aller Bilder eines Clusters multipliziert, um die Rotation im lokalen Raum zu berechnen. Die Koordinaten der Projektionszentren müssen durch die Bedingung der Kolliniaritätsgleichung ebenfalls mit der Normierten Rotationsmatrix in den lokalen Raum überführt werden. Diese Haupttransformation wird für die Normierung der Koordinaten durchgeführ,t um extreme Wertebereiche auszuschließen. Die Normierung der Projektionszentren wird wie die Rotationswinkel durch das Mittel aller Beobachtungen eines Clusters bestimmt.

3 Umsetzung

Am Deutschen Zentrum für Luft- und Raumfahrt wurde ein Modular-Kamerasystem (MACS) entwickelt. Dieses Kamerasystem zeichnet sich durch einen flexiblen Aufbau aus, wodurch den verschiedene Anforderungen und Fragestellungen durch die Anpassung von Sensoren und Kameramodulen begegnet werden kann. Im Frühjahr 2011 wurde das MACS durch zwei Sensorköpfe erweitert, welche eine Oblique-Aufnahme realisieren sollten. Das System MACS-Oblique verfügt über fünf synchron auslösende Kameraköpfe (LEHMANN, 2011).

Im System MACS-Oblique sind drei senkrecht schauende sowie zwei Oblique-Kameras verbaut. Die senkrecht schauenden Kameras bestehen aus einem „Geo-CAM“- und zwei hochauflösenden „DOM- CAM“-Modulen.

Das GeoCam-Modul ist mit einem 50mm-Objektiv und mit einem Infrarotfilter ausgestattet. Diese Konfiguration führt durch geeignete Aufnahmegeometrie zu einer stabilen Lage und Positionsbestimmung durch klassische Aerotriangulationsverfahren. Darüber hinaus bietet der Infrarotfilter die Möglichkeit Infrarotaufnahmen für Klassifikationsaufgaben bereitzustellen.

Das DOM-CAM-Modul besteht aus zwei Kameraköpfen mit 70mm-Objektiven. Diese Module bilden RGB-Informationen durch klassische Senkrechtaufnahmen ab. Um eine Maßstabsänderung durch Verkippung der Sensoren zu umgehen, wurden die beiden Module exzentrisch über dem Sensor gelagert. Dies hat zur Folge, dass die Bildaufnahmen, ähnlich UltraCAM und DMC, mit vershifteten Objektiven ausgestattet sind. Der Vorteil dieses Aufbaus liegt in einer höheren Abdeckung sowie der homogenen Maßstäbe in den Aufnahmen.

Für diesen Test wurde das Kamerasystem um zwei 70mm-Kameramodule, ObliqueCam V und ObliqueCam R, erweitert. Die Ausrichtung der schräg blickenden Kameramodule unterlag hierbei definierten Zwängen. Die ObliqueCam R verkippt um 30° quer zur Flugrichtung zur rechten Seite. Die ObliqueCam V verkippt um 40° in Flugrichtung. Die 40°-Verkippung erzeugt durch das 70mm Objektiv und eine größere Entfernung zum Objekt jedoch eine ähnliche Abdeckung wie die Nadiraufnahme (siehe Abbildung 4).

Eine vollständige Oblique-Aufnahme kann mit einer alternierenden Befliegung gewährleistet werden. Durch eine alternierende Befliegung mit 50% Überlappung zwischen den Streifen ergeben sich Aufnahmen aus allen vier Richtungen.

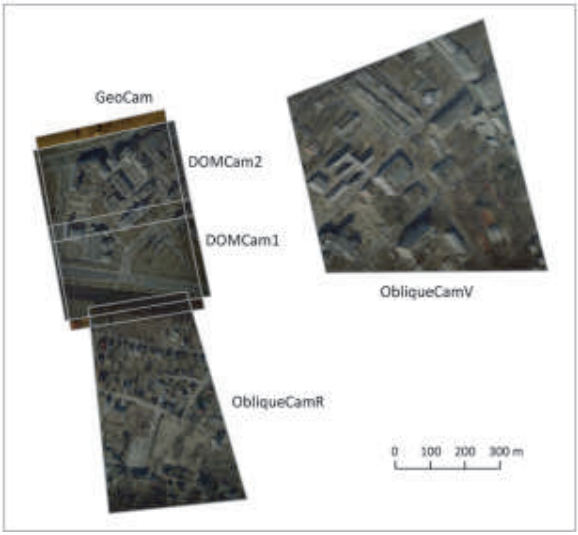


Abbildung 4: Aufnahmegeometrie MACS Oblique Kamerasystem

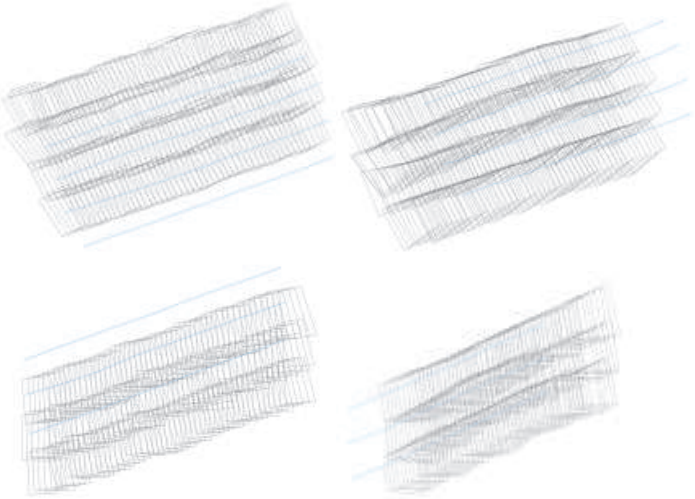


Abbildung 5: Darstellung der Footprints vom Clusterergebnis der ähnlichen Aufnahmegeometrien

Nach Analyse der Aufnahmekonstellationen entstanden unter Bedingung von ähnlichen Aufnahmegeometrien mehrere Aufnahmebündel.

Als Kriterien für die Cluster der Aufnahmen wurden die Winkelinformationen des Inertialsystems genutzt. Nach Betrachtung der Initialdaten konnten alle Flugstreifen in sich zu einem Cluster zusammengeführt werden. Darüber hinaus wurden auch die Aufnahmen in ein Cluster gebündelt, welche aus unterschiedlichen Flugstreifen aufgenommen, wurden jedoch die gleiche Aufnahmerichtung aufwiesen.

Die **Fehler! Verweisquelle konnte nicht gefunden werden.**5 zeigt ein Teilergebnis der zusammengefassten Aufnahmen. Oben-links wurden die Aufnahmen aus Streifen 1; 3; 5 der nach rechts schauenden Kamera (ObliqueCam R) zusammengefasst. Oben-recht die Kamera ObliqueCam R mit den Streifen 1; 3; 5, unten-links die Streifen 2; 4; 6; der ObliqueCam V und unten-rechts die Streifen 2; 4; 6 der ObliqueCAM V.

Insgesamt entstanden so 7 Cluster der Obliquedaten für die Befliegung sowie den klassischen Normalfall der GeoCAM.

Nachdem alle Parameter der äußeren Orientierung nun in den lokalen Raum transformiert wurden, können 2,5D-Informationen in Form von Tiefenkarten berechnet werden. Um zu gewährleisten, dass die Komplanaritätsbedingung erfüllt wird, bietet es sich an, die Epipolargeometrie zu prüfen. Die Epipolargeometrie beschreibt die den geometrischen Zusammenhang zwischen zwei Aufnahmen. Dabei bildet die Epipolarlinie eine Gerade von der optischen Mitte einer Kamera durch eine definierte Position im dreidimensionalen Raum aus der Sicht einer anderen Kamera ab. Dies bedeutet, wird ein Punkt in einem Bild Definiert so sollte dieser durch die Epipolarlinie in allen anderen Bildern geschnitten werden.

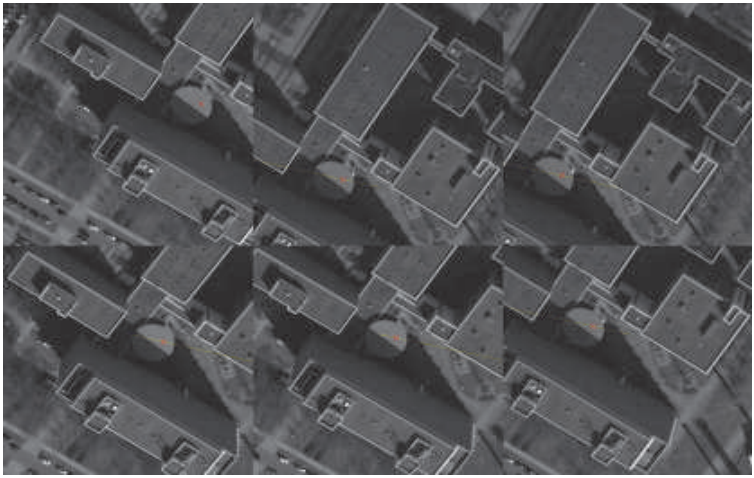


Abbildung 6: Kontrolle der Epipolargeometrie

In Abbildung 6 wird für einen Punkt die Epipolarlinie beispielhaft visualisiert. Dabei wird in einem Referenzbild (oben links) ein Punkt identifiziert. In allen Bildern, welche nach der Komplanaritätsbedingung diesen Punkt ebenfalls abbilden müssen, wird eine Epipolarstrahl (gelb) visualisiert. Dadurch lässt sich prüfen, ob die Transformation unter Berücksichtigung der Komplanaritätsbedingung durchgeführt wurde. Erst wenn alle Strahlen den gesuchten Punkt schneiden, darf mit der Extraktion der Tiefeninformation begonnen werden.

Nachdem für jeden definierten lokalen Raum die Transformation der Orientierungsparameter durchgeführt wurde, können Tiefenkaten extrahiert werden. Für die Testbefliegung der MACS über Adlershof wurden sieben lokale Räume und ein globaler Raum visualisiert. In der **Fehler! Verweisquelle konnte nicht gefunden werden.**7 sind Ausschnitte der Tiefenkarten und der True-Orthoprojektion abgebildet (unten rechts klassischer Fall).

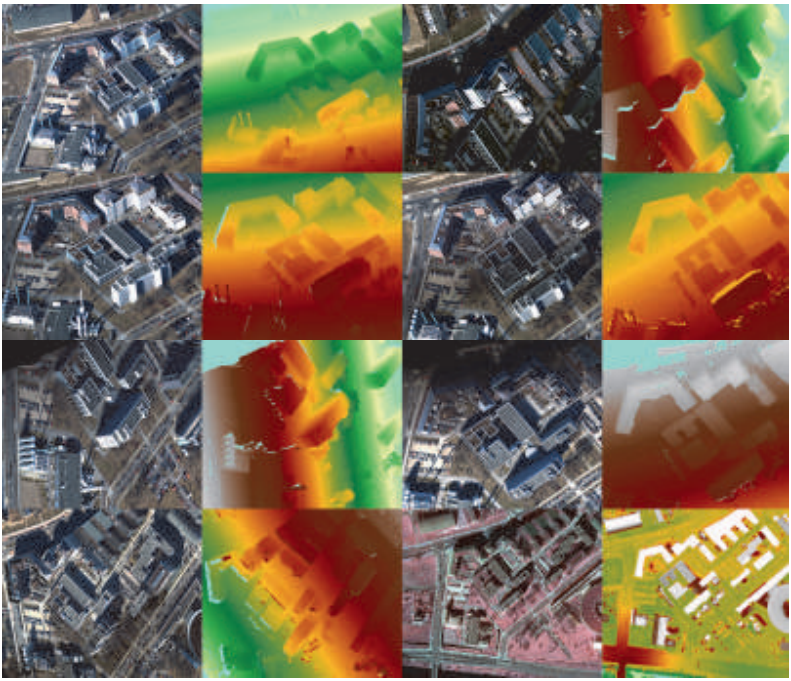


Abbildung 7: Lokale Tiefenkarten und True-Orthobildprojektionen

Diese Datenprodukte haben gleiche Eigenschafften wie klassische Orthobilder und Oberflächenmodelle aus Luftbildern oder Satellitenbilder. Die Oblique-Tiefenkarten sind Co-

registriert zu den Obliquetrueorthobildern und beide Datensätze sind im lokalen Raum georeferenziert. Aufgrund der definierten Transformationsvorschrift stehen alle „Lokalen“ Sichten mit dem Globalen Bezugssystem in Beziehung und daher mathematisch auch miteinander. Um diese gewonnen Informationen nicht nur visuell nutzen zu können, muss die Transformation ebenso von einem lokalen zum globalen Raum durchführbar sein. Hierzu müssen die entstandenen zweidimensionalen Informationen wie Punkte, Linien und Flächen, die Tiefeninformation des Oberflächenmodells übernehmen. Die somit entstehenden 2,5D-Koordinaten lassen sich mit der normierten Rotationsmatrix und dem normierten Translationsvektor in den globalen Raum überführen.

$$\text{Infomation } (X,Y,Z)_{glob} = (\text{Infomation } (X,Y,Z)_{lok} * R_{norm}) + T_{norm}$$

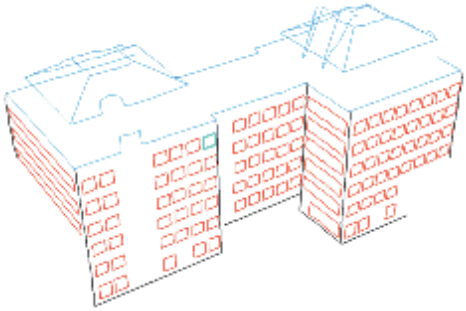


Abbildung 8: transformierte 3D Gebäudegeometrie aus lokalen Raum

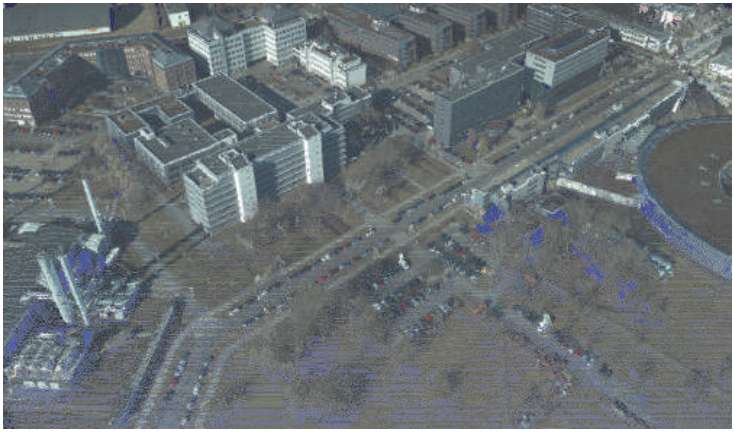


Abbildung 9: Fusionierte 3D Punktwolke aus 4 lokalen Räumen

Als weiteres Beispiel wurden die lokalen Tiefeninformationen in eine Ascii-Punktdatei überführt und mit den RGB-Informationen texturiert. Anschließend wurden die einzelnen lokalen Räume durch die Transformationsvorschrift in den globalen Raum transformiert. Die verschiedenen lokalen Räume erzeugen durch die Fusion im 3D- Raum eine zusammenhängende 3D-Punktewolke.

4 Fazit/Ausblick

Die automatische Gebäudeextraktion ist nach wie vor ein aktuelles Thema. Gerade die Bereiche der Computer-Vision und Geoinformatik arbeiten akribisch an diesem Thema und sind durch unterschiedlichste Ansätze bemüht, diese Fragestellung zu lösen.

Die Nutzung von Schrägluftbildern für die Extraktion von Geoinformationen birgt einige Hindernisse. Besonders die projektiven und perspektivischen Einflüsse. Um die Nutzung von Schrägluftbildern den Geoinformationswissenschaften zugänglich zu machen, war es das Ziel, Datenprodukte in Form von Digitalen Oberflächenmodellen und Orthoprojektionen der Bilddaten für Schrägansichten abzuleiten. Um diese Datenprodukte generieren zu können, wurde der mathematische Zusammenhang zwischen globalen und lokalen Räumen definiert. Unter konsequenter Berücksichtigung der Komplanaritätsgleichung konnten so lokale Datenprodukte für weitere Analysen erzeugt werden. Das Clustern von Schrägluftbildern mit ähnlichen Aufnahmerichtungen ermöglicht das ableiten von True-Oblique- Oberflächenmodellen und True-Oblique- Bildmosaiken. Die neuartigen Datenprodukte können aufgrund klassischer Methoden analysiert werden und somit erste Ergebnisse für die semantische Rekonstruktion von Gebäudedaten aufzeigen.

Darüber hinaus wurde die Rückführung extrahierter Geoinformation in den ursprünglichen Raum aufgezeigt, womit die Nutzung der geometrischen Lage sichergestellt wurde.

Generell konnten aufgrund des Konzepts neue Datenprodukte in Form von Tiefenkarten und Orthobilder ähnlich der klassischen Datenformate erzeugt werden. Inwieweit sich diese Daten eignen, um geometrische Aspekte von Gebäuden vollständig und robust, semantisch wie topologisch abzuleiten, müssen weitere Untersuchungen zeigen. Speziell die klassischen Methoden der Fernerkundung und die Methoden der digitalen Bildverarbeitung könnten einen wesentlichen Vorteil gegenüber der 3D-Segmentierung von Punktwolken mit sich bringen.

5 Literaturverzeichnis

ALBERTZ, J. & WIGGENHAGEN, M. (2009): Taschenbuch zur Photogrammetrie und Fernerkundung = Guide for photogrammetry and remote sensing / Albertz/Wiggenhagen. Bearb. von Jörg Albertz ; Manfred Wiggenhagen (5. Ausg.). Heidelberg ; München ; Landsberg ; Frechen ; Hamburg: Wichmann .

BRENNER, HAALA, & FRITSCH. (2001): Towards fully automated 3D city model generation. In Automatic Extraction of Man-Made Objects from Aerial and Space Images III.

HAALA, N. (1996): Gebäuderekonstruktion durch Kombination von Bild- und Höhendate. Universität Stuttgart, Fakultät Luft- und Raumfahrttechnik und Geodäsie. Stuttgart: Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften.

JÄHNE, B. (2005): Digitale Bildverarbeitung. (6. überarbeitete und erweiterte Auflag).
KONECNY, K. (2011): Evaluierung der Nutzbarkeit von Punktwolken aus Schrägluftbildern zur Kartierung von Grundrisselementen. Berlin: Beuth Hochschule für Technik Berlin.
KRAUS, K. (1996): Photogrammetrie - Verfeinerte Methoden und Anwendungen (2 Ausg., Bd. Bd. 2.). Berlin: Berlin : de Gruyter.
KRAUS, K. (2004): Photogrammetrie - Geometrische Informationen aus Photographien und Laserscanneraufnahmen (7., vollst. bearb. und erw. Aufl. Ausg.). Berlin: Berlin ; New York : de Gruyter .
LEHMANN, F. a. (2011): MACS - Modular Airborne Camera System for generating photogrammetric high-resolution products. Photogrammetrie Fernerkundung Geoinformation, 435-446.
LUHMANN, T. (2006): Nahbereichsphotogrammetrie. Berlin und Offenbach: Wichmann.
MICROSOFT. (April 2007): http://maps.live.com.
PETRIE, P. G. (2008): Systematic Oblique Aerial Photography Using Multiple Digital Cameras. VIII International Scientific & Technical Conference, (S. 22). Porec, Croatia.
PICTOMETRY. (April 2007): Homepage of Pictometry Inc. Von http://www.pictometry.com/. abgerufen
QIMING ZHOU, Q. Z. (2004): Error Analysis on Grid-Based Slope and Aspect Algorithms. PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING, 957-962.
WEWEL, F. a. (1998): Digitale Luftbildaufnahme mit der HRSC –. Ein Schritt in die Zukunft der Photogrammetrie.
WIEDEMANN, A. (Januar 2011): http://www.bsf-swissphoto.ch. Abgerufen am Januar 2012 von Schrägbildsystem AOS.
WINSTON, P. (1993): Artificial Intelligence. Addison Wesley.

Characterization of forest understory using multi-temporal full-waveform airborne laser scanning

REIK LEITERER[1], FELIX MORSDORF[1] & MICHAEL E. SCHAEPMAN[1]

Zusammenfassung: Der Unterwuchs als Teil der Waldstruktur hat eine wichtige Funktion im Hinblick auf die Dynamik der Waldentwicklung. Allerdings ist die Charakterisierung des Unterwuchses mittels Fernerkundungsmethoden problematisch, da die Vegetationsdichte eine Erfassung der vertikalen Struktur stark limitiert. Unter Verwendung von flugzeuggestütztem, multi-temporalen Laserscanning ist es möglich, den Unterwuchs in einem dichten Laubwald zu detektieren und zu charakterisieren. Basierend auf den geometrischen Informationen der Laser-Punktwolke und den zugehörigen full-waveform Charakteristiken wurden folgende Unterwuchsklassen abgeleitet: vegetationsfreie Flächen, Streu, Unterwuchs < 0.5 m, Unterwuchs 0.5-3 m und Verjüngung > 3 m. Für die Validierung wurde sowohl terrestrisches Laserscanning als auch eine umfangreiche Feldmessung entsprechend dem VALERI Ansatzes verwendet. Die Detektion des Unterwuchses erfolgte mit einer Genauigkeit von 78%; die Klassifikation erreichte eine Genauigkeit von 64%.

1 Introduction

Forests play a pivotal role in the global biogeochemical and -physical cycles and provide a range of essential goods and services (e.g. maintenance of biodiversity, soil and water conservation, wood fuel) [ROSS 2011, DE GROOT et al. 2002]. Particularly the forest structure influences the fluxes of energy and matter between the atmosphere and forests and can serve as a proxy to determine forest stand resistance to disturbances or to estimate the conservation potential for biodiversity [KAYES & TINKER 2012, NADKARNI et al. 2008, XUE et al. 2011].
Understory as part of the forest structure plays a crucial role in forest ecosystems (e.g. limitation for establishment, important habitat component) but is difficult to assess. Conventional fieldwork is time-consuming, often subjective and mostly limited in its spatial extent [FOODY 2010, HAARA & LESKINEN 2009], whereas traditional remote sensing methods can only provide information about the top layer of the canopy [JONES et al. 2012; HALL et al. 2011]. Airborne laser scanning (ALS) systems have shown the potential to provide explicit vertical information about the forest structure due to the canopy penetration of the emitted signal [KAARTINEN et al. 2012, LINDBERG & HOLLAUS 2012, LEEUWEN & NIEUWENHUIS 2010]. However, most of the existing approaches to characterize forest understory based on ALS data focus on the detection and description of understory trees only [KORPELA et al. 2012, MALTAMO 2005, HIRATA et al. 2003] or were applied in areas with mainly open forest, enabling high canopy penetration rates of laser pulses [FERRAZ et al. 2012, MORSDORF et al. 2010, KORPELA 2008]. In dense deciduous forests, particularly under foliate conditions, penetration of the laser signal to lower vegetation layers (understory) is strongly limited. In such cases acquisition under defoliated condition is essential to detect the understory, whereas the leaf-on acquisition can still provide significant information to characterize the detected understory, e.g. based on the light availability.

1) Remote Sensing Laboratories, University of Zurich, Winterthurerstrasse 190, 8057 Zürich, Switzerland; E-Mail: (reik.leiterer, felix.morsdorf, michael.schaepman)@geo.uzh.ch

2 Study area and data

The study site is the Laegeren mountain (forming part of the Swiss Jura; 47°28'N, 8°21'E) and is dominated by steep slopes. The semi-natural and dense, deciduous-dominated forest stand is characterized by a high diversity of species (mostly *Fagus sylvatica (L.)*, *Picea abies (L.) Karst*, and *Fraxinus excelsior (L.)*), age (55-160 years), and diameter distribution (7-120 cm) [EUGSTER et al. 2007]. The data acquisition was performed under foliated (or leaf-on) conditions using a RIEGL LMS-Q680i sensor and under defoliates conditions (or leaf-off) using a RIEGL LMS-Q560 sensor. Sensor specifications and settings used in this study are summarized in Table 1 (for more details see WAGNER et al. [2008] and the technical sensor documentation as provided by RIEGL [RIEGL 2012]).

Table 1: Sensor specifications and settings of the RIEGL LMS-Q560 and LMS-Q680i.

	LMS-Q560	LMS-Q680i
pulse repetition rate [Hz]	200 000 Hz	
scan angle [deg]	± 15 deg	
mean operating altitude above ground [m]	500 m	
mean point density for the area of interest [pts/m^2]	~ 20	~ 40
date of acquisition	10.04.2010	01.08.2010

For the study area, an extensive set of ground based reference data is available, mainly measured during field campaigns in September 2011: Digital Hemispherical Photography's, Terrestrial Laser Scanning (TLS) using an Z+F IMAGER 5006, and extensive field measurements of ground cover and understory vegetation. Additional, soil distribution maps and a species classification were available. All field measurements were geo-referenced and co-registered based on the terrestrial land surveying using a total station and a differential GPS system.

3 Method

The understory detection and characterization consists of three methods: data pre-processing, tree delineation within the understory layer for trees > 3 m; and the classification of understory vegetation/ground cover ≤ 3 m.

3.1 Data pre-processing

In comparison to discrete laser systems, full-waveform systems enable the approximation of the entire backscattered signal by digitization, which facilitates the extraction of additional features for each reflecting object within the ALS footprint [MALLET & BRETAR 2009]. Gaussian pulse estimation [WAGNER et al. 2006] was applied in order to obtain representative echo descriptions, resulting in a point cloud with its basic and established geometrical characteristics of reflectors (range and echo types) as well as in a physical characterization of each reflecting object, including information such as amplitude, width and consequently intensity of each specific echo. First of all, we extracted ground returns from the point cloud using single and last echoes, their geometrical characteristics and the corresponding echo width information [MÜCKE et al. 2010].

To detect and filter remaining vegetation echoes, a new adaptive multi-scale filter algorithm based on the method by EVANS & HUDAK (2007) was developed. As initial part of the filtering, a kernel based query was applied to the ground echoes to identify areas with high deviations in height values (≥ 100 % slope). Within the identified areas, a combination of optimized spline function analyses and a scale dependent (using a 3x3 and/or 5x5 kernel) multi-point triangulation was applied to exclude non-ground echoes from the point cloud. The usage of a spline function approach allows for a reliable distinction between height deviations caused by steep terrain and the more likely less continuous height deviations caused by artificial objects or dense vegetation. Finally, the remaining points were interpolated to a 1x1 m digital terrain model (DTM) based on ordinary kriging. A 1x1 m digital surface model (DSM) was processed using the first echo reflections and their corresponding echo width information. The canopy height model (CHM) was calculated as a subtraction of the DTM from the DSM. For each echo of the point cloud we were now able to determine the height above ground as well as the corresponding DTM, DSM and CHM values (Figure 1). Similar to the processing of the ALS data, we calculated the height above ground for each point of the TLS measurements.

3.2 Understory tree delineation

To identify understory trees, we applied an iterative, three-dimensional grayscale dilation of the point cloud using an oblate spheroid-shaped structuring element (minor axis = 1.5 m, major axis = 5 m) [ADAMS 1993]. A local maximum is assumed to belong to an understory tree, if it is > 3 m above ground and at least 3 m below the specific canopy height at this position. The resulting local maxima were used as initial seed-points for a k-means clustering analysis, which segments the point cloud into clusters of ALS returns presumably belonging to one tree (Figure 2). An Euclidean metric is used to compute distances in the three-dimensional feature space. As such a combination favors ball shaped clusters in a three-dimensional feature space, a re-scaling of the height values was applied to account for aspect ratio of features in the point cloud. Finally, we computed the alpha shapes of the resulting individual clusters of points, which facilitate the derivation of additional variables at the crown level [VAUHKONEN et al. 2009].

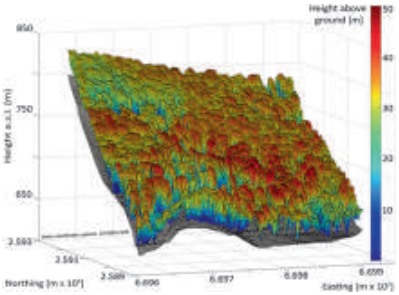


Figure 1: Digital terrain model (in gray) and digital surface model colored with the canopy height information for the area of interest.

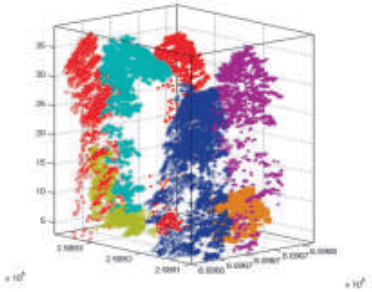


Figure 2: Example of the point cloud segmentation. The green and orange colored point clusters represent detected understory trees.

3.3 Classification of understory vegetation and ground cover

The ground properties and understory vegetation was evaluated on a 1x1 m grid size. For each grid cell the vertical column was subdivided into boxes with a vertical extent of 0.25 m. Within these specific boxes we calculated the percentage amount of echoes, the leaf-on/leaf-off variation in the point cloud and common quantitative statistical measures for the full-waveform variables (e.g. frequency distributions, statistical dispersion and central tendency). The classes 'understory < 0.5 m' and 'understory 0.5 - 3 m' were derived from the echo distribution contained in the boxes of the corresponding height ranges. Considering vertical distribution of scattering objects within one footprint, echoes with larger widths were represented as line segments, oriented in three-dimensional space according to the scan angle. If a line segment (or a part of it) is contained in a box, it is counted as an echo. A box is labeled as vegetation, if the percentage of echoes within the box is at least 5% of all echoes in the height range 0 - 3 m (terrain corrected). Grid cells are defined as 'non-vegetated areas' or 'litter', if > 95% of all echoes in the range 0 - 3 m are located in the 0 - 0.25 m box and > 95% of the echoes have small echo widths and high amplitude values. In addition, the statistical dispersion of the echo types was used as an indicator for light availability, which should indicate the probability of the occurrence of low height vegetation. To distinguish between 'non-vegetated areas' and 'litter', we used a combination of the frequency distribution and central tendency of the echo types, echo widths, and amplitude values.

4 Results and discussion

Figure 3 shows the result of the tree delineation and the understory/ground cover classification (1x1 m grid).

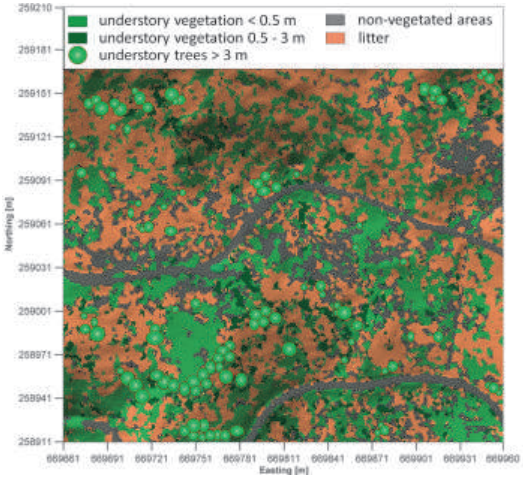


Figure 3: Understory tree delineation and classification result for understory vegetation and ground cover.

The validation of the understory tree detection was based on TLS measurements within two 40x40 m plots and the stratified field measurements outside these areas. The commission and omission error for the delineation is 26.7% and 39.8%, respectively. The commission errors are mainly caused by the detection of crown parts belonging to low reaching crowns of dominant trees, whereby the omission errors have been caused by the large amount of clustered and multi-stemmed trees due to former coppicing activities [VAN CALSTER et al. 2008]. These errors correspond to the findings of existing approaches [e.g. KAARTINEN et al. 2012, FERRAZ et al. 2012] and point out the general limitations for understory tree detection even if multi-temporal ALS data with very high point densities are available. The classification of the understory and the ground cover was compared to a reference data set derived from the TLS data, field measurements and the available plant sociology classification. Utilizing the resulting confusion matrix, we calculated the common statistics metrics overall accuracy, users/producers accuracy and Cohen´s kappa coefficient [LIU et al. 2007] (Table 2).

Table 2: Confusion matrix for the understory/ground cover classification.

Classification / ***Reference***	Understory 0.5-3 m	Understory < 0.5 m	Litter	Non-vegetated areas	***Producer´s accuracy [%]***
Understory 0.5-3 m	**76**	16	1	0	***81.72***
Understory < 0.5 m	21	**49**	6	4	***61.25***
Litter	4	12	**39**	21	***51.31***
Non-vegetated areas	2	7	22	**44**	***58.67***
User´s [%]	***73.77***	***58.33***	***57.35***	***63.77***	
Overall Accuracy [%]: 64.19			**Cohen´s Kappa: 0.52**		

As can be seen from the confusion matrix, we achieved an overall accuracy of ~ 64% and a detection rate of the understory vegetation ≤ 3 m of ~ 89%. The highest degree of class overlap exists for the 'litter' and 'non-vegetated' classes, mainly caused by the similarity of the reflection properties of bare soil and leaf litter. Considering the percentage amount of 'understory trees > 3 m' on the total understory and the related detection rate, we end up with a detection rate for the total understory of 78%.

5 Conclusion

In this study, we present a robust and transferable method to provide a detailed characterization of understory composition and ground cover in a dense, deciduous forest. The validation shows that the detection and characterization of understory/ground cover can be achieved with high accuracy, although the delineation of understory trees is limited by the specific stand characteristics of our study site (e.g. former coppice management, multi-layered crowns). For the land cover classes 'non-vegetated areas', 'litter', 'understory < 0.5 m', and 'understory 0.5-3 m', we achieved a detection rate of 89% for the occurrence of understory and an overall accuracy of 64% for the understory type classification. We conclude that it is possible to detect and characterize forest understory robustly with our method based on ALS data; however, the availability of leaf-on/ leaf-off data with a high point density is indispensable.

Acknowledgements

We acknowledged ESA's STSE program for funding the 3D-VegLab project. IPF Vienna is greatly acknowledged for measuring and pre-processing the TLS data.

6 References

ADAMS, R., 1993: Radial decomposition of discs and spheres. Graphical Models and Image Processing, **55** (5), pp.325-332.

DE GROOT, R.S.; WILSON, M.A. & BOUMANS, R.M.J, 2002: A typology for the classification, description and valuation of ecosystem functions, goods and services. Ecological Economics, **41** (3), pp. 393-408.

EUGSTER, W.; ZEYER, K.; ZEEMAN, M.; MICHNA, P.; ZINGG, A.; BUCHMANN, N. & EMMENEGGER, L., 2007: Nitrous oxide net exchange in a beech dominated mixed forest in Switzerland measured with a quantum cascade laser spectrometer. Biogeosciences Discussion, **4** (2), pp. 1167-1200.

EVANS, J.S. & HUDAK, A.T., 2007: A Multiscale Curvature Algorithm for Classifying Discrete Return LiDAR in Forested Environments. IEEE Transactions on Geoscience and Remote Sensing, **45** (4), pp. 1029-1038.

FERRAZ, A.; BRETAR, F.; JACQUEMOUD, S.; GONÇALVES, G.; PEREIRA, L.; TOMÉ, M. & SOARES, P., 2012: 3-D mapping of a multi-layered Mediterranean forest using ALS data. Remote Sensing of Environment, **121**, pp. 210-223.

FOODY, G.M., 2010: Assessing the accuracy of land cover change with imperfect ground reference data. Remote Sensing of Environment, **114** (10), pp. 2271-2285.

HAARA, A. & LESKINEN, P., 2009: The assessment of the uncertainty of updated stand-level inventory data. Silva Fennica, **43** (1), pp. 87-112.

HALL, F.G.; BERGEN, K.; BLAIR, J.B.; DUBAYAH, R.; HOUGHTON, R.; HURTT, G.; KELLNDORFER, J.; LEFSKY, M.; RANSON, J.; SAATCHI, S.; SHUGART, H.H. & WICKLAND, D., 2001: Characterizing 3D vegetation structure from space: Mission requirements. Remote Sensing of Environment, **115** (11), pp. 2753-2775.

HIRATA, Y.; SATO, K.; SAKAI, A.; KURAMOTO, S. & AKIYAMA, Y., 2003: The extraction of canopy-understory vegetation-topography structure using helicopter-borne LIDAR measurement between a plantation and a broad-leaved forest. Proceedings of the Geoscience and Remote Sensing Symposium, **5**, pp. 3222- 3224.

JONES, T.G.; COOPS, N.C. & SHARMA, T., 2012: Assessing the utility of LiDAR to differentiate among vegetation structural classes. Remote Sensing Letters, **3** (3), pp. 231-238.

KAARTINEN, H.; HYYPPÄ, J.; YU, X.; VASTARANTA, M.; HYYPPÄ, H.; KUKKO, A.; HOLOPAINEN, M.; HEIPKE, C.; HIRSCHMUGL, M.; MORSDORF, F.; NÆSSET, E.; PITKÄNEN, J.; POPESCU, S.; SOLBERG, S.; WOLF, B.M. & WU, J.C., 2012: An international comparison of individual tree detection and extraction using airborne laser scanning. Remote Sensing, **4** (4), pp. 950-974.

KAYES, L.J. & TINKER, D.B., 2012: Forest structure and regeneration following a mountain pine beetle epidemic in southeastern Wyoming. Forest Ecology and Management, **263**, pp. 57-66.

KORPELA, I.; HOVI, A. & MORSDORF, F., 2012: Understory trees in airborne LiDAR data - Selective mapping due to transmission losses and echo-triggering mechanisms. Remote Sensing of Environment, **119**, pp. 92-104.

KORPELA, I., 2008: Mapping of understory lichens with airborne discrete-return LiDAR data. Remote Sensing of Environment, **112** (10), pp. 3891-3897.

LEEUWEN, M. & NIEUWENHUIS, M., 2010: Retrieval of forest structural parameters using LiDAR remote sensing. European Journal of Forest Research, **129** (4), pp. 749-770.

LINDBERG, E. & HOLLAUS, M., 2012: Comparison of methods for estimation of stem volume, stem number and basal area from airborne laser scanning data in a hemi-boreal forest. Remote Sensing, **4** (4), pp. 1004-1023.

LIU, C.; FRAZIER, P. & KUMAR, L., 2007: Comparative assessment of the measures of thematic classification accuracy. Remote Sensing of Environment, **107**, pp. 606-616.

MALLET, C. & BRETAR, F., 2009: Full-waveform topographic lidar: State-of-the-art. ISPRS Journal of Photogrammetry and Remote Sensing, **64** (1), pp. 1-16.

MALTAMO, M.; PACKALÉN, P.; YU, X.; EERIKÄINEN, K.; HYYPPÄ, J. & PITKÄNEN, J., 2005: Identifying and quantifying structural characteristics of heterogeneous boreal forests using laser scanner data. Forest Ecology and Management, **216**, (1-3), pp. 41-50.

MÜCKE, W.; BRIESE, C. & HOLLAUS, M., 2010. Terrain echo probability assignment based on full waveform airborne laser scanning observables. Int. Archives of Photogrammetry and Remote Sensing, **XXXVIII/7A**, pp. 157-162.

MORSDORF, F.; MÅRELL, A.; KOETZ, B.; CASSAGNE, N.; PIMONT, F.; RIGOLOT, E. & ALLGÖWER, B., 2010: Discrimination of vegetation strata in a multi-layered Mediterranean forest ecosystem using height and intensity information derived from airborne laser scanning. Remote Sensing of Environment, 114 (7), pp. 1403-1415.

NADKARNI, N.M.; MCINTOSH, A.C.S. & CUSHING, J.B., 2008: A framework to categorize forest structure concepts. Forest Ecology and Management, **256** (5), pp. 872-882.

ROSS, A.N., 2011: Boundary-layer flow within and above a forest canopy of variable density. Quarterly Journal of the Royal Meteorological Society, DOI:10.1002/qj.989.

RIEGL, 2012: Products. Airborne Scanning. Datasheets. < http://www.riegl.com>.

VAUHKONEN, J.; TOKOLA, T.; PACKALÉN, P. & MALTAMO, M., 2009: Identification of scandinavian commercial species of individual trees from airborne laser scanning data using alpha shape metrics. Forest Science, **55** (1), pp. 37-47.

WAGNER, W.; HOLLAUS, M.; BRIESE, C. & DUCIC, V.: 2008: 3D vegetation mapping using small-footprint full-waveform airborne laser scanners. International Journal of Remote Sensing, **29**(5), pp. 1433-1452

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** (2), pp. 100-112.

XUE, B.-L.; KUMAGAI, T.; IIDA, S.; NAKAI, T.; MATSUMOTO, K.; KOMATSU, H.; OTSUKI, K. & OHTA, T., 2011: Influences of canopy structure and physiological traits on flux partitioning between understory and overstory in an eastern Siberian boreal larch forest. Ecological Modelling, **222** (8), pp. 1479-1490.

Detection of Stationary Vehicles in Airborne Decimeter Resolution SAR Intensity Images using Morphological Attribute Filters

OLIVER MAKSYMIUK [1], ANDREAS BRENNER [2] & UWE STILLA [1]

Abstract: The increasing resolution of modern SAR sensors leads to new approaches for object detection in SAR images. Objects appear not just as blobs but rather show a certain structure which can be exploited by advanced approaches already established in image processing. But an increase in detail comes at the cost of an increased complexity. In this work we use Mathematical Morphology (MM) to detect stationary vehicles in airborne decimeter resolution SAR intensity images. Mathematical Morphology has become popular because of its proven utility and sound mathematical foundations. Here we use the concept of connected components to simplify the image and to describe parts of the vehicles. The detection is a subsequent process. It is shown that a robust detection based on the explicit description of the vehicle appearance is possible.

Zusammenfassung: Bedingt durch die hohe Auflösung moderner SAR Sensoren ergeben sich viele neue Ansätze für die Detektion von Objekten. Es ist nunmehr möglich abseits von statistischen Methoden auch leistungsfähige Verfahren aus der Bildverarbeitung zu verwenden. In dieser Arbeit werden Methoden aus der mathematischen Morphologie verwendet um in SAR Intensitätsbildern stationäre Fahrzeuge zu detektieren. Dabei wird vor allem von Attributfiltern Gebrauch gemacht, die basierend auf den verbundenen Grauwertkomponenten eine Bildtransformation durchführen. Das Ergebnis ist eine Objektbeschreibung in der eine Komposition relevanter Fahrzeugmerkmale zur finalen Detektion genutzt wird. Es wird gezeigt, dass sich mittels dieser Zerlegung und anschließender Korrelation ein robustes Detektionssystem ergibt.

1 Introduction

The increase of the overall traffic volume imposes new challenges on traffic research and planning. Therefore, airborne and spaceborne imaging systems that provide synoptic views of complex traffic situations and the associated context have become interesting (STILLA et al., 2004) and can be considered as complementary systems to road sensors. In recent years, many approaches for detection of moving stationary and moving vehicles using data from different sensors have been developed, e.g. from optical satellite imagery (LEITLOFF et al., 2010), aerial image sequences (LENHART et al., 2008), aerial thermal videos (KIRCHHOF et al., 2006), and LiDAR point clouds (YAO et al, 2011).

1) Photogrammetrie & Fernerkundung, Technische Universität München (TUM), Arcisstraße 21, 80333 München, http://www.pf.bv.tum.de
2) Fraunhofer Institut für Hochfrequenzphysik und Radartechnik FHR, Fraunhoferstraße 20, 53343 Wachtberg, http://www.fhr.fraunhofer.de

Furthermore, there are furthermore many investigations in the context of Moving Target Detection (MTI) based on along track interferometry (CERUTTI-MAORI et al., 2008) but the detection of stationary vehicles has received only little attention so far. In (MAKSYMIUK et al., 2012) it was shown that robust detection of stationary vehicles in SAR intensity images is possible. The approach makes use of supervised learning and employs a set of different features ranging from statistical to spectral features as well as clustering mechanisms. It is a property of the former approach that there is no explicit formulation of the characteristic parts of the object and that most of the features capture additionally some of the surroundings which do not belong to the object itself. In supervised learning a set of features is generated and based on a training set the algorithms determine the importance of the features. Nevertheless it is not enforced that the features capture only the characteristic or essential parts of the objects, although the machine learning procedures ensure that the features are suitable to distinguish between the classes. In this work we want to perform detection only on the characteristic parts which were examined explicitly a priori for the different classes. A multiscale approach is used to represent an object by connected gray level components and therefore as a set of bright and dark components. Since the body of vehicles generally consists of metal it is expected that to cause strong backscattering which in turn is represented as bright connected component in the image. Since SAR images suffer from speckling pixel based methods are likely to fail. A better approach is to work on a higher level than on single pixel values. For this reason it seems promising to work on connected gray level components. They implicitly contain additional information about the objects' shape.

2 Mathematical Morphology and Image Representation

In this section, we present some basics of Mathematical Morphology (MM), the hierarchical image representation by Min- or Max-trees and attribute filters. These are essential for our proposed detection approach, which is introduced afterwards. We will show how to extract relevant features from the images and use them to detect stationary vehicles.

2.1 Morphological Operators and Connected Components

Image segmentation is a low-level processing task and of critical importance for most high-level applications such as object detection. It aims to identify and separate different regions of the images by their visual properties. Ideally the processing should not introduce artificial structures or other kind of artifacts. In this work we use a morphological decomposition related to granulometry and morphological reconstruction to divide the image into connected components with respect to gray levels. Generalized to a gray level image $f(x,y)$ where x, y are the pixel coordinates, a connected component at level t is defined as a set of pair wise connected pixels from the cross section of f at level t, which is the set of all pixels whose values are greater than or equal to t. Fig. 1 shows an example of connected components for a gray level image showing three vehicles oriented across track. At level 0 the whole image is a connected component. In general connected components as image segmentation result are of interest because of the assumption that pixels which are in a connected component belong to the same object.

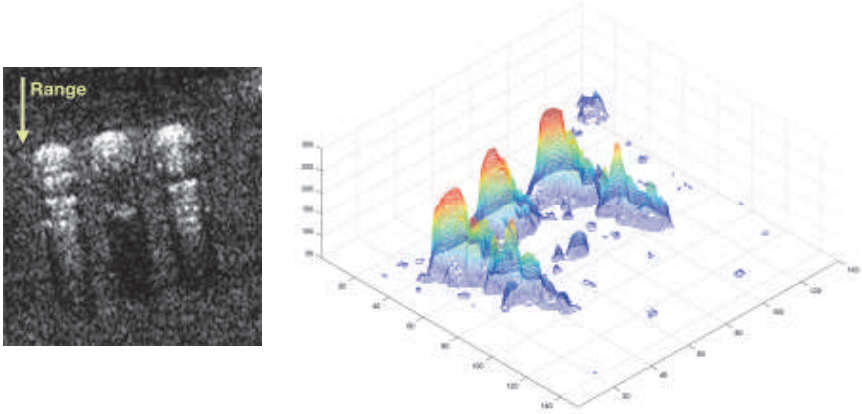


Fig. 1: Image with three vehicles oriented across track (left) and their corresponding connected components at different grayscale levels (right). For better visual representation it is obtained from the median filtered image and only the upper gray levels are presented. The property being a convex set is obvious in this representation.

2.2 Attributes and Attribute Filters

The first morphological reconstruction methods were based on the size of the objects. The well known operator *opening by reconstruction* removes all objects which do not meet the following criterion T, which uses the size as attribute: *the given structure element must fit inside the region*. This criterion is *increasing* since if a connected set meets this criterion then all its supersets also satisfy it. Other properties of the openings are that they are *anti-extensive* and *idempotent*. But since the majority of descriptive criterions like e.g. eccentricity are *nonincreasing,* the attribute opening must be generalized to attribute thinning (BREEN, 1996). However, a nonincreasing criterion T can introduce some artificial structures, which is undesired. In this case the filtering has to take care of this by imposing a kind of increasing property (see section 2.3.1).

2.3 Hierarchical Image Representation

The operators introduced in the last section belong to the class of *connected operators*. These operators describe and filter the image in terms of the *flat zones* rather than single pixels. As already mentioned they do not introduce any new contours and thus are very attractive as information preserving filter. They implicitly work on a data structure called *Max-Tree* (SALEMBIER et al., 1998) for anti-extensive (opening, thinning) or *Min-Tree* for extensive (closing, thickening) operators. This data structure is a hierarchical representation of the image in terms of the connected components, where the pixels assigned to node N_h^k are the smallest superset of all pixels attached to its successors. Here, h is the gray level and k an index variable. This property also explains the multi-scale characteristic since small components are

located near to the leafs whereas the root represents the whole image and is the node with the lowest intensity (Max-tree). It is possible to reconstruct an image Q perfectly from the tree.

2.3.1 Attribute Filtering

Within the Max-tree representation an attribute filtering contains three different stages: construction, filtering, and image restitution. There are several strategies of filtering a Min/Max-Tree with respect to the attached attributes. If the criterion used for filtering is increasing for Max-Trees, all the approaches will deliver the same result. But if the criterion does not meet the condition the attribute filtering becomes more complicated and leads to different results depending on the strategy (DALLA MURA, 2010). In this work we use the s*ubtractive* rule as *nonpruning* strategy removing only nodes which do not satisfy the criterion and update its successors (URBACH, 2007). If a node is removed its pixels are assigned to the highest ancestor meeting the criterion and are therefore lowered in their gray level. The same amount is also used to lower the gray level of the successors. This means that a leaf node having many predecessors which do not meet the criterion has a low gray level at the end of the filtering. This can be interpreted as a kind of *contrast* with respect to the attribute and is an indicator for the exposedness of the component. An example for tree creation, filtering and image restitution is shown in Fig. 2. It should be noted that the tree creation has to be done only once. Attributes can

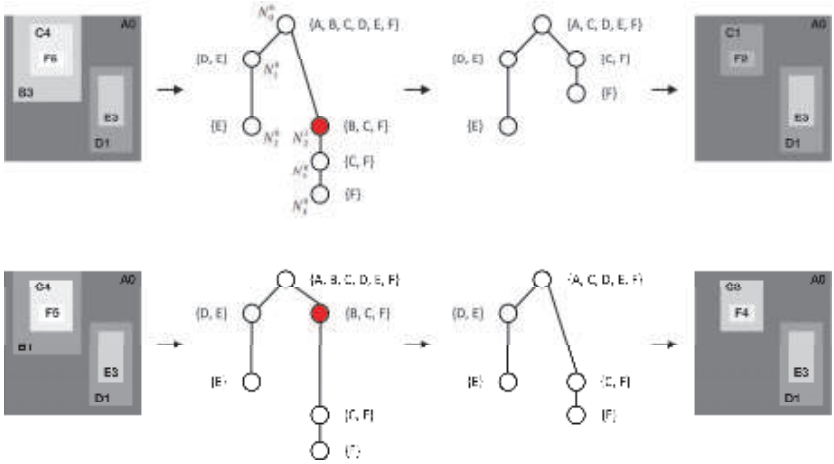


Fig. 2: Exemplary tree creation, subtractive attribute filtering and image restitution. The number given in the identifier of the gray level patches indicates their gray level. The red node (component B) is assumed to fail a criterion and is therefore removed. Note the difference between the upper and lower example with respect to the gray level of components C and F.

be computed at the time of tree creation. After this process the tree can be used to filter for different attributes. Due to this the approach is computational effective.

3 Part-based Representation by Attribute Filters

The set of methods presented in the last section, especially the attribute filters, allow the description of an image as a hierarchy of connected components with additional attributes attached. However the detection of objects which are not a single connected component is not straightforward. Our proposed method is to decompose the class of objects into multiple common parts P_j as can be seen in Fig. 3, where j is some index variable. Since the connected components work on gray levels the parts represent bright (convex) or dark (concave) parts and every part is considered as a single connected component with a set of attributes. Attribute filters (thinning, thickening) allow extracting them easily from the image neglecting the other components and without introducing artificial artifacts but it remains to model their spatial distribution or arrangement. One approach to tackle this problem is to use tree matching procedures to find node structures which best represent the hierarchy and match the attributes (KUMAR et al., 2011). Unfortunately this is a nontrivial problem especially if not all parts of the object are enforced to exist for every single object. Although there are some procedures using heuristics the problem remains difficult. However the problem statement can be reduced in complexity if only the spatial arrangement has to be analyzed. Assuming that the parts match the criterios, a correlation can be used to analyze and assess the position of the parts within the object. For this process we use a mask M_j with a predetermined size (depending on the size of the vehicles) containing a distribution centered at the expected position of the specific part. The assessment of every part is therefore given by $C_j = Q_j * M'_j$, where M'_j is the normalized mask and Q_j the reconstructed image as the result of an attribute filtering. The final detection result is then a linear combination $Y = \sum_j \lambda_j C_j$ with $\sum_j \lambda_j = 1$.

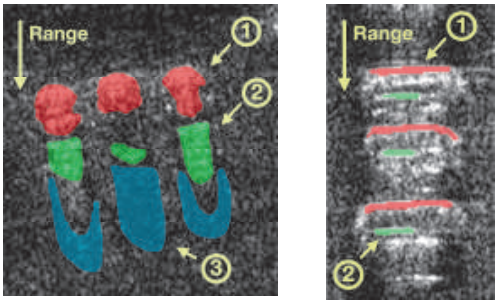


Fig. 3: Decomposition of vehicles of different orientation. Cross track (left) and along track (right). The decomposition is based on bright (convex) and dark (concave) gray level sets. For both classes the components 1 and 2 are convex components residing in the Max-tree whereas component 3 (shadow area) is based on the Min-tree as it is a concave component.

With the linear combination it is possible to put more emphasis on reliable parts e.g. the bright line for vehicles oriented along track is less probable caused by random events than the smaller one below. In order to use the exposedness of the component as a measure in the detection process, Q_j is not normalized. To scale C_j it is simply divided afterwards by $t_{\max}$. It should be noted that there is no investigation in determining a model for an understanding which parts of the vehicle are causing the backscattering visible in the image. One subtle difficulty is the choice of the common parts and their corresponding descriptive attributes. On the one hand they should represent a whole class of objects but should on the other hand allow discriminating between them and other objects in the scene.

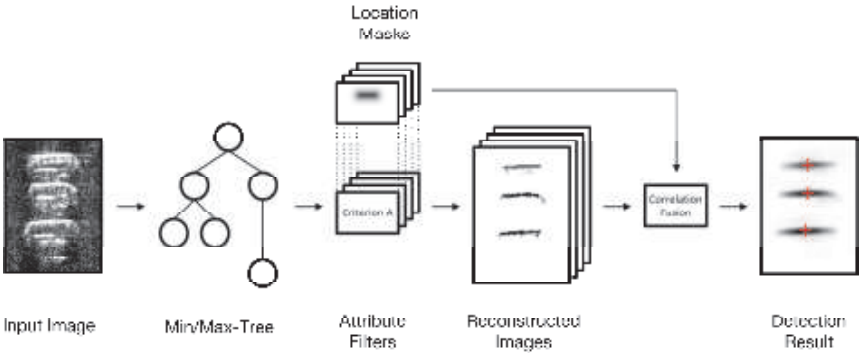


Fig. 4: Flow chart of the vehicle detection based on the decomposition profile. The process of detecting and evaluating the upper bright line is shown. Every attribute filter has its own location mask assessing the position of that part within the object. The final detection result is a linear combination of all part detections. Looking at the reconstructed image it can be clearly seen that only specific parts pass the filter.

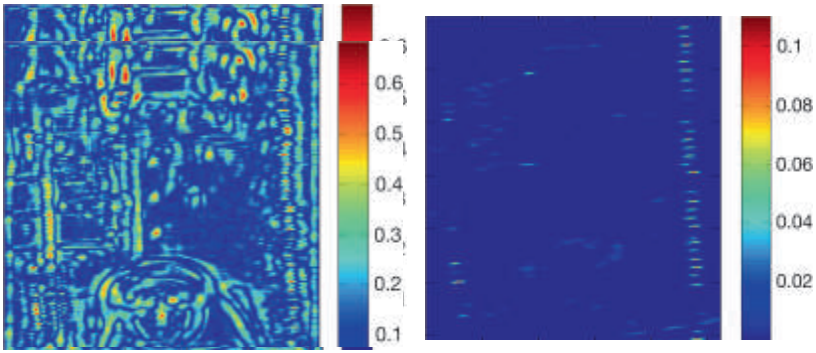


Fig. 5: Comparison between normalized correlation with the mean image of vehicles oriented along track (left) and our proposed approach (right) working with decompositions. Note the difference in scale.

One solution is to apply supervised learning procedures to determine the best representation. If there are not enough data for training a classifier the parameters need to be defined just by user knowledge or inspection of the visual appearance. The advantage of the decomposition is the reduction of artifacts, thus the correlation is performed only on data which is essential and characteristic for the object with high probability. Fig. 5 shows a comparison between a standard correlation and our proposed approach. Obviously the responses from the vehicles in the image are much more distinctive and compared to their local neighborhood making it remarkably easier to detect them.

4 Experiments

The dataset for the evaluation consists of three SAR images obtained from the PAMIR sensor (BRENNER, 2012) working in X-Band with a resolution of 15 cm in range and 7 cm in azimuth. The images contain 77 vehicles oriented along track and 29 vehicles oriented cross track.

4.1 Attributes for Part Based Detection

In this work the attributes and their values are determined statistically by using mean images of vehicles oriented across track and along track. The following attributes are used for components description: Area, direction obtained from main axes, moment of inertia (measure for the elongation of a shape), standard deviation in gray levels, length of the diagonal of the bounding box, and the volume. The latter is just the product of all pixel values residing in the component. These attributes are sufficient to obtain a filtering which removes many components not belonging to vehicles.

4.2 Detection Results

After receiving the result from the fusion of all decomposition parts the local maxima are detected and marked after thresholding with a centered rectangle. The size of the rectangle is choosen according to the size of the location masks. We used a mask size of 35 times 70 pixels for vehicles oriented along track and 80 times 40 pixels for vehicles oriented cross track. The threshold was chosen such that the objects have to exhibit a lower attribute contrast.

5 Conclusions

This paper presented an approach to object detection by decomposing the object into multiple connected components in a morphological sense. Every component has a set of descriptive attributes. The hierarchical representation as Min/Max-Tree allows using powerful attribute filters as information preserving image transformation. The spatial composition of the objects

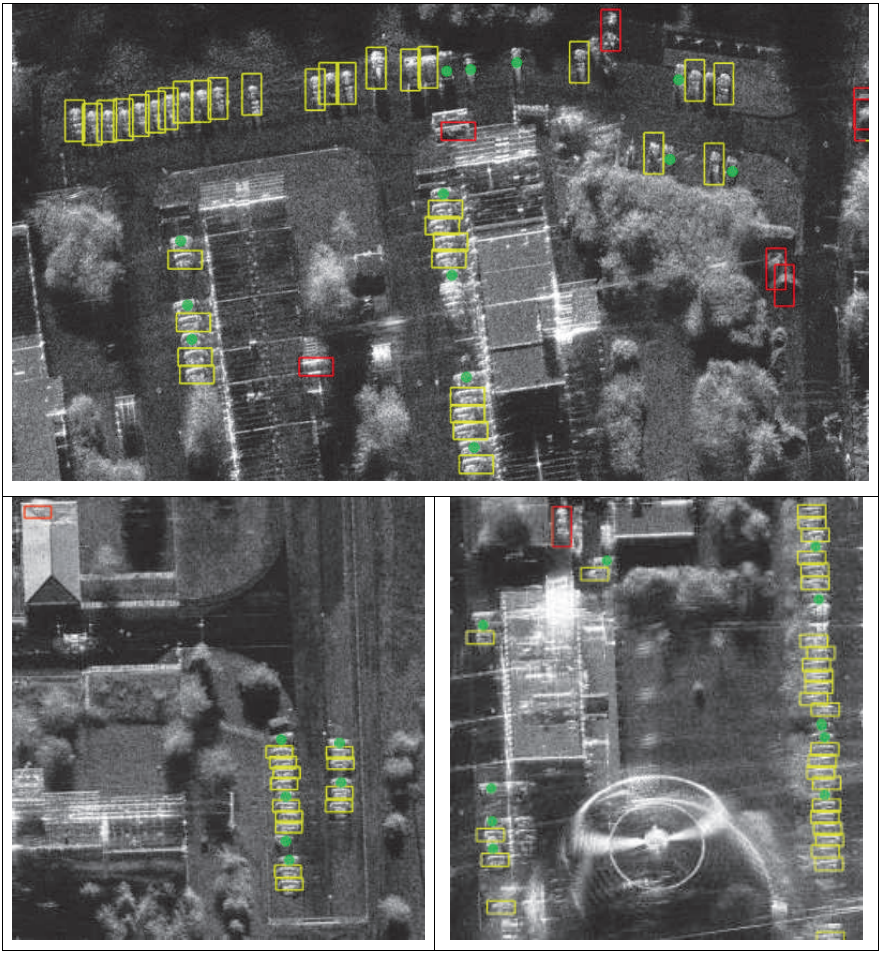


Fig. 6: Detection results of the proposed approach. Although there are some missing detection the method works remarkably well. False detections are marked with a red rectangle. Missing detection are indicated by green points. However, there are no viable ground truth data.

parts is handled by using simple correlation with location masks. The detection performance is only slightly lower than the more sophisticated approach using Support Vector Machines or Random Forests in a supervised learning setup (MAKSYMIUK et al., 2012). It can thus serve as a approach to guide more sophisticated by computationally expensive methods by ruling out areas

which contain no vehicle. In future both approaches can also be combined using the connected components and their attributes as input for supervised learning classifiers. As a consequence one may conclude that the use of higher level features in SAR images has advantages to the use of single pixel values. Further investigations are necessary to determine how to model the objects' parts in the best manner.

6 References

BREEN, E. J. & JONES, R., 1996: Attribute Openings, Thinnings, and Granulometries, Computer Vision and Image Understanding, **64** (3), pp. 377-389.

BRENNER, A. R., 2010: Proof of concept for airborne SAR imaging with 5 cm resolution in X-band, 8th European Conference on Synthetic Aperture Radar (EUSAR), pp. 1-4.

BRENNER, A. R.; ESSEN, H. & STILLA, U., 2012: Representation of stationary vehicles in ultra-high resolution SAR and turntable ISAR images, 9th European Conference on Synthetic Aperture Radar (EUSAR), pp. 1-4.

CERUTTI-MAORI, D.; KLARE, J.; BRENNER, A. R. & ENDER, J. H. G., 2008: Wide-Area Traffic Monitoring with the SAR/GMTI System PAMIR, IEEE Trans. on Geoscience and Remote Sensing, **46** (10), pp. 3019-3030.

DALLA MURA, M.; BENEDIKTSSON, J. A.; WASKE, B. & BRUZZONE, L., 2010: Morphological Attribute Profiles for the Analysis of Very High Resolution Images, IEEE Trans. on Geoscience and Remote Sensing, **48** (10), pp. 3747-3762.

KIRCHHOF, M. & STILLA, U., 2006: Detection of moving objects in airborne thermal videos, ISPRS Journal of Photogrammetry and Remote Sensing, **61** (34), pp. 187-196

KUMAR, R.; TALTON, J. O.; AHMAD, S.; ROUGHGARDEN, T. & KLEMMER, S. R.,2011: Flexible Tree Matching, International Joint Conference on Artificial Intelligence

LEITLOFF, J.; HINZ, S. & STILLA, U., 2010: Vehicle Detection in Very High Resolution Satellite Images of City Areas, IEEE Trans. on Geoscience and Remote Sensing, **48** (7), pp. 2795-2806

LENHART, D.; HINZ, S.; LEITLOFF, J. & STILLA, U., 2008: Automatic traffic monitoring based on aerial image sequences, Pattern Recognition and Image Analysis, **18** (3), pp. 400-405

MAKSYMIUK, O.; SCHMITT, M.; BRENNER, A. R. & STILLA, U., 2012: First Investigations on Detection of Stationary Vehicles in Airborne Decimeter Resolution SAR Data by Supervised Learning, Proc. IEEE Geoscience and Remote Sensing Symposium (IGARSS), pp. 3584-3587.

PESARESI, M. & BENEDIKTSSON, J. A., 2001: A New Approach for the Morphological Segmentation of High-Resolution Satellite Imagery, IEEE Trans. on Geoscience and Remote Sensing, **39** (2), pp. 309-320.

SALEMBIER, P.; OLIVERAS, A. & GARRIDO, L., 1998: Antiextensive Connected Operators for Image and Sequence Processing, IEEE Trans. on Image Processing, **7** (4), pp. 555-570.

SALEMBIER, P. & GARRIDO, L., 2000: Binary Partition Tree as an Efficient Representation for Image Processing, Segmentation, and Information Retrieval, IEEE Trans. on Image Processing, **9** (4), pp. 561-576.

SOILLE, P. & PESARESI, M., 2002: Advances in Mathematical Morphology Applied to Geoscience and Remote Sensing, IEEE Trans. on Geoscience and Remote Sensing, **40** (9), pp. 2042-2055.

STILLA, U.; MICHAELSEN, E.; SOERGEL, U.; HINZ, S. & ENDER,J. H. G., 2004: Airborne monitoring of vehicle activity in urban areas, International Archives of Photogrammetry and Remote Sensing, **35**, pp. 973-979.

URBACH, E. R; ROERDINK, J. B. T. M. & WILKINSON, M. H. F, 2007: Connected Shape-Size Pattern Spectra for Rotation and Scale-Invariant Classification of Gray-Scale Images, IEEE Trans. on Pattern Analysis and Machine Intelligence, **29** (2), pp. 272-285.

YAO, W. & STILLA, U. 2011: Comparison of Two Methods for Vehicle Extraction From Airborne LiDAR Data Toward Motion Analysis, IEEE Geoscience and Remote Sensing Letters, **8** (4), pp. 607-611.

Erkennen von liegendem Totholz in full-waveform Airborne Laserscanning Daten

WERNER MÜCKE[1], ANKE SCHROIFF[2], MARKUS HOLLAUS[1]& NORBERT PFEIFER[1]

Zusammenfassung:

Totholz ist einer der Hauptindikatoren für die Zustandsbewertung von Lebensräumen und Biodiversität in Wäldern. Es bietet verschiedensten Organismen Nahrung, Schutz und Unterkunft. Vermoderndes Totholz ist außerdem ein wichtiger Kohlenstoffspeicher und liefert wertvolle Nährstoffe für den Waldboden, was für die Verjüngung des Waldes notwendig ist. Die Erfassung von Totholz ist daher Teil der nachhaltigen Forstwirtschaft und von ökologischem Monitoring im Sinne der Fauna-Flora-Habitat (FFH) Richtlinie. In der Praxis gestaltet sich der Erfassungsprozess jedoch oft als zeitaufwändig, ohne großen Personalaufwand nicht flächendeckend durchführbar und kostenintensiv.

In dieser Studie wird gezeigt, dass sich die großflächige Erfassung von liegendem Totholz mit Airborne Laserscanning (ALS) unterstützen lässt. Auf der Basis von full-waveform ALS-Punktwolken wird liegendes Totholz auf dem Waldboden identifiziert. Zuerst werden, unter der Nutzung der Echobreite aus der Aufzeichnung der full-waveform, jene ALS-Punkte selektiert, welche den Waldboden inklusive der umgefallenen Stämme repräsentieren. Damit wird bodennahe Vegetation eliminiert, um ein detailliertes Höhenmodell des Waldbodens und des Totholzes zu erstellen. Dieses wird anschließend mit morphologischer Bildverarbeitung und Rasterkarten-Algebra aufbereitet, so dass sich eine binäre Karte des liegenden Totholzes ergibt. Zusätzlich kann anhand von geometrischen Kriterien zwischen noch gut erhaltenen und bereits stark vermoderten Stämmen oder Zweighaufen unterschieden werden.

Zum Trainieren und Testen der Methode steht eine Untersuchungsfläche in einem FFH-Gebiet in Ungarn zur Verfügung. Dort wurden im Rahmen des EU-Projektes „ChangeHabitats2" terrestrische Referenzdaten aufgenommen (GPS-Positionen der liegenden Stämme, Fotos, etc.). Ende März 2012 fanden in diesem Gebiet ALS-Befliegungen mit einem RIEGL LMS-Q680i full-waveform Laserscanner statt, jeweils im belaubten und im laublosen Zustand. Die so erstellte 3D-Punktwolke mit einer Punktdichte von mehr als 20 Punkten pro m² erlaubt eine direkte Identifikation der liegenden Stämme.

In der Studie wird gezeigt, dass mithilfe der full-waveform ALS-Daten liegendes Totholz im Wald mit guter Vollständigkeit (>70%) gefunden werden kann. Die Methode funktioniert auch dort, wo die umgefallenen Stämme bereits von umliegender Vegetation überwachsen werden. Durch den vorgestellten Ansatz kann das Vorhandensein und die Menge von liegendem Totholz großflächig abgeschätzt und beispielsweise auch Flächen ausgewiesen werden, wo sich (durch Schädlingsbefall oder Windwurf) signifikante Änderungen ergeben haben.

1) Werner Mücke, Markus Hollaus und Norbert Pfeifer, Department für Geodäsie und Geoinformation, Forschungsgruppen Photogrammetrie und Fernerkundung, Technische Universität Wien, Gußhausstraße 27-29, A-1040 Wien. (wm,mh,np)@ipf.tuwien.ac.at

2) Anke Schroiff, YGGDRASIL Diemer, Dudenstraße 38, D-10965 Berlin. ankeschroiff@yahoo.de

1 Einleitung

Natürliche und naturnahe Wälder in Europa unterliegen dem Schutz der Fauna-Flora-Habitat Richtlinie (FFH) und sind Bestandteil des Natura2000-Netzwerkes (N2000). Das Hauptziel von N2000 besteht im langfristigen Erhalt der europäischen Biodiversität. Das Vorkommen von Totholz ist eine wesentliche Voraussetzung für die Biodiversität in Wäldern, da zahlreiche Tiergruppen, Pilze und Pflanzen daran gebunden sind (PAILLET et al. 2010). In Totholz finden holzbewohnende Wirbellose Nahrung und Lebensraum (LASSAUCE et al. 2011; RIFFELL et al. 2011), die ihrerseits eine wichtige Beute für zahlreiche Vogelarten darstellen (LOHR et al. 2002). Totholz bietet Bruthabitate, Schutz, Wanderkorridore und Futter für viele Vogel- und Säugetierarten (BUTTS & MCCOMB 2000; SALLABANKS et al. 2001; CHRISTENSEN et al. 2005). Es stellt außerdem günstige Bedingungen hinsichtlich Mikroklima, Feuchte und Nährstoffe für Pilz-, Flechten-, Moos- und einige Gefäßpflanzenarten (JUUTILAINEN et al. 2011). Darüber hinaus erfüllt Totholz als Kohlenstoffspeicher wichtige Ökosystemleistungen mit großer Bedeutung für die Klimaregulierung (WOODALL et al. 2008). Die Erfassung von Totholz ist daher Teil von naturschutzfachlichem Monitoring, sowie von nachhaltiger Forstwirtschaft. Die Aufnahme von Totholz im Gelände erfordert jedoch einen großen Personalaufwand und ist damit zeit- und kostenintensiv, vor allem in sehr großen oder schwer zugänglichen Waldgebieten. Die FFH-Richtlinie schreibt den Mitgliedsstaaten einen regelmäßigen Monitoringzyklus von 6 Jahren für alle N2000 Gebiete vor (aktuell ca. 583888 km² oder 13,4 % der EU-Fläche). Diese Aufgabe ist angesichts der Ausdehnung des Gebietes nur schwer bis gar nicht in der vorgegebenen Zeit zu bewältigen.

Das von der EU finanzierte und noch bis Ende 2014 laufende Projekt „*ChangeHabitats2"* (www.changehabitats.eu) hat die Entwicklung von FFH-Monitoringmethoden zum Ziel, die auf Fernerkundungstechnologien beruhen. Methoden aktiver (Airborne Laser Scanning – ALS) und passiver (Hyperspektralbilder) Fernerkundung sollen benutzt werden, um das konventionelle Bewertungsverfahren zu unterstützen und, sofern sinnvoll und möglich, zu ersetzen.

In dieser Studie soll die Eignung von ALS-Daten für die Identifizierung von Totholz in Wäldern nachgewiesen werden. Bereits existierende Arbeiten zu diesem Thema basieren fast ausschließlich auf statistischen Methoden (z.B. mittels Regressionsanalyse). Für die Modellkalibrierung ist dabei eine bestimmte Menge von Gelände- oder Referenzdaten nötig. VEHMAS et al. (2011) nutzten aus ALS ermittelte Höhenangaben für das Auffinden von Lücken im Kronendach, in denen sich signifikante Mengen von starkem Totholz (CWD, engl. Coarse Woody Debris) befanden. BATER et al. (2009) identifizierten den Variationskoeffizient von ALS-Punkthöhen als beste Variable für die Vorhersage des Totholzanteils. KIM et al. (2009) untersuchten die Möglichkeit, ALS-Daten für die Unterscheidung von lebenden und toten Bäumen zu verwenden und ermittelten die Echo-Intensität (oft auch als Echo-Amplitude bezeichnet) als das entscheidende Maß für eine genaue Einschätzung. Auch PESONEN et al. (2008) verwendeten ALS-Punktverteilungsmaße für die Vorhersage von stehendem und liegendem Totholzvorrat. Nach bisherigem Kenntnisstand gibt es aber lediglich eine Studie, die die Eignung von ALS für die direkte Kartierung von umgefallenen Bäumen in der Punktwolke getestet hat. BLANCHARD et al. (2011) haben Methoden der objekt-basierten Bildanalyse für die Kartierung von liegenden Stämmen in gerasterten ALS-Daten verwendet.

Im vorliegenden Artikel wird ein schrittweiser Prozess für die direkte Extraktion von liegendem Totholz vorgestellt, der auf der Kombination von Punktwolken-Filterung, morphologischer Bildverarbeitung und Rasterkarten-Algebra basiert. Analyse und Verarbeitung der Daten wurden mit Hilfe von Mathworks Matlab (www.mathworks.com), OPALS (www.ipf.tuwien.ac.at/opals) und ESRI ArcGIS (www.esri.com/software/arcgis) durchgeführt.

2 Testgebiet und verfügbare Daten

2.1 Testgebiet „Nagyerdö“ in Ungarn

Die Untersuchung wurde in Nagyerdo (Ungarn, bei Debrecen, N2000-Gebiet „Debrecen-Hajdúböszörményi tölgyesek“), eines der CH2-Projektgebiete, durchgeführt. Es umfasst die Lebensraumtypen 91I0 (Euro-Sibirische Eichen-Steppenwälder) und 91F0 (Hartholzauenwälder). Neben dem natürlichen Bestand gibt es Pflanzungen mit Robinie (*Robinia pseudo-acacia*), Rot-Eiche (*Quercus rubra*) und Kiefer (*Pinus sylvestris*). Der untersuchte Ausschnitt beschränkt sich auf den Lebensraumtyp Eichen-Steppenwälder. Der Waldbestand ist ca. 120 Jahre alt, mehrschichtig (2 Kronenschichten, dichte Strauchschicht, Krautschicht), wird nicht bewirtschaftet und hat einen hohen Anteil an Totholz (Ausschnitt siehe Abb. 1).

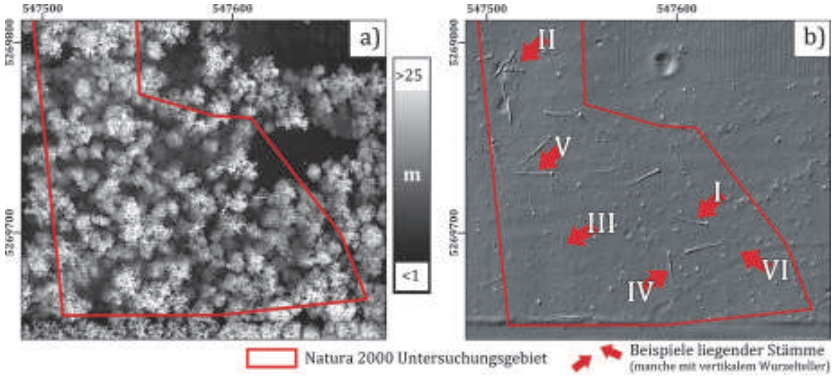


Abb. 1: Übersicht des Untersuchungsgebietes. (a) normalisiertes Höhenmodell; (b) Schummerung des $nDHM_{cand}$ mit Beispielen von liegenden Stämmen (I-VI).

2.2 Erhobene Felddaten

Im Frühjahr, Sommer und Winter 2012 wurden vor Ort Felddaten für die N2000-Bewertung erhoben. Liegendes Totholz wurde durch seinen Anfangs- und Endpunkt bestimmt. Die Messung erfolgte mittels differentiellem globalem Navigationssateliten System (dGNSS). Zusätzlich wurden Durchmesser und Länge mittels Maßband bestimmt. Baumstümpfe und Zweighaufen umgefallener Bäume wurden durch eine dGNSS-Punktmessung ihres Mittelpunktes erfasst. Alle Totholzfunde wurden photographisch dokumentiert (inklusive Kompassablesung des Azimut).
Der Zersetzungsgrad des Totholzes wurde wie folgt eingeschätzt:

Grad 1: Kürzlich umgestürzter Baum, kaum oder noch gar keine Zersetzung, Rinde intakt, hartes Holz.
Grad 2: Älteres Totholz, stellenweise beginnende Zersetzung, lose Rinde, stellenweise weiches Holz.
Grad 3: Altes Totholz, Zersetzung entlang des gesamten Stammes, keine Rinde mehr, weiches Holz, stellenweise in Stücke zerfallen.

Insgesamt wurden 86 liegende Stämme und Zweighaufen in Nagyerdö terrestrisch erhoben.

2.3 Airborne Laserscanning-Daten

Die ALS-Daten über Nagyerdő wurden im Frühling 2012 (28. März, unbelaubt) aufgenommen. Zum Einsatz kam ein Riegl LMS-Q680i full-waveform-Laserscanner (www.riegl.com). Der maximale Scanwinkel betrug 30° und die Streifenüberlappung zumindest 50%. Mit der Pulswiederholrate von 400kHz und einer mittleren Flughöhe von 800 m wurde eine Punktdichte von 29,4 Echos/m² erreicht (alle Echos gezählt). Durch Gaußsche Zerlegung (WAGNER et al. 2006) in RiPROCESS (www.riegl.com) wurden die Echos aus den Wellenformen extrahiert.
Ein digitales Geländemodell (DGM) mit einer Gitterweite von 25 cm wurde mittels hierarchischer robuster Filterung (KRAUS & PFEIFER 1998) bestimmt und anschließend wurden die normalisierten Punkthöhen (Höhen der Echos über dem Boden) bestimmt. Die Punktdaten umfassen somit für jedes Echo: 3D Koordinaten im übergeordneten System, Echo-Amplitude und -breite (beide aus der Full-Waveform-Analyse), Echonummer und Gesamtzahl der Echos für diesen Schuss, sowie die normalisierte Höhe.

3 Extraktion von liegendem Totholz in der Punktwolke

Der Extraktionsprozess beginnt mit der Selektion von letzten Echos basierend auf deren normalisierter Punkthöhe und deren Echobreite. Die Echobreite liefert Informationen über die Rauigkeit bzw. die Höhenvariation innerhalb der durch den Laserstrahl beleuchteten Oberfläche (WAGNER 2005). Gibt es innerhalb des Lichtkegels viele kleine zurückstreuende Elemente mit Höhenunterschied (z.B. Blätter oder Äste), nimmt die Echobreite zu. Diese Eigenschaft wird genutzt, um zwischen glatten (hier der Waldboden oder die liegenden Stämme, welche als glatte Flächen in Relation zur beleuchteten Fläche erscheinen) und rauen Flächen (hier die niedrige Vegetation oder Krautschicht) zu unterscheiden. Durch das Anbringen von empirisch ermittelten Schwellenwerten können so Echos von bodennaher Vegetation effizient eliminiert werden (DONEUS et al. 2008; RUTZINGER et al. 2008; LIN & MILLS 2009; ALEXANDER et al. 2010; HOLLAUS et al. 2011). Im hier vorgestellten Ansatz wird die Echobreite in doppelter Weise genutzt: (1) als Schwellenwert für alle Echos unterhalb von 2 m normalisierter Höhe und mit Echobreiten kleiner als 4.75 ns und (2) als ein 1x1 m² Rastermodell der Echobreitenvarianz EW_{var}, der Echos zwischen 0.1 m und 2 m normalisierter Höhe für die Klassifikation im letzten Schritt. Der Grenzwert von 2 m entstammt der Tatsache, dass die größten liegenden Stämme im Untersuchungsgebiet inklusive ihrer vertikalen Wurzelteller niedriger sind als 2 m.

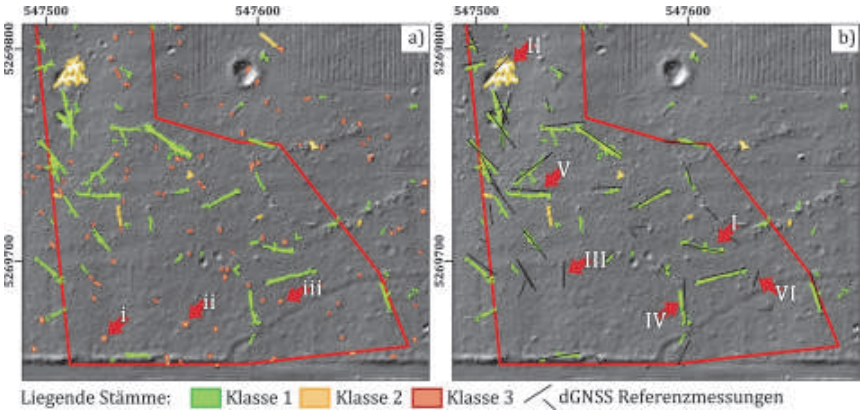


Abb. 2: : Schummerung des $nDHM_{cand}$ überlagert mit dem Resultat der automatischen Extraktion liegender Stämme. (a) inklusive der eliminierten Flächen aus Klasse 3; (b) ohne die Klasse 3 Polygone, aber mit der Referenzklasse der im Feld gemessenen Stämme.

Die so selektierten Echos werden anschließend verwendet, um ein digitales Höhenmodell (DHM) des Waldbodens und des Totholzes zu erstellen (DHM_{cand}). Das DHM_{cand} zeigt den Waldboden inklusive der liegenden Stämme als lineare Objekte (Abb. 1). Anschließend wird ein Differenzenmodell aus dem DHM_{cand} und dem DGM gebildet, das normalisierte DHM_{cand} ($nDHM_{cand}$). Dieses $nDHM_{cand}$ wird in eine binäre Karte konvertiert (1 wenn $nDHM_{cand} > 0.1$ m, sonst 0) und morphologische Bildverarbeitung (closing) wird verwendet, um gefundene Stammobjekte zu verbinden und Löcher zu füllen. Die binäre Karte wird anschließend vektorisiert und im nächsten Schritt werden mithilfe des Umfang-zu-Fläche-Ratio ap_{ratio}, längliche Elemente identifiziert:

$$ap_{ratio} = 4 * \pi * a / p^2$$

wobei a die Fläche und p der Umfang des vektorisierten Objektes ist. Der Wert ap_{ratio} liegt näher bei 1 für runde (kreisförmige) Polygone und näher bei 0 für langgestreckte Polygone. Eine ap_{ratio} von mehr als 0.3 und eine Fläche von weniger als 4 m² führt dazu, dass Polygone automatisch gelöscht werden, da diese Polygone stehende Bäume repräsentieren. Für die weiteren Selektionsschritte wird die Vektorkarte mit der EW_{var}-Rasterkarte verbunden und ein regelbasiertes Klassifikationsschema angewandt. Dieses soll die gut erkennbaren Stämme vom Rest trennen. Es basiert auf dem Formkriterium und der Rauigkeit der Oberfläche:

Klasse 1: Objekt ist länglich und hat niedrige Rauigkeit: $ap_{ratio} < 0.3$ und $EW_{var} < 1$.
Klasse 2: Objekt ist länglich und hat mittlere bis hohe Rauigkeit: $ap_{ratio} < 0.3$ und $EW_{var} > 1$.
Klasse 3: Objekt ist nicht länglich und hat mittlere bis hohe Rauigkeit: $ap_{ratio} >= 0.3$ und $EW_{var} > 1$.

Klasse 1 soll die gut erkennbaren Stämme beinhalten, wohingegen Klasse 3 Flächen ausweisen soll, die niedrige Vegetation darstellen. In Klasse 2 sollen jene Stämme fallen, die zwar prinzipiell durch die längliche Form erkennbar, aber stark überwachsen sind und daher eine höhere Rauigkeit aufweisen. Objekte, die nicht länglich sind, eine geringe Rauigkeit aufweisen, und größer als 4 m² sind, sind im untersuchten Datensatz kaum enthalten und nicht von weiterem Interesse. Das endgültige Resultat ist eine Vektorkarte der Umrisslinien von extrahiertem liegenden Totholz (Abb. 2).

4 Evaluierung und Diskussion

Ein rein visueller Vergleich der im Feld mit dGNSS gemessenen Stämme hat Positionsabweichungen von 0.5 m bis 3 m aufgezeigt. Diese sind höchstwahrscheinlich durch Abschattungseffekte, schlechte Satellitenkonstellationen oder andere genauigkeitsmindernde Effekte während der dGNSS-Beobachtungen entstanden. Daher war ein direkter (und automatischer) Vergleich der Referenzdaten und der extrahierten Umrisslinien nicht möglich. Allerdings konnte mit Hilfe eines manuellen Abgleichs von räumlicher Nähe und Form jedem im Feld gemessenen Stamm ein korrespondierendes längliches Objekt im $nDHM_{cand}$ zugeordnet werden. Die Evaluierung wurde daher wie folgt durchgeführt: ein im $nDHM_{cand}$ manuell erkannter Stamm wurde

- als gefunden bezeichnet, wenn die Übereinstimmung mit seiner (automatisch extrahierten) korrespondierenden Umrisslinie größer als 75% war,
- als teilweise gefunden bezeichnet, wenn die Übereinstimmung mit seiner korrespondierenden Umrisslinie kleiner als 75% war, und
- als nicht gefunden bezeichnet, wenn keine korrespondierende Umrisslinie gefunden werden konnte.

Automatisch gefundene Stämme, die auch in der Referenz enthalten sind, gelten als „True Positives“ (TP). Stämme, die automatisch gefunden wurden, aber nicht in der Referenz enthalten sind, sind „False Positives“ (FP) und Stämme die nur in der Referenz enthalten sind, sind „False Negatives“ (FN). Completeness (producer’s accuracy) und Correctness (user’s accuracy) wurden anschließend wie folgt berechnet (HEIPKE et al. 1997):

$$\text{Completeness} = ||TP|| / (||TP||+||FN||)$$

$$\text{Correctness} = ||TP|| / (||TP||+||FP||)$$

Ein klassifiziertes Objekt ist ein TP, wenn es gefunden wurde (Übereinstimmung > 75%) und der Zerfallsgrad 1, 2 oder 3 ist. Als TP wurden aber ebenso Objekte gezählt, die teilweise gefunden (Übereinstimmung < 75%) und mit Zerfallsgrad 2 oder 3 bewertet wurden.
Abb. 2a zeigt die resultierenden Umrisslinien der liegenden Stämme inklusive der Klassifikation (Klasse 1 = grün, Klasse 2 = orange, Klasse 3 = rot). Mithilfe der terrestrischen Fotos und durch visuelle Interpretation der ALS-Punktwolke konnte das Klassifizierungsschema evaluiert

werden. Klasse 3 (rot) stellt meistens bodennahe Vegetation dar (dichte Krautschicht) oder Stämme von stehenden Bäumen (inklusive der aufstrebenden Wurzelteile knapp über dem Boden) (Abb. 4 i). Manchmal sind die als Klasse 3 eingestuften Polygone auch abgetrennte Teile von liegendem Totholz, allerdings sind dies so wenige, dass ihr Einfluss vernachlässigt werden kann (Abb. 4 ii + iii). Es wurde festgestellt, dass Klasse 2 (orange) hauptsächlich Zweighaufen oder liegende Stämme beinhaltet, welche besonders dicht überwachsen sind (Abb. 3 II & V). Klasse 1 (grün) repräsentiert die Polygone, die sehr wahrscheinlich liegende Stämme sind (Abb. 3 I & IV). Klasse-3-Polygone wurden in weiterer Folge eliminiert, da sie mit sehr hoher Wahrscheinlichkeit keine liegenden Stämme darstellen. Somit wurden nur Klasse 1 und 2 Polygone für die Qualitätskontrolle berücksichtigt, welche eine Completeness von 75.6% und eine Correctness von 89.9% ergab (62 TPs, 7 FPs und 20 FNs). Langgestreckte und gut erhaltene Stämme (Zersetzungsgrad 1 und 2) können sehr zuverlässig gefunden werden (31 von 40 Stück), unabhängig davon, ob sie von Zweighaufen oder vitaler Vegetation überdeckt sind (Abb. 3 V). Je stärker zersetzt der Stamm ist oder je kleiner sein Durchmesser, desto weniger hebt er sich vom Boden ab, sein Umriss wird undefiniert und es wird unwahrscheinlicher, dass er automatisch gefunden werden kann (Abb. 3 III & VI).

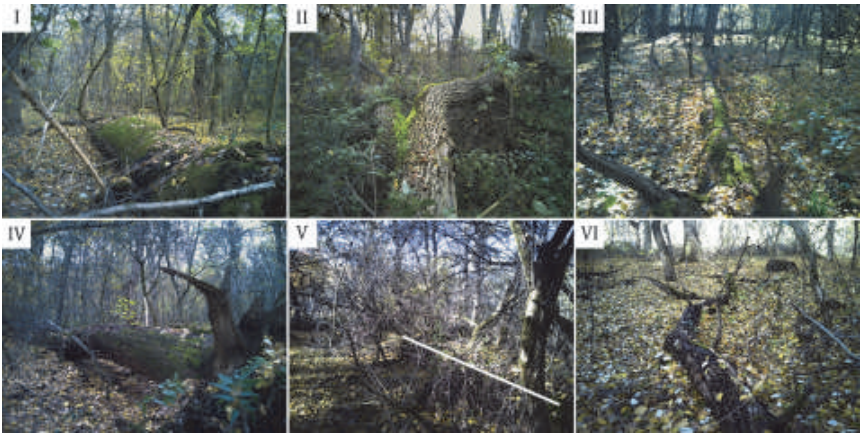


Abb. 3: Beispiele gefundener liegender Stämme im Untersuchungsgebiet. In V wurde der liegende Stamm durch eine weiße Linie gekennzeichnet, da er fast vollständig von Zweigen abgedeckt ist. Abb. 2b zeigt wo diese Bäume gefunden wurden.

5 Zusammenfassung

Diese Studie untersucht die Möglichkeit, basierend auf ALS-Daten liegendes Totholz im Wald automatisch zu finden. Der vorgestellte Workflow produziert eine Vektorkarte, welche die Umrisslinien liegender Stämme enthält. Es wurde gezeigt, dass gut erhaltene Stämme gut erkannt

werden können. Die Echobreite des full-waveform-Laserscanning wurde zur Elimination niedriger Vegetation genutzt.
Effiziente Datenstrukturen für die Speicherung und den Zugriff auf die ALS-Daten vorausgesetzt (wie sie die OPALS Module zur Verfügung stellen), arbeitet die Methode sehr effizient, da sie nach dem anfänglichen Echoselektionsschritt nur mehr auf Rasterkarten oder Standard-GIS-Funktionen zurückgreift. Das Prozessieren einer 1x1 km² großen Fläche mit der hier beschriebenen Punktdichte dauert im Durchschnitt 6 Minuten auf einem Standard PC (QuadCore Intel CPU, 8 GB RAM).
Die vorgestellte Methode eröffnet auch die Möglichkeit, Länge und Durchmesser der erkannten Stämme zu bestimmen, was eine Abschätzung des Totholzvolumens pro Fläche ermöglicht.



Abb. 4: Beispiele gefundener Zweighaufen im Untersuchungsgebiet. Abb. 2a zeigt wo diese gefunden wurden.

6 Danksagung

Die Arbeit wurde durch die Projekte NEWFOR (finanziert durch die “European Territorial Cooperation Alpine Space”) und ChangeHabitats2 (Marie Curie - FP7-PEOPLE-2009-IAPP - Grant Agreement Number 251234) unterstützt. Dank geht auch an die RIEGL Laser Measurements GmbH für die Aufnahme der ALS-Daten und an Prof. Hermann Heilmeier (TU Bergakademie Freiberg) für wertvolle Hinweise während der Entstehung des Artikels.

7 Literaturverzeichnis

ALEXANDER, C., K. TANSEY, J. KADUK, D. HOLLAND & N. J. TATE (2010). "Backscatter coefficient as an attribute for the classification of full-waveform airborne laser scanning data in urban areas." ISPRS Journal of Photogrammetry and Remote Sensing **65**(5): 423-432.

BATER, C. W., N. C. COOPS, S. E. GERGEL, V. L. LEMAY & D. COLLINS (2009). "Estimation of standing dead tree class distributions in northwest coastal forests using lidar remote sensing." Canadian Journal of Forest Research **39**(6): 1080-1091.

BLANCHARD, S. D., M. K. JAKUBOWSKI & M. KELLY (2011). "Object-based image analysis of downed logs in disturbed forested landscapes using lidar." Remote Sensing **3**(11): 2420-2439.

BUTTS, S. R. & W. C. MCCOMB (2000). "Associations of forest-floor vertebrates with coarse woody debris in managed forests of western Oregon." The Journal of wildlife management **64**(1): 95-104.

CHRISTENSEN, M., K. HAHN, E. P. MOUNTFORD, P. ÓDER, T. STANDOVÁR, D. ROZENBERGAR, J. DIACI, S. WIJDEVEN, P. MEYER & S. WINTER (2005). "Dead wood in European beech (Fagus sylvatica) forest reserves." Forest Ecology and Management **210**(1): 267-282.

DONEUS, M., C. BRIESE, M. FERA & M. JANNER (2008). "Archaeological prospection of forested areas using full-waveform airborne laser scanning." Journal of Archaeological Science **35**(4): 882-893.

HEIPKE, C., H. MAYER, C. WIEDEMANN & O. JAMET (1997). "Evaluation of automatic road extraction." International Archives of Photogrammetry and Remote Sensing **32**(3 SECT 4W2): 151-160.

HOLLAUS, M., C. AUBRECHT, B. HÖFLE, K. STEINNOCHER & W. WAGNER (2011). "Roughness Mapping on Various Vertical Scales Based on Full-Waveform Airborne Laser Scanning Data." Remote Sensing **3**(3): 503-523.

JUUTILAINEN, K., P. HALME, H. KOTIRANTA & M. MÖNKKÖNEN (2011). "Size matters in studies of dead wood and wood-inhabiting fungi." Fungal Ecology **4**(5): 342-349.

KIM, Y., Z. YANG, W. B. COHEN, D. PFLUGMACHER, C. L. LAUVER & J. L. VANKAT (2009). "Distinguishing between live and dead standing tree biomass on the North Rim of Grand Canyon National Park, USA using small-footprint lidar data." Remote Sensing of Environment **113**(11): 2499-2510.

KRAUS, K. & N. PFEIFER (1998). "Determination of terrain models in wooded areas with airborne laser scanner data." ISPRS Journal of Photogrammetry & Remote Sensing **53**(4): 193-203.

LASSAUCE, A., Y. PAILLET, H. JACTEL & C. BOUGET (2011). "Deadwood as a surrogate for forest biodiversity: meta-analysis of correlations between deadwood volume and species richness of saproxylic organisms." Ecological Indicators **11**(5): 1027-1039.

LIN, Y. C. & J. P. MILLS (2009). "Integration of full-waveform information into the airborne laser scanning filtering process." International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **36**(3/W8): 224-229.

LOHR, S. M., S. A. GAUTHREAUX & J. C. KILGO (2002). "Importance of coarse woody debris to avian communities in loblolly pine forests." Conservation Biology **16**(3): 767-777.

PAILLET, Y., L. BERGÈS, J. HJÄLTEN, P. ÓDOR, C. AVON, M. BERNHARDT-RÖMERMANN, R. J. A. N. BIJLSMA, L. U. C. DE BRUYN, M. FUHR & U. L. F. GRANDIN (2010). "Biodiversity Differences between Managed and Unmanaged Forests: Meta-Analysis of Species Richness in Europe." Conservation Biology **24**(1): 101-112.

PESONEN, A., M. MALTAMO, K. EERIKÄINEN & P. PACKALÉN (2008). "Airborne laser scanning-based prediction of coarse woody debris volumes in a conservation area." Forest Ecology and Management **255**(8): 3288-3296.

RIEFELL, S., J. VERSCHUYL, D. MILLER & T. B. WIGLEY (2011). "Biofuel harvests, coarse woody debris, and biodiversity–A meta-analysis." Forest Ecology and Management **261**(4): 878-887.

RUTZINGER, M., B. HÖFLE, M. HOLLAUS & N. PFEIFER (2008). "Object-Based Point Cloud Analysis of Full-Waveform Airborne Laser Scanning Data for Urban Vegetation Classification." Sensors **8**: 4505-4528.

SALLABANKS, R., E. B. ARNETT, T. B. WIGLEY & L. IRWIN (2001). Accomodating Birds in Managed Forests of North America: A Review of Bird-forestry Relationships, National Council for Air and Stream Improvement.

VEHMAS, M., P. PACKALÉN, M. MALTAMO & K. EERIKÄINEN (2011). "Using airborne laser scanning data for detecting canopy gaps and their understory type in mature boreal forest." Annals of Forest Science **68**(4): 825-835.

WAGNER, W. (2005). "Physical principles of airborne laser scanning." University Course: Airborne laser scanning.

WAGNER, W., A. ULLRICH, V. DUCIC, T. MELZER & N. STUDNICKA (2006). "Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne

laser scanner." ISPRS Journal of Photogrammetry & Remote Sensing **60**(2): 100-112.

WOODALL, C. W., L. S. HEATH & J. E. SMITH (2008). "National inventories of down and dead woody material forest carbon stocks in the United States: Challenges and opportunities." Forest Ecology and Management **256**(3): 221-228.

Video-like exploitation of MiSAR sequences from small UAVs

G. Saur[1], N. Heinze[1] & W. Krüger[1]

Der miniaturisierte Radarsensor mit synthetischer Apertur MiSAR wurde von CASSIDIAN Deutschland für die Nutzung in Kleindrohnen entwickelt. In einem speziellen bildgebenden Verarbeitungsmodus werden SAR-Bildfolgen mit etwa 1 Hertz Bildtakt und mehr als 90% Überlappung erzeugt. In einer früheren Arbeit wurde eine Video-Verarbeitungskette vorgestellt, welche Relativtransformationen zwischen Bildern berechnet und zu Mosaiken zusammensetzt (SAUR et al. 2007). Diesen Experimenten lagen experimentelle MiSAR-Sequenzen zu Grunde, welche von einem Flugzeug mit relativ ruhigen Flug- und Antennenbewegungen gewonnen wurden. Neue MiSAR-Aufnahmen mit der Drohne LUNA als Sensorträger boten die Möglichkeit, die Auswerteverfahren unter realistischen Aufnahmebedingungen und mit für kleine UAVs typischen Flugbewegungen zu untersuchen. Kernkomponente ist die Ko-Registrierung aufeinanderfolgender Einzelbilder basierend auf der Zuordnung lokaler Bildmerkmale. Die darauf aufbauenden Verfahren zur Bildmosaikierung und Hintergrundstabilisierung werden für verschiedene Bewegungsmuster des Sensorabdrucks am Boden angewendet. Auf den erzeugten Bildprodukten vom Typ Mosaikstreifen und Bildstapel finden darüber hinaus Verfahren zur Referenzierung und Auflösungsverbesserung Anwendung, und es ergeben sich Perspektiven für Änderungs- und Bewegungsdetektion.

The miniature SAR system MiSAR has been developed by CASSIDIAN Germany for light weight UAVs (Unmanned Aerial Vehicle) like the LUNA system. In a special imaging mode MiSAR generates SAR image sequences with approximately 1 Hz framerate and an overlap area of more than 90% for successive frames. Starting from our previous work of SAUR et al. 2007 on MiSAR data taken from an aircraft carrying MiSAR and having relatively smooth flight and antenna movements, we now consider MiSAR taken from the UAV LUNA with agile flight characteristics typical for small UAVs. The approach combines image based co-registration of consecutive frames using local features and the a-priori geocoding information of the frames. The registration enables both image stitching to mosaics and creating co-registered images stacks. The resulting geocoded MiSAR mosaics with enhancements by blending and super-resolution give outlook to fusion and change detection.

1 Introduction

In the past years, Synthetic Aperture Radar (SAR) sensors have found a wide application field for remote sensing purposes. SAR images from satellites (TerraSAR-X/TanDEM-X, RadarSAT, Cosmo-Skymed etc.) are available, commercially and for research applications. On the other hand, the civil application of airborne SAR images until now is limited to few cases, due to the availability and the cost of sensors and measurement campaigns. But some technological trends developed for military reconnaissance may be interesting for future civil applications, too.
Reconnaissance with small UAVs (Unmanned Aerial Vehicles) is done until now mainly with optical video and IR-video sensors. In situations with mist and low clouds reconnaissance is prevented due to the missing of the all-weather capability of these sensors. This may happen in

[1] Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Fraunhoferstr 1, 76131 Karlsruhe, Germany, correspondence: guenter.saur@iosb.fraunhofer.de

situations where situation-awareness is essential. Especially in tropical regions such conditions appear regularly. This gap in reconnaissance-coverage can be closed by a SAR-sensor. Until now, size, power consumption and weight of SAR sensors prohibit their installation on small UAVs. To overcome this limitation, CASSIDIAN Germany (an EADS company) has developed the MiSAR sensor with a weight of only 4 kg. Due to the installation on a small UAV, there are challenges to overcome, to create high quality images in real-time.
In a previous work, we reported on adaption and application of video-like image processing to MiSAR images sequences, see SAUR et al. 2007. Another work is presented by MAGNARD et al. 2010 who used the range-Doppler equeations and the flight parameters for mosaicking and geocoding of MiSAR images. In our case, MiSAR data are used which are taken from an aircraft having relatively smooth flight and antenna movements. In the meanwhile some experimental flights with the UAV LUNA have been undertaken. These data were acquired under flight conditions more typical for small UAVs with a very flexible flight path and movements. The video processing steps image matching, mosaicking etc. have to be adapted to these data. Additional measures have been taken to get high-quality image products e.g. by super-resolution. Furthermore the generation of real-time imagery is only one step towards the objective to extract information from the flight-mission. Next step was to develop processing for this imagery and to develop an image exploitation station to give photo interpreters (PI) best working conditions for their image exploitation task.

2 MiSAR-Sensor

The overall goal was to design a miniaturized SAR sensor which can be integrated even in small UAVs currently operated on a regular basis. The image resolution and swath should be comparable to the respective capabilities of electro optical sensors used in small UAVs. To transmit the SAR data to the UAV ground control station, the standard analogue video data links used for the E/O sensors should be usable. With these preconditions, the design goals were determined as shown in the table, Figure 1 left. Using a FMCW radar transmitting in the Ka-band led to a product shown in Figure 1 right, see WEGIN & EDRICH 2003 and SAUR et al. 2007. With an antenna depression angle in (18°,30°) the resulting maximum radar range reaches up to 4 km.

Carrier platforms	small size UAVs
Image resolution	~ 0.5m x 0.5m
Swath	500m ... 1000m
max. onboard weight	4kg
max. onboard volume	10 liter
max. onboard power demand	100W
data link type	analogue
data link bandwidth	5MHz



Figure 1: Left Table: Design goals for the MiSAR sensor system. Right: UAV LUNA with MiSAR

3 Experimental Data and Preprocessing

Our previous work based on experimental data taken from an aircraft carrying MiSAR, see SAUR et al. 2007. The actual data reported here have been taken at a flight campaign with LUNA, see Figure 2 left. Typical for flying UAVs is the highly flexible flight path, and the additional movements of the sensor due to its exposure to wind, see the sensor footprints in Figure 2 right.



Figure 2: Left: UAV LUNA starting from catapult. Right: Flight path of LUNA (red line), footprints of image frames (yellow), and single MiSAR image frame overlaid over map and TerraSAR-X image.

The images delivered from the SAR processor have the following characteristics:

- Coverage of about 500 (azimuth) x 800 (range) meters. The area is a trapezoid with less width in near range than in far range. The angle of the mid axes depends on the squint angle.
- At a speed of up to 30 meter/second and a frame rate of 1 Hz, consecutive images will have an overlap of approximately 90 %.
- The MiSAR image frames have a gray value range that is coded by 16bit.
- For each image frame, the Meta data carry information about the SAR parameters, the illumination coverage by the antenna, and the flight path (see Figure 2 right).
- The geocoding is specified by the coordinate triples of the four image corners.

Thus, the image sequences have the characteristics of a “SAR video” of 1 Hz and so we applied some video-like processing to these data.

Image pre-processing is done in an initial step for displaying purposes and before applying further algorithms (e.g. mosaicing). This includes some filter and a gray value scaling, so that consecutive MiSAR images have the same gray value characteristics. This is essential, since due to the different variations e.g. in the scene, the aspect angle, the squint angle etc. the images must have a uniform gray scaling, so that the mosaicking result has not too extreme gray steps.

4 Methods

The main goals for MiSAR image processing are geocoding, mosaicking, and image enhancement. MAGNARD et al. 2010 presented a “direct” approach based on the MiSAR raw data using the flight parameters and the range-Doppler equations. Our approach is using also the

flight parameters by considering the geocoding of the MiSAR frames by its corner coordinates. Additionally, this geocoding information is combined with the image-to-image-transformation of consecutive frames. The first one gives a long-term and the second one a short-term stabilization. Using this basic idea, the following methods can be applied.

4.1 Image Coregistration

To find local features for image registration, we apply the Förstner-Harris corner detector (HARRIS & STEPHENS 1988) which detects image locations where the image function has significant changes in both directions. Robust parameter estimation is used to find a transformation from the generally outlier-contaminated set of putative point correspondences (MÜLLER et al. 2006). To cope with large image motions, the whole estimation process is iterated at the different resolutions of an image pyramid.

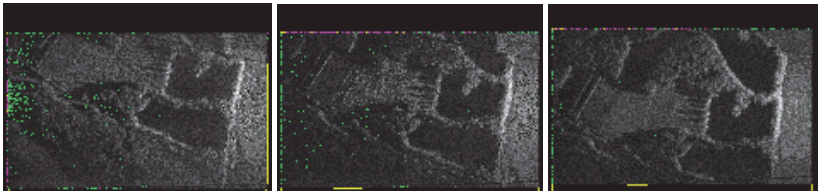


Figure 3: Sequence of three successive MiSAR images with trapezoidal coverage (yellow lines) and cutout frame (magenta box).

Additionally, an alternative image-to-image transformation is given by combining the geocoding information. In Figure 3 three frames are shown with linearly moving footprint in flight direction. In Figure 2 right we see that the footprint could also move backwards or even stay in place. In presence of wind, the speed of the UAV relative to the ground is varying, and the synthetic aperture image is not well defined if the speed is, e.g., close to zero. For this reason and due to some other influences, the algorithm has to cope with varying image quality in the MiSAR sequence.

4.2 Mosaicking

Mosaicking uses the coregistration information to stitch the image frames to a single image carpet which covers the scene stripe swept by the sensor's footprint. Filtering the coregistration chain and using the frame geocoding as (flexible) constraints leads to an elastic transformation chain and results a MiSAR geomosaic, see Figure 4 left. Even in the (partial) presence of water with no local features, the geomosaic fits to the map. The registration can be verified by using another georeferenced image of the scene. Figure 4 right shows a TerraSAR-X image with an appearance similar to that of MiSAR but of lesser resolution. Application of a line-based image matching (MÜLLER et al. 2006) verifies the geocoding and is capable to compensate global registration errors.

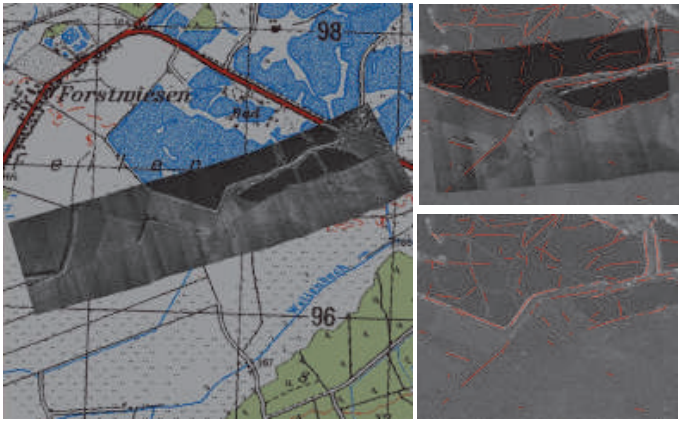


Figure 4: Left: Mosaicking and geocoding result MiSAR image overlaid over map. Right: fine-registration by line segments using MiSAR mosaic (top) and TerraSAR-X image (bottom).

Beside the correct image transformations, the image quality is of high importance. One measure for ensuring this consists in selecting high-quality image parts with sufficient radar illumination by the antenna and in suppressing low-quality images. Combining this selection with image blending at the stitch edges of the mosaic, high-quality MiSAR mosaics will result. These are suited for further image evaluation, e.g. by a human photo interpreter.

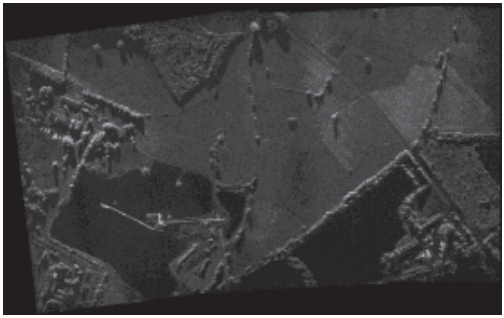


Figure 5: MiSAR mosaicking result using image blending.

4.3 Super-Resolution

Resolution is always a crucial part of video-systems and of SAR, of course. In image sequences, many low resolution frames are available. By combining the information of a co-registered image stack, the noise can be reduced and the resolution can be improved. Our implementation combines the robust image registration with the super-resolution approach from FARSIU et al. 2004. Figure 6 top shows the principle where several frames are registered on a subpixel high

resolution grid and are projected to the fused result image. Figure 6 bottom shows a MiSAR super-resolution result which has been fused from 9 single MiSAR images. Since the overlapping area leaves only a narrow strip, this method has to be combined with the mosaicking leading to extended super-resolution mosaics. The example shows that the speckle has been reduced, that the strong reflexes have reduced brightness with sharper edges, that the object contours have become clearer, and that fine structures and gray differences got visible.

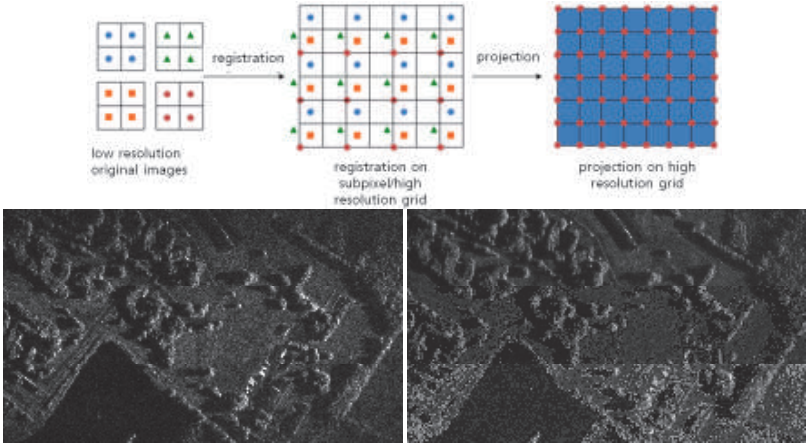


Figure 6: Top: Principle of super-resolution approach from Farsiu et al. 2004. Bottom: Application to MiSAR with original (left) and result (right) fused from 9 frames.

5 System Integration and Experimental Application

Parts of the described MiSAR algorithms have been integrated into the LUNA ground station (HEINZE et al. 2010). There, additional algorithms are available, like image filtering and enhancement, screening, object detection, and tracking. The operator can choose between different algorithms, in particular the mosaicing, screening, grey-value manipulation etc. as described above. Furthermore, navigation in the acquired imagery and annotation of images is possible. The system is designed for on-line real time and off-line exploitation.
The ground segment is entirely made up of commercial PC components. For the image formation part the intrinsic SAR processor unit is a workstation which is linked by a gigabit ethernet switch to two peripheral workstations for data preprocessing and data storage. The exploitation station part are two PCs used for Human-Machine interface, visualization and running of automatic image processing and exploitation algorithms, see Figure 7.

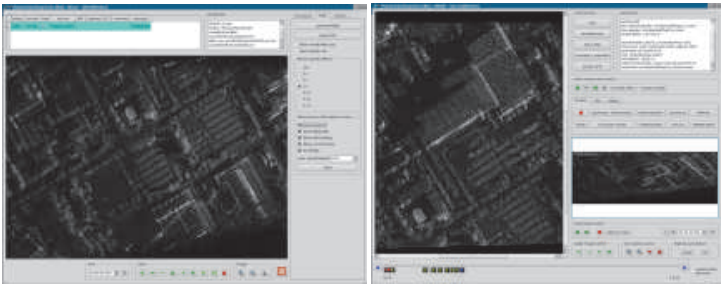


Figure 7: Dual screen user interface for MiSAR image exploitation. Right: Display of image frames and processing in the sensor's real time. Left: Mosaic and replay function with timeline.

6 Summery

The lightweight MiSAR sensor from CASSIDIAN takes high quality SAR image sequences, filling the actual gap in reconnaissance coverage during mist and low cloud conditions, considering light weight UAVs. To cope with the difficulties of geocoding and mosaicking SAR sequences of agile flight paths under the presence of noisy images, a combination of direct geocoding and image based coregistration extents our former approach. On this basis, MiSAR mosaicking and geocoding is possible. Blending and super-resolution approaches produce high-quality geomosaic image which are suited for further exploitation by human photo interpreters (PIs). Further image processing investigations are related to fusion with other sensor images and to change detection.

7 References

FARSIU, S., ROBINSON, M. D., ELAD, M., MILANFAR, P., 2004: Fast and Robust Multiframe Super Resolution, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 10, Oct. 2004, pp. 1327-1344

HARRIS, C. & STEPHENS, M., 1988: A combined corner and edge detector. Proc. 4th Alvey Vision Conference, University of Manchester, 31st August - 2nd September 1988, 147-151

HEINZE, N., ESSWEIN, M., KRÜGER, W., SAUR, G., 2010: Image exploitation algorithms for reconnaissance and surveillance with UAV, Proc. SPIE Vol. 7668, Orlando FL, April 2010

MAGNARD, C., FRIOUD, M., MEIER, E., 2010: Processing, geocoding and mosaicking of MiSAR data. Dreiländertagung DGPF, SGPBF, OVG, Vienna (A), July 1-3 2010, pp. 524-531

MÜLLER, M., KRÜGER, W., SAUR, G., 2006: Robust image registration for fusion, *Information Fusion* 8 (4): 347-353 (2007)

SAUR, G., EDRICH, E. HEINZE, N., KRÜGER, W., 2007: Video-like MiSAR image sequence processing, DGPF Tagungsband 16 / 2007, Dreiländertagung SGPBF, DGPF und OVG, pp. 111-118, 2007

WERGIN, A. & EDRICH, M., 2003: First LUNA Flight Results with a Miniaturised SAR, in Proc. of the EURO UVS International Technical Conference on Unmanned Vehicle System Technologies, Brussels, 2003

TanDEM-X als Datengrundlage für Referenzkartierungen im Kontext von REDD

MICHAEL SCHLUND[1], FELICITAS VON PONCET[1], STEFFEN KUNTZ[1], DIRK H. HOEKMAN[2] & CHRISTIANE SCHMULLIUS[3]

Zusammenfassung: Entwaldung und Walddegradation sind Quellen für anthropogene Kohlenstoffemissionen. Daher ist die Überwachung der tropischen Wälder von großer Bedeutung, da sie eine wichtige Rolle als Kohlenstoffsenke einnehmen. Ein Werkzeug hierfür könnten SAR Sensoren sein. Die bistatische TanDEM-X Mission besteht aus zwei Sensoren, die im Laufe ihrer Mission die Erde mehrmals aufnehmen. Daher könnten diese Datensätze als Grundlage für die Überwachung tropischer Wälder dienen. Die Ergebnisse implizieren, dass die Kohärenz von der Kronenbedeckung abhängig ist und Informationen über die Vegetationsstruktur liefert. Eine Merkmalsselektion und ein Vergleich von Klassifikationen bestätigen, dass die Kohärenz zur Trennbarkeit von thematischen Klassen beiträgt und die Genauigkeit um 10 % im Vergleich zur Klassifikation ohne Kohärenz erhöht.

1 Einleitung

Das Wissen über Kohlenstoffspeicher und deren Veränderung durch Landnutzungsänderungen ist im Zusammenhang des Klimawandels von Bedeutung (WATSON et al. 2000). Vor allem die tropischen Wälder spielen dabei eine große Rolle, die mithilfe von synthetischen Aperturradarsystemen (SAR) aufgrund ihrer Wetterunabhängigkeit konsistent überwacht werden können (KUNTZ 2010). Im Laufe der TanDEM-X Mission wird die Erde mehrmals von zwei SAR Sensoren aufgenommen, um die Erstellung eines genauen globalen digitalen Höhenmodells zu gewährleisten (KRIEGER et al. 2007). Somit können diese Daten als homogene und zeitlich konsistente Grundlage für Landbedeckungskartierungen in hoher Auflösung genutzt werden.
Eine automatische Kartierung auf Basis von SAR Daten ist anspruchsvoll (DEL FRATE et al. 2008). Die TanDEM-X Konstellation könnte es möglich machen, eine Kartierung mit hoher Genauigkeit zu erzielen. Beide Sensoren der TanDEM-X Mission fliegen in räumlich geringem Abstand zueinander und bilden ein bistatisches SAR System bzw. einen SAR Interferometer im All, wobei ein Sensor als Sender und Empfänger (aktiv/monostatisch), hingegen der andere lediglich als Empfänger dienen kann (passiv/bistatisch). Beide empfangen das Signal des gleichen Punktes auf der Erde mit einem geringfügigen Unterschied des Einfallswinkels (KRIEGER et al. 2007). Dieser Winkelunterschied ermöglicht, dass Ziele erfasst werden, die mit einem Sensor nicht erfasst werden können und das sich das empfangene Signal abhängig vom Winkelunterschied und den Streueigenschaften der Objekte verändert. Eine verbesserte Klassifikation der Daten könnte daher erwartet werden (KRIEGER & MOREIRA 2006).

1) Michael Schlund, Felicitas von Poncet, Steffen Kuntz, Astrium GEO-Information Services, Claude-Dornier-Str., 88090 Immenstaad, michael.schlund@astrium.eads.net, felicitas.poncet@astirum.eads.net, steffen.kuntz@astrium.eads.net
2) Dirk H. Hoekman, Wageningen University, Droevendaalsesteeg 4, 6708 PB, Wageningen, Niederlande, dirk.hoekman@wur.nl
3) Christiane Schmullius, Friedrich-Schiller-Universität Jena, Löbdergraben 17, 07743 Jena; c.schmullius@uni-jena.de

Größter Vorteil der TanDEM-X Mission im Vergleich zu anderen SAR Missionen ist die Reduktion der temporalen Dekorrelation bei der interferometrischen Auswertung (KRIEGER et al. 2007). Bereits u.a. WEGMÜLLER & WERNER (1995) konnten einen Informationsgewinn bei ERS Daten durch die Verwendung der interferometrischen Kohärenz für Landbedeckungsklassifikationen feststellen. Bei der TanDEM-X Mission ist im Gegensatz zu bekannten Studien mit *repeat-pass* interferometrischen Daten (bspw. WEGMÜLLER & WERNER 1995, 1997; ENGDAHL & HYYPPA 2003) die zeitliche Dekorrelation vernachlässigbar und die interferometrische Kohärenz hauptsächlich durch Volumendekorrelation beeinflusst. Andere Einflüsse wie die Bewegung von Objekten oder Veränderung des Feuchtegehaltes spielen keine Rolle und Konfusionen könnten somit vermieden werden.
Ziel der Studie ist die effektive Nutzung der bistatischen TanDEM-X Daten zur Erstellung einer automatisierten Landbedeckungsklassifikation. Des Weiteren soll der Informationsgehalt der speziellen bistatischen Merkmale im Vergleich mit monostatischen Merkmalen untersucht werden.

2 Untersuchungsgebiet & Datengrundlage

Das Untersuchungsgebiet befindet sich auf Kalimantan, Indonesien und ist hauptsächlich durch Torfsumpfwälder bedeckt. Torfsümpfe machen einen geringen Teil der globalen Landbedeckung aus, jedoch haben sie eine große Bedeutung in der globalen Kohlenstoffbilanz (VAN DER WERF et al. 2009). Im El Niño Jahr 1997-1998 wurden nach PAGE et al. (2002) 0,81 – 2,57 Gt Kohlenstoff durch die Torfsümpfe Kalimantans emittiert, was 13 – 40 % der mittleren jährlichen globalen Emission von Kohlenstoff entspricht. Die Torfsumpfwälder Kalimantans sind jedoch weiterhin gefährdet, da sie für landwirtschaftliche Zwecke und Plantagen drainiert werden (HOOIJER et al. 2010). Die Abbildung 1 zeigt die Lage des Untersuchungsgebietes.

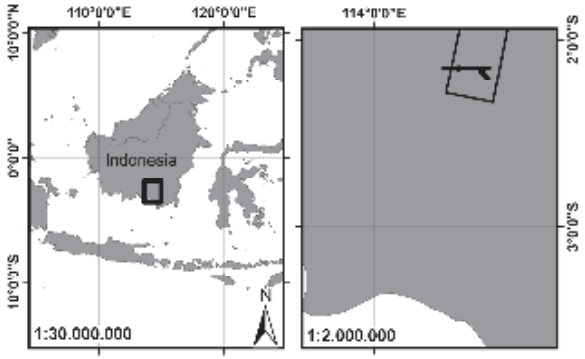


Abb. 1: Lage des Untersuchungsgebietes auf Kalimantan (links) & Lage der TanDEM-X & LiDAR Aufnahmen sowie der Felderhebungen (rechts)

Weiterhin ist die Ausdehnung des verwendeten SAR Datensatzes der TanDEM-X Mission zu erkennen. Die SAR Daten wurden am 21.12.2010 aufgenommen. Sie wurden im bistatischen Modus im StripMap Format mit HH (horizontal) Polarisation aufgenommen, der für die operationelle Mission verwendet wird (KRIEGER et al. 2007). Die Auflösung der Daten beträgt ca. 3 m, die Wellenlänge 3,1 cm und die effektive Basislinie beider Sensoren 163 m. Die Daten wurden koregistriert, *resampled* und als sogenannte CoSSCs geliefert (Co-registered Single-look Slant range Complex; FRITZ 2012).
Aufgrund der schwierigen Erreichbarkeit des Gebietes standen für die Analyse lediglich 8 Transekte, in denen Feldmessungen erhoben wurden, mit einer Größe von 100 m x 10 m zur Verfügung. In diesen Gebieten wurden die Baumhöhe, die Höhe bis zum ersten Ast, der Stammdurchmesser auf Brusthöhe sowie die Kronenausdehnung gemessen. Zusätzlich wurden Luftbilder mit einem Ultraleichtflugzeug gemacht. Diese wurden am 13. bis 15.01.2005 in Nadir aufgenommen und deren Position sowie die Höhe des Flugzeuges per GPS gemessen. Insgesamt konnten 343 Aufnahmen verwendet werden. Aktuellere Aufnahmen wurden ferner am 04.12.2011 gemacht, als das Gebiet mit einem Flugzeug in geringer Flughöhe ein weiteres Mal beflogen wurde.
Um weitere quantitative Analysen außerhalb der 8 Felderhebungen machen zu können, stand ein LiDAR Datensatz mit einem Meter Auflösung vom 05.08.2007 zu Verfügung. Hierbei konnte sowohl das digitale Oberflächenmodell als auch das digitale Geländemodell verwendet werden. Die Differenz beider ergibt das *canopy height model* (CHM), welches in bewaldeten Gebieten die Baumhöhe wiedergibt. In Abbildung 1 ist die Ausdehnung des verwendeten LiDAR Datensatzes sowie die Lokation der Felddaten zu erkennen. Offensichtliche Änderungen zwischen den Aufnahmen der Referenz- sowie der TanDEM-X Daten konnten mithilfe von ALOS PALSAR Zeitreihen und SPOT 5 Daten ausgeschlossen werden.

3 Methodik

3.1 Vorprozessierung & Klassendefinition

Die aktive und passive Amplitude wurden vom TanDEM-X CoSCC abgeleitet. Die Differenz beider Rückstreukoeffizienten soll durch die Berechnung des Ratios der Intensitäten (=Amplituden²) hervorgehoben werden. Weiterhin wurde die interferometrische Phase der beiden komplexen Datensätze berechnet und die Kohärenz abgeleitet. Vorherige Studien belegen, dass Texturinformationen bei hohen Frequenzen hilfreicher für Waldklassifikationen als die Rückstreuintensitäten sind (bspw. VAN DER SANDEN & HOEKMAN 1999). Um Texturinformationen abzuleiten, wurden die *co-occurence* Maße von HARALICK et al. (1973) sowie der Variaktionskoeffizient berechnet. Die Daten wurden anschließend geokodiert und auf 5 Meter *resampled*.
Zur Vorbereitung auf die Klassifikation wurde ein Interpretationsschlüssel mithilfe der Felderhebungen und den Luftbildern erstellt. Demnach wurden folgende Klassen definiert: „Wald mit mehr als 50 % Kronenbedeckung“, „Wald mit weniger als 50 % Kronenbedeckung“, „Grünland“, „Buschvegetation“, „Feuchtgebiete“ und „Wasserflächen“. Diese sind konform mit den Methoden der IPCC (*Intergovernmental Panel on Climate Change*, 2003) für Landnutzung, Landnutzungsänderungen und Forst (*Land use, land use change and forestry*). Exakte

Definitionen für diese Klassen werden jedoch nicht gegeben. Daher wurde die Definition nach FAO (*Food and Agriculture Organization of the United Nations*, 2000) für die Waldklasse verwendet. Mithilfe der Feldmessungen konnte die Kronenbedeckung pro Meter Einheitsfläche bestimmt werden, die sich aus der Anzahl der Bäume pro Hektar sowie aus der durchschnittlichen Kronenfläche in Meter berechnet. Aufgrund der räumlichen Nähe einiger Feldmessungen wurde jeweils der Mittelwert berechnet, wodurch lediglich 4 Transekte zur Verfügung standen.
Mithilfe des *canopy height models* des LiDAR Datensatzes wurden einzelne Bäume detektiert. Hierfür wurde ein *local maximum* Filter verwendet. Ein lokales Maximum wird definiert, wenn alle Nachbarpixel einen geringeren Wert aufweisen. Die Anzahl der detektierten lokalen Maxima pro Einheitsfläche kann Aufschluss über die Anzahl der Bäume pro Einheitsfläche geben (POPESCU & WYNNE 2004).

3.2 Merkmalsselektion & Klassifikation

Um die Relevanz der bistatischen TanDEM-X Merkmale wie bistatischer Rückstreukoeffzient und Kohärenz im Vergleich zu monostatischen Merkmalen zu bestimmen, wurde eine Trennbarkeitsanalyse mit Merkmalsselektion durchgeführt. Aufgrund der hohen Auflösung der Daten wird ein objekt-basiertes Verfahren als geeigneter als ein pixel-basiertes Verfahren betrachtet. Die geokodierten Merkmale wurden mit der *eCognition* Software segmentiert und anschließend die Objektmittelwerte exportiert, um sie in R (R CORE TEAM 2012) statistisch analysieren zu können.
Für die Merkmalsselektion wurde der *random forest* Algorithmus nach BREIMAN (2001) verwendet, der eine Bewertung der Wichtigkeit der Merkmale beinhaltet. Diese kann mithilfe des *z-score* berechnet werden. Dieser Wert alleine betrachtet, kann allerdings zu Fehlinterpretationen führen. Die Korrelation einzelner Merkmale kann beispielsweise die Trennung von wichtigen und unwichtigen Merkmalen erschweren. Daher wurde der Boruta Algorithmus nach KURSA et al. (2010) angewendet, bei dem die Wichtigkeit der abgeleiteten Merkmale mit der Wichtigkeit von Zufallsmerkmalen verglichen wird. Erzielen die abgeleiteten Merkmale einen höheren *z-score* als die Zufallsmerkmale, so wird das jeweilige Merkmal als relevant betrachtet.
Die fünf wichtigsten Merkmale wurden anschließend genutzt, um eine Landbedeckungsklassifikation durchzuführen. Hierbei wurden zwei Merkmalsselektionen getrennt voneinander betrachtet. Zum einen wurden Merkmale verwendet, die ausschließlich aus der aktiven Amplitude abgeleitet wurden. Dies soll eine monostatische TerraSAR-X Aufnahme simulieren. Zum anderen wurden auch die bistatischen Merkmale in die Merkmalsselektion aufgenommen. Die monostatische Merkmalssauswahl besteht aus der aktiven Amplitude und deren abgeleitete Texturen Variationskoeffizient, Korrelation, Mittelwert und Entropie. Bei der bistatischen Merkmalsauswahl wurde die Entropie durch die interferometrische Kohärenz ersetzt. Für den Vergleich beider Merkmalskonstellationen wurde eine *maximum likelihood* Klassifikation verwendet, bei der die gleichen Trainingsgebiete genutzt wurden. Anschließend wurden beide Klassifikationen mithilfe der Luftbilder bewertet. Hierfür wurde eine stratifizierte zufällige Auswahl von Stichprobenpunkten verwendet. Trainingsgebiete wurden aus der

Bewertung ausgeschlossen. Auch hier wurden für beide Merkmalsselektion die gleichen Punkte beurteilt, um Vergleichbarkeit der Ergebnisse zu gewährleisten.

4 Ergebnisse & Diskussion

Die Tabelle 1 zeigt unterschiedliche Kronenbedeckungsgrade aus den Felderhebungen und den dazugehörigen Mittelwert und Standardabweichung der interferometrischen Kohärenz. Die Kohärenz ist bei höherer Bedeckung niedriger. Dies deutet darauf hin, dass die Kohärenz als Indikator für die Waldstruktur dienen kann. Diese Information kann für die Landbedeckungsklassifikation genutzt und somit die Trennbarkeit für Vegetationsklassen verbessert werden. Die Kohärenz ist hauptsächlich durch die Volumendekorrelation bestimmt. Volumendekorrelation tritt demnach stärker auf, je höher die Kronenbedeckung ist und die elektromagnetische Welle dadurch mit einem dichteren Volumen interagiert. Dagegen steigt die Wahrscheinlichkeit der direkten oder multiplen Rückstreuung durch Äste, Zweige und Stämme von Bäumen, wenn die Kronenbedeckung geringer ist und es somit zu weniger Volumendekorrelation kommt. Tabelle 1 zeigt aber auch, dass mithilfe der Kohärenz lediglich zwei Klassen getrennt werden können, da die Mittelwerte für 130 % und 80 % sowie 45 % und 30 % sich nicht signifikant unterscheiden. Aufgrund der Überlappung von Kronen kann der Wert des Kronenbedeckungsgrades 100 % überschreiten.

Tab. 1: Kronenbedeckungsgrad & dazugehörige Werte der Kohärenz

Kronenbedeckungsgrad	Kohärenz (Mittelwert)	Kohärenz (Standardabweichung)
130 %	0,787	0,099
80 %	0,791	0,088
45 %	0,840	0,062
30 %	0,834	0,064

Abbildung 2 zeigt die Korrelation der interferometrischen Kohärenz mit der Anzahl der detektierten lokalen Maxima pro Hektar. Die Anzahl der lokalen Maxima pro Hektar kann Aufschluss über die Anzahl der Bäume pro Hektar geben. Allerdings werden gegebenenfalls nicht alle Bäume erfasst. Bäume, die von der Krone eines anderen Baumes verdeckt werden, können nicht detektiert werden. Des Weiteren können kleinere Bäume, die in der Nachbarschaft von größeren Bäumen stehen, Auslassungsfehler erzeugen, da sie kein lokales Maximum ergeben. Abbildung 2 zeigt trotz dessen, dass es eine Korrelation zwischen der Kohärenz und der Anzahl der lokalen Maxima als Parameter zur Beschreibung der Waldstruktur gibt. Die Anzahl der Bäume ist ein Faktor in der Berechnung des Kronenbedeckungsgrades und es könnte geschlossen werden, dass die Kohärenz von der Anzahl der Bäume als auch vom Bedeckungsgrad, die miteinander verknüpft sind, beeinflusst wird.

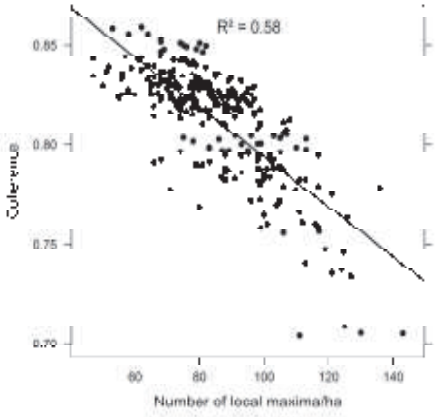


Abb. 2: Zusammenhang zwischen interferometrischer Kohärenz & Anzahl lokaler Maxima pro Hektar

Die Merkmalsselektion mit *random forest* zeigte, dass alle abgeleiteten Merkmale im Vergleich zu den zufälligen Merkmalen relevant sind. Die interferometrische Kohärenz stellte sich dabei mit Abstand als wichtigstes Merkmal heraus und erzielte einen *z-score* von 1,6. Des Weiteren konnte bestätigt werden, dass Texturmerkmale mehr zur Klassenseparabilität beitragen als Rückstreuintensitäten. Wichtigste Texturmerkmale sind der Variationskoeffizient mit einem *z-score* von 1,32 sowie die Haralick-Merkmale *correlation* (1,25), Mittelwert der aktiven Amplitude (1,22) und *entropy* (1,19). Die aktiven und passiven Rückstreukoeffizienten hingegen erzielten einen Wert von 1,11 bzw. 0,97. Das Ratio-Merkmal, das die Unterschiede beider Rückstreukoeffizienten hervorheben soll, erreichte einen Wert von 0,56 und ist damit das Merkmal mit der geringsten Relevanz. Diese Resultate könnten die Folgerung zulassen, dass der Winkelunterschied einer operationellen TanDEM-X Aufnahme demzufolge nicht ausreicht, um einen Informationsgewinn durch die Verwendung der passiven Amplitude bzw. des Ratio zu erzielen.

Die wichtigsten Merkmale wurden in den Klassifikationen verwendet, deren Ergebnisse in Abbildung 3 dargestellt sind. Links ist die Klassifikation mit ausschließlich monostatischen Merkmalen abgebildet. Diese bestehen aus aktiver Amplitude und deren Texturmaße Variationskoeffizient, *correlation*, Mittelwert und *entropy*. Rechts ist die Klassifikation mit mono- und bistatischen Merkmalen dargestellt, bei der die *entropy* durch die interferometrische Kohärenz ersetzt wurde. Es wird deutlich, dass die räumliche Verteilung der Landbedeckung bei beiden Klassifikationen ähnlich ist. Allerdings ist die monostatische Klassifikation verrauschter und enthält weniger Details. Beispielsweise werden degradierte Waldflächen nicht detektiert. Hingegen sind in der bistatischen Klassifikation diese Flächen als „Wald mit offener Bedeckung“ klassifiziert worden (bei ca. 2°10’S, 114°28’O).

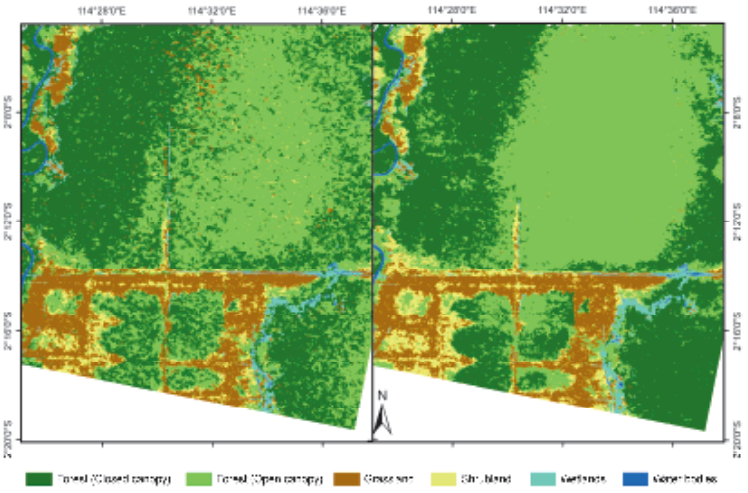


Abb. 3: Ausschnitt der Klassifikationen mit monostatischen (links) & bistatischen (rechts) Merkmalen

Die bistatische Klassifikation erzielte in der Genauigkeitsanalyse in jeder Klasse höhere Genauigkeiten als die monostatische. Speziell die Einschließungsfehler der beiden Waldklassen werden durch die interferometrische Kohärenz um 14 % bzw. 16 % vermindert. Darüber hinaus konnten Konfusionen in der Klasse „Buschvegetation“ durch die Verwendung der Kohärenz verringert werden, wodurch sich die *producer's accuracy* dieser Klasse von 59 % auf 94 % erhöht. Mit einer Gesamtgenauigkeit von 85 % ist die bistatische Klassifikation um ~10 % höher als die monostatische Klassifikation.

5 Zusammenfassung

Die interferometrische Kohärenz mit *repeat-pass* interferometrischen X-Band Daten ist meist aufgrund der zeitlichen Dekorrelation bei Vegetation gering, so dass lediglich Vegetation und Nicht-Vegetation in besten Fällen unterschieden werden kann. Die *single-pass* interferometrischen Daten der TanDEM-X Mission bieten jedoch die Möglichkeit die Kohärenz auch für die Unterscheidung von Vegetationsklassen zu nutzen. Die Auswertung der Felderhebungen sowie der LiDAR Merkmale zeigte, dass die Kohärenz unter anderem von der Anzahl der Bäume bzw. dem Bedeckungsgrad abhängig ist. Die Kohärenz könnte demzufolge genutzt werden, um unterschiedliche Waldbedeckungen voneinander zu trennen. Dies bestätigte sowohl die Merkmalsselektion als auch die Auswertung der Klassifikationsergebnisse. Hierbei

erzielte die Kohärenz den höchsten Wert der Relevanz und erhöhte die Klassifikationsgenauigkeit um ca. 10 %. Die global verfügbaren TanDEM-X Datensätze könnten somit als potentielle Grundlage für die Referenzkartierung in REDD Projekten dienen. Dies soll in nachfolgenden Untersuchungen bestätigt werden.

Danksagung

Das Projekt wird gefördert von Astrium GEO-Information Services. Die TanDEM-X Daten wurden vom Deutschen Zentrum für Luft- und Raumfahrt (DLR) bereitgestellt. Ein Dank gilt der JAXA für die Bereitstellung der ALOS PALSAR Zeitreihen. Ein weiterer Dank gilt BOS für die Unterstützung vor Ort in Kalimantan sowie H.-D. V. Boehm von Kalteng Consultants für die Lieferung der LiDAR Daten.

6 Literaturverzeichnis

BREIMAN, L., 2001: Random Forests. Machine Learning, **45**, S. 5–32.

DEL FRATE, F., PACIFI, F. & SOLIMINI, D., 2008: Monitoring Urban Land Cover in Rome, Italy, and Its Changes by Single-Polarization Multitemporal SAR Images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, **2**, S. 87-97.

ENGDAHL, M. E. & HYYPPA, J. M., 2003: Land-cover classification using multitemporal ERS-1/2 InSAR data. IEEE Transactions on Geoscience and Remote Sensing, **41**, S. 1620-1628.

FAO, 2000: FRA 2000. On definitions of forest and forest change. Forest Resources Assessment Programme Working Paper 33: Rome.

FRITZ, T., 2012: TanDEM-X. Ground Segment. TanDEM-X Experimental Product Description. Issue: 1.2. Oberpfaffenhofen: Deutsches Zentrum für Luft- und Raumfahrt (DLR).

HARALICK, R. M., SHANMAGUN, K. & DINSTEIN, I., 1973: Textural Features for Image Classification. IEEE Transactions on Systems, Man & Cybernetics, **6**, S. 610-621.

HOOIJER, A., PAGE, S. E., CANADELL, J. G., SILVIUS, M., KWADIJK, J., WÖSTEN, H. & JAUHIAINEN, J., 2010: Current and future CO2 emissions from drained peatlands in Southeast Asia. Biogeosciences, **7**, S. 1505–1514.

IPCC, 2003: Good practice guidance for land use, land-use change and forestry. Hayama: Institute for Global Environmental Strategies (IGES).

KRIEGER, G. & MOREIRA, A., 2006: Spaceborne bi- and multistatic SAR: potential and challenges. IEEE Proceedings - Radar, Sonar and Navigation, **153**, S. 184-198.

KRIEGER, G., MOREIRA, A., FIEDLER, H., HAJNSEK, I., WERNER, M., YOUNIS, M., & ZINK, M., 2007: TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry. IEEE Transactions on Geoscience and Remote Sensing, **45**, S. 3317-3341.

KUNTZ, S., 2010: Potential of spaceborne SAR for monitoring the tropical environments. Tropical Ecology, **51**, S. 3-10.

KURSA, M. B., JANKOWSKI, A. & RUDNICKI, W. R., 2010: Boruta – A System for Feature Selection. Fundamenta Informaticae, **101**, S. 271-285.

PAGE, S. E., SIEGERT, F., RIELEY, J. O., BOEHM, H.-D. V., JAYAK, A. & LIMINK, S., 2002: The amount of carbon released from peat and forest fires in Indonesia during 1997. Nature, **470**, S. 61-65.

POPESCU, S. C. & WYNNE, R. H., 2004: Seeing the Trees in the Forest: Using Lidar and Multispectral Data Fusion with Local Filtering and Variable Window Size for Estimating Tree Height. Photogrammetric Engineering & Remote Sensing, **70** (5), S. 589-604.

R CORE TEAM, 2012: R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

VAN DER SANDEN, J. J. & HOEKMAN, D. H., 1999: Potential of Airborne Radar To Support the Assessment of Land Cover in a Tropical Rain Forest Environment. Remote Sensing of Environment, **68**, S. 26-40.

VAN DER WERF, G. R., MORTON, D. C., DEFRIES, R. S., OLIVIER, J. G. J., KASIBHATLA, P. S., JACKSON, R. B., COLLATZ, G. J. & RANDERSON, J. T., 2009: CO_2 emissions from forest loss. Nature, **2**, S. 737-738.

WATSON, R. T, NOBLE, I. R., BOLIN, B., RAVINDRANATH, N. H., VERARDO, D. J. & DOKKEN, D. J., 2000: Land Use, Land Use Change and Forestry. Cambridge Univ. Press: Cambridge.

WEGMÜLLER, U. & WERNER, C., 1997: Retrieval of vegetation parameters with SAR interferometry. IEEE Transactions on Geoscience and Remote Sensing, **35**, S. 18-24.

WEGMÜLLER, U. & WERNER, C., 1995: SAR interferometric signatures of forest. IEEE Transactions on Geoscience and Remote Sensing, **33**, S. 1153-1161.

Potential of Airborne Single-Pass Millimeterwave InSAR Data for Individual Tree Recognition

MICHAEL SCHMITT [1], AXEL BRÜCK [1], JOHANNES SCHÖNBERGER [1] & UWE STILLA [1]

Abstract: This paper shows the potential of airborne single-pass millimeterwave InSAR data for the recognition of individual trees. In addition to the very high available resolution in the decimeter range, the benefits of both single-pass interferometry as well as millimeterwave radar are discussed. Using knowledge of the side-looking imaging geometry, the representation of trees in airborne SAR data is modeled and compared to optical and LiDAR remote sensing data. Finally, two different concepts for automated single tree recognition are proposed in order to show the versatile potential of synthetic aperture radar interferometry for the required task.

Zusammenfassung: Dieser Aufsatz zeigt das Potential flugzeuggetragener Single-Pass Millimeterwellen InSAR-Daten für die Erkennung einzelner Bäume. Neben der sehr hohen verfügbaren Auflösung im Dezimeterbereich werden die Vorteile sowohl der Single-Pass-Interferometrie als auch des Millimeterwellen-Radar diskutiert. Unter Nutzung von Wissen über die seitwärts-blickende Abbildungsgeometrie wird die Erscheinung von Bäumen in flugzeuggetragenen SAR-Daten modelliert und mit optischen und LiDAR-Fernerkundungsdaten verglichen. Schließlich werden zwei unterschiedliche Ansätze zur automatisierten Einzelbaumerkennung vorgestellt um das vielseitige Potential von Radar-Interferometrie für die geforderte Aufgabe zu zeigen.

1 Motivation for Tree Recognition in InSAR Imagery

The automatic recognition of single trees in remote sensing data is an important topic for several application fields. The most important certainly is sustainable forest management: In many countries, single-tree related parameters are used as a basis for forest inventory, e.g. tree species, mean tree height or timber volume. To this date, most of these variables are collected manually by measurement of sample plots in cost- and time-intensive field surveys, although remote sensing-based methods have been investigated for some decades now (FAGAN & DEFRIES, 2009). Another potential application is the extraction of single trees in urban areas in order to generate information for city tree cadastres or to add tree layers to geoinformation systems and 3D city models (STRAUB & HEIPKE, 2001).

While most of the research concerning the utilization of synthetic aperture radar (SAR) for the remote sensing of trees and forests centers either around large-scale forest classification (PERKO et al., 2010), biomass and forest volume estimation using L band tomography (NEUMANN et al., 2010; MERCER et al., 2010), or canopy height model reconstruction using X band interferometry (IZZAWATI et al., 2006), most work on the individual tree level to this date is based

1) Photogrammetrie & Fernerkundung, Technische Universität München (TUM), Arcisstr. 21, 80333 München, http://www.pf.bv.tum.de

on optical imagery or LiDAR data, respectively. Some promising approaches in this field have been proposed: For example, WANG et al. (2004) employ a combination of edge and local maxima detection on the first principal component of airborne multi-spectral imagery in order to detect tree crowns and tree tops, whereas ERIKSON (2003) proposes a region growing method for the segmentation of tree crowns in aerial false color photographs, including a subsequent classification of the tree species. KOCH et al. (2006) transfer the local maxima approach to rasterized airborne LiDAR data. After possible tree tops are detected, the tree crowns are delineated using a pouring algorithm that makes use of model knowledge about tree shapes. Finally, one of the latest developments concerning single tree detection from remote sensing data is based on 3D segmentation of full waveform LiDAR measurements (REITBERGER et al., 2009).

In opposition to the rich literature on single tree recognition in optical and LiDAR remote sensing data, only little has been published on this topic based on the exploitation of synthetic aperture radar, which is certainly mostly due to the rather coarse resolutions of conventional SAR sensors available so far. One of the few exceptions is (HALLBERG et al., 2005), where VHF band SAR images acquired from multiple viewing angles are coregistered and combined to improve spatial and radiometric resolution. The enhanced imagery is then used to estimate the stem volume of individual trees from their related backscattering amplitudes.
The analysis of individual trees is, however, becoming more and more interesting with the new generation of SAR sensors that offer resolutions in the decimeter range. Therefore, this paper investigates the potential of single tree recognition in airborne single-pass millimeterwave InSAR data. As will be shown later, this sensor/data configuration is considered to be particularly suitable for the task.
In section 2 the data and the relevant sensor characteristics are introduced, in section 3 the geometrical imaging of trees in airborne SAR imagery is discussed. Section 4 shows a first concept for automatic tree recognition. Finally, section 5 discusses the findings of this study, before in section 6 conclusion and outlook are given.

2 Airborne Single-Pass Millimeterwave InSAR Data Characteristics

The main theme of this paper already implies that our work takes advantage of three core sensor properties: First of all, we intend to use interferometric SAR data, which enables the derivation of tree heights and therefore delivers an additional indicator that can be exploited for tree detection. Second, we use single-pass data, giving us the possibility to work with high-coherent InSAR observations, a pre-requesite that has not been available to the satellite community until the launch of TanDEM-X in June 2010 (KRIEGER et al., 2007). Third, we employ millimeterwave radar data in the Ka band. In opposition to the more common X band (wavelength: about 3 cm), millimeterwave data is known to penetrate the tree canopy even less (see Fig. 1), thus giving way to a more accurate height estimation.

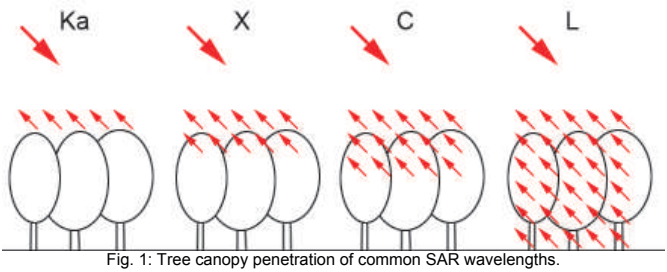


Fig. 1: Tree canopy penetration of common SAR wavelengths.

2.1 MEMPHIS SAR System

For the considerations in this paper, the German MEMPHIS (***M**illimeterwave **E**xperimental **M**ultifrequency **P**olarimetric **H**igh Resolution **I**nterferometric **S**ystem*) by the Fraunhofer Institute for High Frequency Physics and Radar Techniques is used as an exemplary sensor (SCHIMPF et al., 2002). Although it is able to be operated in different modes and configurations, we employ the basic airborne side-looking configuration with a carrier frequency of 35 GHz (Ka band), and a bandwidth of 900 MHz, leading to a slant range resolution of 16.5 cm. Since MEMPHIS is equipped with four receiving antennas, a number of single-pass interferograms with different baselines can be created. However, in this work, only the longest baseline (length: 27.5 cm, corresponding ambiguity height: 42 m) was used; the utilization of the remaining interferograms or a full multi-baseline configuration is part of ongoing research. An overview of the relevant sensor parameters is given in Table 1.

MEMPHIS sensor system	
Band	Ka band (35 GHz)
Wavelength	8.54 mm
Pixel spacing (slant range)	5.1 cm (azimuth) 16.7 cm (range)
Baseline used	27.5 cm
Ambiguity Height	42 m
Test scene "Alte Pinakothek"	
Flying altitude	768 m
Heading angle	340°
Depression angle	30°

Tab. 1: Sensor and test scene parameters

2.2 Test Scene

The available MEMPHIS test data were acquired on a campaign over Munich, Germany, in 2011. While the scene is centered at the main campus of Technische Universitaet Muenchen (TUM) close to the city center and therefore contains mostly dense building blocks, there are some interesting sub-scenes containing a number of trees: The one that is used in this paper is a small group of trees next to the "Alte Pinakothek", a museum close to TUM main campus. The radar intensity in slant range geometry with approximately square pixels and an orthophoto of the scene are shown in Fig. 2. Although the visual impression of the trees is unexpectedly nice in the SAR image, in the following section we will show that due to the imaging geometry an interpretation is still non-trivial. In Fig. 3 ground photographs of the scene are displayed in order to facilitate a deeper understanding of the tree configuration.

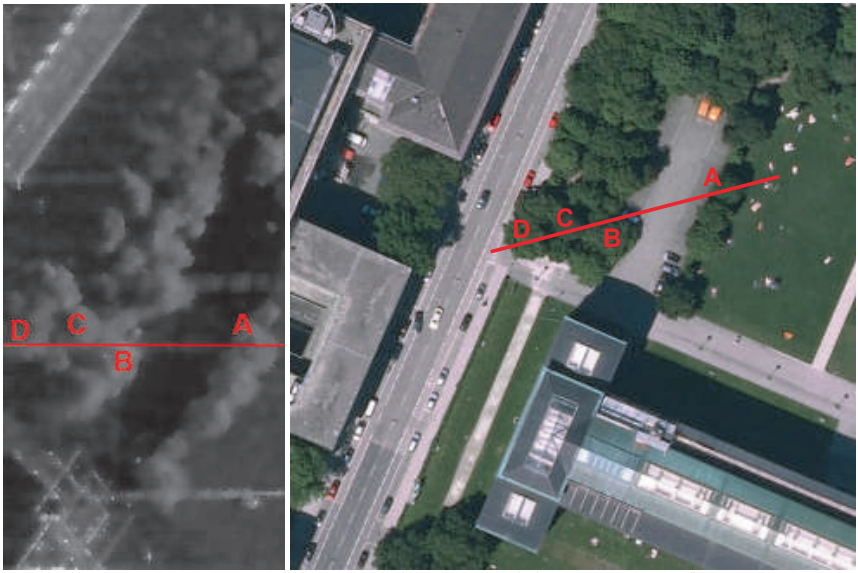


Fig. 2: SAR intensity image in slant range geometry (left), orthophoto of of the "Alte Pinakothek" test scene (right). The red line indicates a range profile corresponding to the radar's viewing direction. A, B, C and D are trees appearing in this range line.

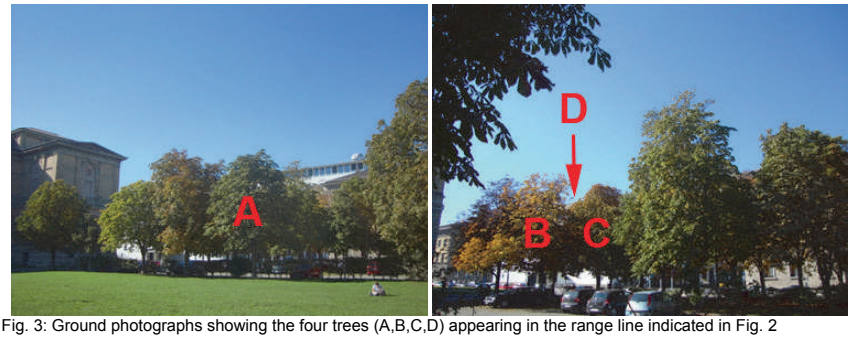


Fig. 3: Ground photographs showing the four trees (A,B,C,D) appearing in the range line indicated in Fig. 2

3 Modelling of the Representation of Trees in InSAR Images

While the pictorial appearance of trees has long been studied for optical remote sensing imagery and LiDAR data, only little is known about the SAR imaging characteristics on single tree level. The main difference to the conventional sensor types is certainly the side-looking imaging geometry leading to well-known effects like foreshortening, layover and shadowing: Layover

generally occurs whenever the terrain slope (or any other object) is steeper than a line perpendicular to the radar line-of-sight defined by its depression angle. Surfaces that are less steep but elevated nonetheless, will not be represented in true size as well. Instead, they will be displayed compressed, while still containing the full backscattering energy. Surfaces facing away from the antenna finally will return no signal to the radar at all. Figure 4 shows a model for the line-of-sight depicted in Fig. 2.

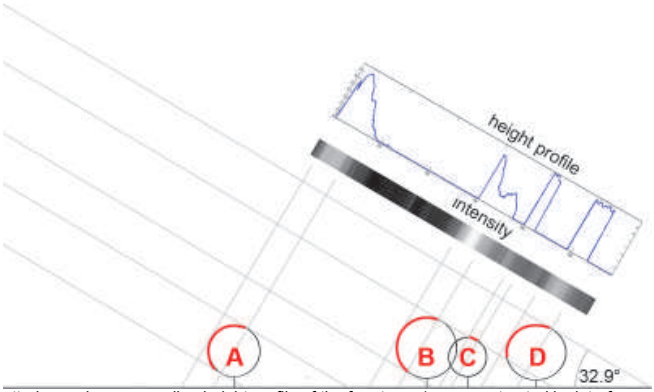


Fig. 4: Backscattering and corresponding height profile of the four trees (as reconstructed by interferometric phase-to-height conversion) for the range profile introduced in Fig. 2.

It can clearly be assessed that only the upper front parts of the trees are directly imaged by the SAR – as long as the respective tree is not affected by the shadowing effect of another tree. Here, tree C is depicted significantly smaller than the other trees due to the fact that the radar line-of-sight does not go through its center but only touches the outer part of its canopy. Therefore, tree C is mostly affected by the shadowing of tree B. Appearances like this naturally exacerbate any kind of automatic image analysis. The locations of the backscattering maxima determined by local maxima detection are shown in Fig. 5. It is obvious that the maximum backscattering energy for each tree appears oriented towards the sensor. Thus, local intensity maxima cannot be considered valid hypotheses for the center (trunk) position of the corresponding trees without further correction.

4 Preliminary Concepts for Automatic Tree Recognition

Considering the imaging peculiarities of the side-looking SAR geometry, we propose two different preliminary concepts for automated tree recognition. The first is meant to take place in the slant range plane, i.e. directly in the SAR image and in products that may be derived from exploitation in the sense of interferometry (e.g. coherence maps, interferograms, height maps).



Fig. 5: Local maxima (white boxes) detected in the SAR intensity image (left). These maxima are projected into an orthophoto (July 2009) overlayed with a LiDAR point cloud (September 2009) in order to show their real world position (right). The SAR viewing direction is indicated by the red arrows. The red circles indicate manually determined tree top positions. Note how all detected maxima are located towards the sensor.

The second is intended to make use of interferometric 3D reconstruction possibilities. Here, in a first step 3D point clouds are derived from the available interferograms, and afterwards trees are detected in the 3D data in the vein of LiDAR-based approaches.

4.1 Image-based Tree Recognition

A flowchart of the image-based concept is shown in Fig. 6. It is inspired by work on tree recognition from optical remote sensing data and centers around the detection of local maxima in the SAR intensity data. As proposed by (POPESCU & WYNNE, 2004), the search radius of the local maxima detection is derived by prior knowledge about the canopy height. In the InSAR case, this height information can easily be derived from the interferometric phase information if two (or more) coregistered single-pass acquisitions of the scene are available. Due to the fact that trees (and, in the case of urban areas, buildings) lead to shadow occlusions, leaving just noisy phase measurements, the according coherence map can be used to eliminate shadow pixels from the interferogram. After the local maxima are found, they serve as preliminary tree hypotheses, and patches of a certain radius are extracted around them. Within each of these patches, edge detection is applied in order to extract the contour caused by strong reflections oriented towards the sensor. Using this contour as input for a circle fitting procedure, finally the tree crown is delineated, whose center is considered to be the true tree position. Using the height previously generated from the interferogram, the tree can be reconstructed with respect to its georeferenced position in a world coordinate system.

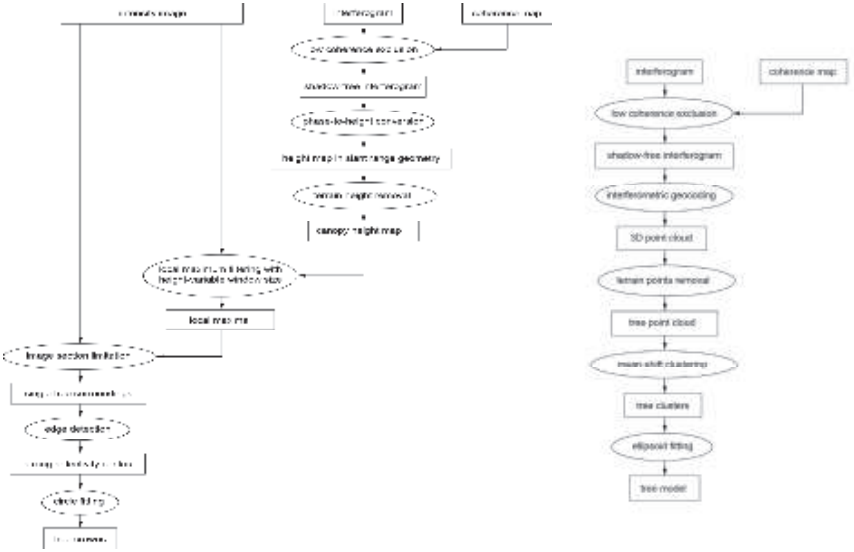


Fig.6: Workflows of the proposed tree recognition concepts. Left: Image-based approach. Right: 3D-based approach.

4.2 3D-based Tree Recognition

The flowchart of the 3D-based concept is also shown in Fig. 6. Here, the (interferometric) SAR data is used to reconstruct georeferenced 3D points, which are then used as input information for tree recognition. Therefore, the general nature of this concept is inspired by the related works of the LiDAR community (GUPTA, 2010). In a first step, the pre-processed InSAR data is geocoded using the interferometric range-Doppler equations, resulting in a point cloud that corresponds to the SAR image resolution concerning point density. For the available MEMPHIS data, this is up to about 100 points per m², highly depending on the terrain slope due to foreshortening and shadowing. An exemplary conventional InSAR point cloud is compared to a LiDAR point cloud in Fig. 7. It can easily be seen that both point clouds differ in their structure, which is due to the side-looking nature of SAR systems. The most important difference is caused by the layover effect that can nicely be seen in form of points shifted towards the sensor. But also the radar shadow has a significant impact causing comparably large holes in the point cloud, which could, however, easily be filled using multi-aspect InSAR methods (SCHMITT & STILLA, 2011). Nevertheless, the overall height trend corresponds to what we see in the LiDAR data. In order to exploit the dense InSAR point cloud, we propose the application of a clustering algorithm, e.g. mean-shift clustering (COMANICIU & MEER, 2002) leading to clusters of 3D points grouped to single trees. The resulting clusters can then be used as input for the parameterization of the tree: For example, height and diameter can be corrected with respect to the known imaging geometry, such that e.g. the mathematical model of a rotation ellipsoid can be fitted to the cluster to deduce a simple tree model.

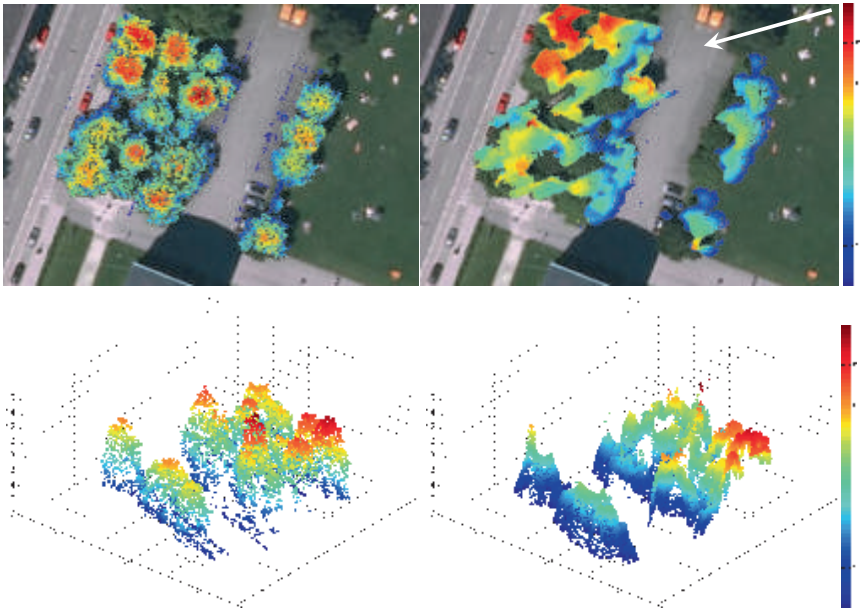


Fig. 7: Top row: Two-dimensional representation of LiDAR point cloud (left) and InSAR point cloud (right). Note how many parts of the InSAR tree heights are affected by layover. The white arrow shows the SAR viewing direction. Bottom row: Three-dimensional view of LiDAR point cloud (left) and InSAR point cloud (right), approximately from the SAR perspective.

5 Discussion

Due to the imaging characteristics of optical remote sensing systems, which are quite familiar to the human visual system or the convenience of high resolution point clouds as they are created by LiDAR sensors, manual, semi-automatic and automatic tree recognition is obviously fancied using data of this kind of sensors. There are, however, a number of arguments that speak in favour of the utilization of synthetic aperture radar for this task as well. First of all, the night-and-day and relative all-weather capability of SAR has to be mentioned. Although forest inventory tasks are usually not as time-critical as disaster response situations, time is still an expense factor. Therefore, the possibility to carry out flight campaigns without respect to the actual weather or daylight conditions provides a noteworthy advantage, especially if we consider that most forested areas see relatively high cloud coverage. Furthermore, SAR has been described to have a good cost-performance ratio since large swaths can be imaged with one single track.

In addition to that, there is another advantage more from the methodological point of view: Single-pass InSAR data provides both radiometric as well as height information in form of

intensity images and interferograms. Both kinds of data are precisely coregistered by their very nature, which enables a joint exploitation without any error-prone pre-processing steps and without the need to buy additional external data. On the other hand, it has to be admitted that the side-looking SAR imaging geometry introduces challenges for both manual and automated image analysis that seem more severe than in conventional remote sensing data. Especially Figs. 5 and 7 nicely illustrate the effects caused by foreshortening (intensity maxima shifted towards sensor), layover (tree heights shifted towards sensor) and shadowing (missing data behind trees). In spite of these drawbacks, a further investigation of forest remote sensing and tree recognition based on single-pass InSAR data seems promising.

6 Conclusion & Outlook

In this paper we have discussed the potential of airborne single-pass millimeterwave InSAR data for the recognition of single trees. The considerations were carried out using experimental data in an urban environment and can readily be extended to forest scenarios. It has been shown that airborne InSAR data with range and azimuth resolutions in the decimeter range can potentially be used for the analysis of wooded areas on a single tree level. Furthermore, the usefulness of single-pass data as well as millimeterwave frequencies has been shown. However, it has to be noted that the side-looking imaging geometry inherent to synthetic aperture radar sensors leads to challenging imaging effects that have to be considered when the tree detection process is modelled.

Finally, two different concepts for automated recognition of individual trees have been introduced: One in the SAR image space exploiting radiometric information as well as height maps in slant range geometry derived from interferometric phase information; the other one in object space exploiting 3D point cloud data reconstructed by interferometric geocoding. The investigation of their applicability, especially with respect to established methods from the field of optical and LiDAR remote sensing, will be in the focus of future research.

7 Acknowledgment

The authors would like to express their thanks to Stephan Stanko and Thorsten Brehm of Fraunhofer FHR as well as Erich Meier and Christophe Magnard of RSL Zurich for providing the MEMPHIS test data. We would also like to thank Marcus Hebel of Fraunhofer IOSB for providing the reference LiDAR data and LVG Muenchen for the orthophoto.

8 Literaturverzeichnis

COMANICIU, D. & MEER, P., 2002: Mean Shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5, pp. 603-619

ERIKSON, M., 2003: Segmentation of individual tree crowns in colour aerial photographs using region growing supported by fuzzy rules. Canadian Journal of Forest Research, Vol. 33, No. 8, pp. 1557-1563

FAGAN, M. & DEFRIES, R., 2009: Measurment and monitoring of the world's forests. A review and summary of remote sensing technical capability 2009-2015. RFF Report

GUPTA, S., 2010: Single tree detection and modeling using airborne laser scanner data. PhD Thesis, Faculty of Forest and Environmental Sciences, Albert-Ludwigs-Universität Freiburg im Breisgau

HALLBERG, B.; SMITH-JONFORSEN, G. & L.M.H. ULANDER, 2005: Measurements on individual trees using multiple VHF SAR images. IEEE Transactions on Geoscience and Remote Sensing, Vol. 43, No. 10, pp. 2261-2269

IZZAWATI; WALLINGTON, E.D. & WOODHOUSE, I.H., 2006: Forest height retrieval from commercial X-band SAR products. IEEE Transactions on Geoscience and Remote Sensing, Vol. 44, No. 4, pp. 863-870

KOCH, B.; HEYDER, U. & WEINACKER, H., 2006: Detection of individual tree crowns in airborne LiDAR data. Photogrammetric Engineering & Remote Sensing, Vol. 72, No. 4, pp. 357-363

KRIEGER, G.; MOREIRA, A.; FIEDLER, H.; HAJNSEK, I.; WERNER, M.; YOUNIS, M. & ZINK, M., 2007: TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry. IEEE Transactions on Geoscience and Remote Sensing, Vol. 45, No. 11, pp. 3317-3340

MERCER, B.; ZHANG, Q.; SCHWÄBISCH, M. & DENBINA, M., 2010: Estimation of forest biomass from an airborne single-pass L-band Pol-InSAR system. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38, No. 1, on CD

NEUMANN, M.; FERRO-FAMIL, L. & REIGBER, A., 2010: Estimation of forest structure, ground, and canopy layer characteristics from multi-baseline polarimetric interferometric SAR data. IEEE Transactions on Geoscience and Remote Sensing, Vol. 48, No. 3, pp. 1086-1104

PERKO, R.; RAGGAM, H.; GUTJAHR, K. & SCHARDT, M., 2010: The capabilities of TerraSAR-X imagery for retrieval of forest parameters. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38, No. 7B, pp. 452-456

POPESCU, S.C. & WYNNE, R.H., 2004: Seeing the trees in the forest: Using LiDAR and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogrammetric Engineering & Remote Sensing, Vol. 70, No. 5, pp. 589-604

REITBERGER, J.; SCHNÖRR, C.; KRZYSTEK, P. & STILLA, U., 2009: 3D segmentation of single trees exploiting full waveform LiDAR data. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 64, No. 6, pp. 561-574

SCHIMPF, H.; ESSEN, H.; BOEHMSDORFF, S. & BREHM, T., 2002: MEMPHIS – a fully polarimetric experimental radar. Proceedings of IEEE International Geoscience and Remote Sensing Symposium, pp. 1714-1716

SCHMITT, M. & STILLA, U., 2011: Fusion of airborne multi-aspect InSAR data by simultaneous backward geocoding. Proceedings of Joint Urban Remote Sensing Event, pp. 53-56

STRAUB, B.-M. & HEIPKE, C., 2001: Automatic extraction of trees for 3D-city models from images and height data. In: BALTSAVIAS, E.P.; GRUEN, A. & VAN GOOL, L. (Eds.): Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), pp. 267-277

WANG, L.; GONG, P. & BIGING, G.S., 2004: Individual tree-crown delineation and treetop detection in high-spatial-resolution aerial imagery. Photogrammetric Engineering & Remote Sensing, Vol. 70, No. 3, pp. 351-357

A study on the detection of understory regenerations in temperate forest areas using airborne full-waveform LiDAR

WEI YAO[1], PETER KRZYSTEK[1]& MARCO HEURICH[2]

Abstract: Airborne remote sensing has been developing as an efficient instrument to supplement the conventional field work towards forest inventory. Unfortunately, understory trees such as young regenerations and low trees in multilayered forest stands are often ignored. However, information about post-disturbance or natural forest regeneration and low vegetation will provide insight into the dynamics of future stand composition and structure. Therefore, a detailed understanding of the availability and spatial distribution of understory regeneration in forested areas is required. Currently, field-based inventory is the principal means used to collect information regarding regeneration status. In this work, we examine the potential of an airborne small-footprint full-waveform laser scanner for the mapping of advanced regeneration in temperate mixed mountain forests. Usually, overstory trees prevent from obtaining a wall-to-wall sample of understory. Therefore, discriminating between the crowns of dominant and suppressed trees in LiDAR point clouds is very challenging in terms of both the sensor technology and detection methodology. This paper aims to undertake a study on detecting understory trees from airborne LiDAR point clouds and is dedicated to presenting an automatic strategy for inferring the spatial distribution of understory trees by extracting 3D individual trees. Understory regenerations are obtained based on classifying the single trees with object-level features. We deal with understory trees in central-Europe spruce-beech stands with overstory mortality due to bark beetles. We applied the developed approach to two sample plots acquired under leaf-on and leaf-off conditions and validate the results by two different evaluation methods based on reference data manually acquired from TLS data of the same sites. For both test sites, we have obtained a detection rate of up to 45% with 25% false alarms based on single-tree evaluation. However, when using the area-wide evaluation we achieved an improved result, namely up to 60% for the detection rate with 18% false alarms. This study made the initial attempt to characterize the spatial distribution of 3D individual understory regeneration trees in a recovering heterogeneous forest from airborne LiDAR and has demonstrated the feasibility and capability of the single tree detection strategy for understory regeneration tree analysis.

1) Wei Yao and Peter Krzystek Department of Geoinformatics, Munich University of Applied Sciences, 80333 Munich, Germany, {Yao, Krzystek}@hm.edu
2) Marco Heurich Bavarian, Forest National Park, 94481 Grafenau, Germany, marco.heurich@npv-bw.bayern.de

1 Introduction

Airborne laser scanning (ALS) or LiDAR has been widely used in mapping the Earth's surface and especially in forestry applications. Techniques for tree extraction from LIDAR data have been investigated for mapping forests at both plot and tree levels to identify important structural parameters (HEURICH et al., 2008; KORPELA et al., 2010; YU et al., 2011). Recent advances in LIDAR technology have demonstrated that new full waveform scanners can provide a higher spatial point density and additional information about the reflectional characteristics and vertical structure of trees (STILLA et al., 2007; REITBERGER et al., 2008; YAO et al., 2010). Novel methods for single tree detection tackle conceptually the segmentation problem with a 3D approach instead of using only the CHM. In combination with full waveform data, REITBERGER et al. (2009) successfully demonstrated that the overall detection rate of single trees could be significantly improved, especially in heterogeneous forest types. The improvement happened mostly to the lower forest layers with 20%. The fusion of 3D techniques with full waveform data pushed the single tree approach to a new level of completeness. Meanwhile, the estimation of tree shape parameters benefits from 3D tree segments. Moreover, the tree reflectional characteristics gain more insights into tree structure which are significant for tree species classification. Therefore, unbiased species-specific models can be used to accurately determine forest biophysical parameters (YAO et al., 2012).

So far, little attention has been paid to identify the growth condition of individual trees and detect young regenerations from heterogeneous forest areas using LiDAR data, mainly due to the low point density and laser transmission loss which lead to the lack of LiDAR sampling of understory tree structure. However, information about post-disturbance tree regeneration will provide insight on dynamics of future stand composition and structure (COLLINS et al. 2011). Information about them would also be valuable to forest and biodiversity management. Therefore, a detailed understanding of the availability and spatial distribution of regeneration trees in forest stands is required. In the past several years there are certain authors dealing with the similar topics as young regeneration characterization. KING et al. (2005) used high resolution airborne imagery to delineate 3D outlines of individual trees and showed good detection accuracy for trees larger than 30 cm crown diameter. The detection is limited to the boreal forest 5–10 years after harvest. In POTTER et al. (2012) the post-fire density of lodgepole pine sapling regeneration in Yellowstone National Park is predicted based on hyperspectral vegetation indices by multiple regression models. Finally, the succession patterns and processes in post-fire forest regeneration are revealed at a large scale. Most recently, KORPELA et al. (2012) and LI et al. (2011) have used airborne and spaceborne LiDAR data, respectively, to map young regenerations re-established following recent disturbances and estimate their height. However, it was still an area-based detection and assessment of the understory regenerations. NÆSSET (2011) has investigated the regression of total above-ground biomass of spruce, pine and birch in young boreal forests based on area-based ALS metrics. However, it has been found out that tree species had a significant influence on the models, so a more species-species model based on the individual tree detection is desirable.

The full waveform LiDAR systems have the possibility to overcome drawbacks of conventional laser scanners since they detect significantly more reflections for trees in the lower forest layer, and provide the intensity and the echo width as reflectivity parameters. The objectives of this work are (i) exploit the potential of airborne full-waveform LiDAR for characterizing understory regenerations in a multi-layered heterogeneous forest single tree detection method, (ii) provide a study analysis result on young regeneration detection based on two evaluation methods.

The paper is divided into five sections. Section 2 focuses on the segmentation of the single trees and classification of their growth conditions. Section 3 presents the experimental results which are obtained from full-waveform LiDAR data of Bavarian Forest National Park. The results are discussed and concluded in the last two sections.

2 Method

2.1 Waveform decomposition

We assume that full waveform LIDAR data have been captured in a region of interest (ROI). A single waveform is decomposed by fitting mixed Gaussian pulse models to the waveform which contains N_R reflections (Fig.1).

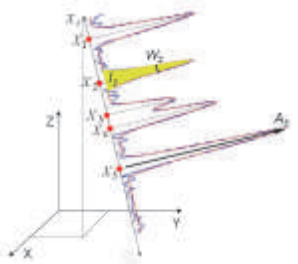


Fig.1: 3D points and attributes derived from a waveform

The vector $\boldsymbol{X}_i^T = (x_i, y_i, z_i, W_i, I_i)(i = 1,..., N_R)$ is provided for each reflection i with (x_i, y_i, z_i) as the 3D coordinates of the echo. Additionally, the points $\boldsymbol{x}_i$ are given the width $W_i = 2 \cdot \sigma_i$ and the intensity $I_i = \sqrt{2 \cdot \pi} \cdot \sigma_i \cdot A_i$ of the echo pulse with σ_i as the standard deviation and A_i as the amplitude of the reflection i (REITBERGER et al., 2008). Note that basically each reflection in the travel path of the laser beam can be detected by the waveform decomposition. This is remarkable since some conventional LIDAR systems have a dead zone of about 3 m which makes these systems effectively blind after triggering a reflection. The LiDAR data are calibrated by referencing W_i and I_i to the pulse width W^e and the intensity I^e of the emitted Gaussian pulse and correcting the intensity with respect to the run length s_i of the laser beam and a nominal distance s_0.

$$W_i^c = W_i / W^e \tag{1}$$

$$I_i^c = (I_i \cdot s_i^k) / (I^e \cdot s_0^k) \tag{2}$$

Note that the correction assumes a target size larger or equal to the footprint. The points from a waveform are subdivided into 4 point classes depending on the number of reflections within a waveform (Tab.1).

Tab.1: Subdivision of points into classes in dependence on the number of reflections N_R and the position i of the reflection in the waveform

Class	Single	First	Middle	Last
Definition	N_R=1 i=1	$N_R \geq 2$ i=1	$N_R \geq 3$ i=2,…, N_R-1	$N_R \geq 2$ i=N_R

2.2 Single tree detection based on a hybrid approach

The hybrid approach for single tree detection consists of two steps: watershed transformation with stem detection and normalized cut segmentation within merged watershed segments.

Firstly, the coarse detection of single trees is achieved from CHM by watershed transformation. The CHM is derived by subdividing the ROI into a grid having a cell spacing of cp and N_C cells. Within each grid cell, the highest 3D point is selected and adapted with respect to the ground level z_j^{ground}, i.e. $z_j^{CHM} = z_j - z_j^{ground} (j = 1,\ldots,N_C)$. The ground level z_j^{ground} is estimated from a given DTM by bilinear interpolation. In the next step, all the highest 3D points $\boldsymbol{X}_j^T = (x_j, y_j, z_j^{CHM}) (j = 1,\ldots,N_C)$ of all N_C cells are robustly interpolated in a grid that has N_X and N_Y grid lines and a grid size g_w. For this purpose an algorithm called 'gridfit' (D'ERRICO, 2006) is adopted which smoothens the surface by maintaining the surface gradients as small as possible. The trade-off between interpolation and regularization is determined by the adjustable smoothing factor. Both steps are carried out simultaneously in a least squares adjustment. The result is a smoothed CHM having equally spaced cells. The watershed segments derived on this CHM act as candidate regions where single trees could be contained. The results can also be improved by an additional stem detection method to further detect smaller trees which are not represented by local maximums.

Within every watershed segments the 3D segmentation technique using normalized cuts (SHI & MALIK, 2000) is used to detect point clouds associated to single trees. This makes it possible to detect also smaller understory trees which cannot be indicated by local maxima in the CHM. This segmentation uses the positions (x_i, y_i, z_i) of the laser reflections and optionally the pulse width W_i and the intensity I_i of the waveform decomposition. Additionally, stem positions or local maximums derived by the watershed segmentation of CHM can be used as a-prior knowledge. The normalized cut segmentation applied to the voxel structure of a (merged) watershed segment is based on a graph G describing the adjacent topology between each voxel. The two disjoint segments A and B of the graph are found by minimizing the cost function:

$$NCut(A,B) = \frac{Cut(A,B)}{Assoc(A,V)} + \frac{Cut(A,B)}{Assoc(B,V)} \tag{3}$$

with $Cut(A,B) = \sum_{i \in A, j \in B} w_{ij}$ as the total sum of weights between the segments A and B and $Assoc(A,V) = \sum_{i \in A, j \in V} w_{ij}$ as the sum of the weights of all edges ending in the segment A. The weights w_{ij} specify the similarity between the voxels and are a function of the LiDAR point distribution and various features. A minimum solution for Eq.(3) is found by means of a corresponding generalized eigenvalue problem (REITBERGER et al., 2009). It turned out that the spatial distribution of the LiDAR points mainly influences the weighting function. The features derived from the LiDAR points attributes of W_i and I_i only support in second instance the segmentation result. Note that the 3D segmentation approach is not limited to full waveform LiDAR data. It can also successfully be applied to conventional LiDAR data just providing 3D point coordinates.

2.3 Feature extraction and classification

Deriving significant features describing each tree individually is a key step in the classification of tree growth condition. The single tree detection provides for each segmented tree corresponding laser points. Crown points can be separated from possible existing stem reflections by finding the crown base height or stem detection. All the laser hits above the crown base height form the crown points.

Based on our previous works (REITBERGER et al., 2008; YAO et al., 2012), the four-group salient features $\boldsymbol{S}_{\text{tree}} = \{\boldsymbol{S}_g, \boldsymbol{S}_I, \boldsymbol{S}_h, \boldsymbol{S}_c\}$ of a tree can be defined to reflect the geometric and physical properties of trees. The features will be defined as follows, respectively, while the detailed definition of the first two features can be referred to REITBERGER et al. (2008):

- $\boldsymbol{S}_g = \{ S_g^1 \}$, which describe the outer tree geometry. The parameter S_g^1 records the mean distances of points in all equidistance height-layers to the tree stem position.
- $\boldsymbol{S}_I = \{ S_I^2 \}$, which describe the tree reflectivity properties against laser beam and benefits from intensity information provided by the waveform decomposition. The parameter S_I^2 is introduced as overall mean intensity value for the entire tree segment.
- $\boldsymbol{S}_h = \{ S_h \}$, it uses tree height as a feature, which is derived from subtracting the maximum height of tree segments from the height of DTM at the stem position.
- $\boldsymbol{S}_c = \{ S_C^d, S_C^V \}$, where S_C^d is the maximal diameter of the tree crown, S_C^V is the volume of the tree crown derived based on the 3D alpha shape. The tree crowns are extracted from 3D single tree segments.

Tree growth conditions are distinguished between mature trees and young regenerations by a supervised classification using support vector machine (SVM). SVM is selected due to its computational simplicity and superior accuracy. It is not constrained to prior assumptions on the distribution of input data and is, hence, well suited for complex feature space, e.g. nonlinear recognition problems. The SVM was originally designed for binary classification; therefore, in this study, the SVM is directly implemented for our task in a genuine way (mature and young trees) without handling the multiclass problem. Moreover, the RBF kernel is used and controlling parameters including the kernel parameters were selected automatically based on the LOO (leave-one-out) algorithm (CHAPELLE et al., 2002). After obtaining reasonable initial parameters, they can be refined to get better accuracy.

3 Experiment

3.1 Material

Experiments were conducted in the Bavarian Forest National Park which is located in south-eastern Germany along the border to the Czech Republic (49° 3' 19" N, 13° 12' 9" E). Two sample plots with an area size between 1000 m^2 and 2400 m^2 were selected in the mixed mountain heterogeneous forests. The plots comprise forest in the regeneration phase, the late pole phase. The test sites have suffered and regenerated from tree disease due to bark beetle attack in the last decade. Tab.2 summarizes the characteristics of the individual sample plots.

Tab.2: Characteristics of sample plots

Plot	Flight time	Size[ha]	Altitude[m]	Trees/ha	Deciduous[%]	Percentage of young regenerations [%]
I	2009	0.22	610	830	5	30
II	2011	0.12	830	2400	10	75

Full waveform data have been collected by Milan Flug GmbH with the Riegl LMS-Q560 scanner in August 2009 during foliation and in April 2011 after snowmelt but prior to foliation with an average point density of 25-30 points/m^2 (Tab.3). The vertical sampling distance was 15 cm, the pulse width at half maximum reached 4 ns and the laser wavelength was 1550 nm. The flying altitude of 400 m resulted in a footprint size of 20 cm.

Tab.3: Configurations of two airborne LiDAR campaigns

Time of flight	Aug '2009	Apr '2011
Dataset/Plot	I	II
Foliage	Leaf-on	Leaf-off
Scanner	Riegl LMS-Q560	Riegl LMS-Q560
Pts/m^2	25	30
AGL [m]	400	400
Footprint [cm]	20	20

3.2 Calibration

The relative intensity calibration of the Riegl full waveform system was determined from the special calibration flight performed over an airfield. Several tracks were flown at different flying heights (200 m and 400 m) along and across the airfield. The mean intensity I_i, is corrected with respect to the emitted intensity I_e, and the mean run length s_i were calculated in four homogeneous areas (100 m^2–144 m^2) for each track. According to Eq. (2), the best coefficient k was estimated from all possible observation equations

$$I_i \cdot s_i^k = I_j \cdot s_j^k \tag{5}$$

which can be formulated for two tracks i and j flown at different heights. Tab.4 shows the results obtained for the two flights of datasets.

Tab.4: Estimation of calibration parameter *k*

	Flight 2009	Flight 2011
Calibrated parameter *k*	1.871	1.935

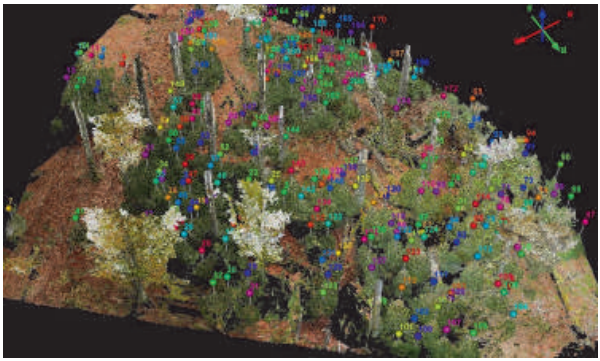
3.3 Reference data and evaluation

The referenced data for detecting young regenerations in the two sample plots were manually either cropped out from or marked in very high point-density TLS data of the same sites. We applied the strategy to the sample plots acquired under leaf-on and leaf-off conditions and validate the results by two different methods: single-tree based and area-based evaluations. The single tree based evaluation is performed based on single-tree level by finding matched tree segments in the reference. The tree detection results were evaluated by the comparison with reference data using two criterions simultaneously: i). the distance between detected and reference trees should be smaller than 60% of the mean tree spacing of the plot; ii) the height difference between detected and reference trees should be smaller than 15% of h_{top}, where h_{top} is the average height of the 100 highest trees per hectare in the sample plot. If a reference tree is assigned to more than one tree position, the tree position with the minimum distance to the reference is selected. Detected trees that are liked to one reference tree position are so-called "detected trees" and detected trees without any link to a tree position are treated as "false positives" (HEURICH, 2006). For single-tree based evaluation, the reference data for single trees including young and mature ones are marked one by one via human operator within TLS point clouds whenever possible. The tree parameters like tree height and stem position can be measured and determined. However, information about tree species have not acquired yet due to observation uncertainty and difficulty in TLS data. The area-based evaluation is performed by area (pixel)-wise comparing detected regenerations to the ground truth in the horizontal (x-y) dimension by checking the degree of pixel overlap. The reference data for young regenerations are directly cropped from TLS data. At first, all the points lying in the interval between 0.5 m and 3 m above the ground level are cropped out as candidate points. Then, a post-editing operation is performed to eliminate outliers and remained stems of mature trees. Both point clouds of detected and reference young regenerations are projected onto the horizontal (x-y) dimensions and converted into a georeferenced raster format with a grid cell size of 1 m. If corresponding grid cells in both projected georeferenced raster are occupied by points, it is a detected pixel for young regenerations. If only the grid cell/pixel in the projected georeferenced raster of ALS/TLS data is occupied, it is treated as "false positives/misdetections".

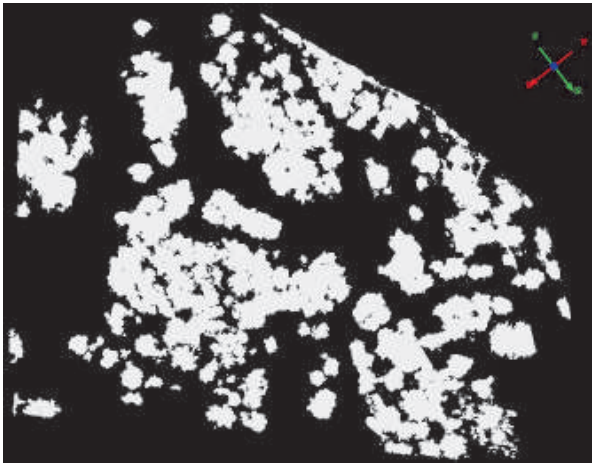
Tab.5 Field data for single tree based evaluation

Time of acquisition	Aug. 2009	Apr. 2011
Dataset	I	II
	Tree height(m)	Tree height(m)
Min	0.81	0.12
Max	3.43	3.30
Mean	2.35	1.32
Standard deviation	0.71	0.61

Reference data for single tree based evaluation have been collected for 50 and 247 trees, for Dataset I and II, respectively. A DTM with a grid size of 1 m and an absolute accuracy of 25 cm was available for the test sites (HEURICH et al., 2008). Tab.5 gives us the descriptive statistics of the filed trees in reference data for single tree based evaluation. However, DBH of young regenerations have not been determined in the reference data, since even the high-density TLS data cannot guarantee the accurate geometric reconstruction of stems of small young regenerations.



(a)



(b)

Fig. 3: Exemplary reference data for Dataset II (a) single tree based evaluation, (b) area-coverage based evaluation

Fig.3(a) shows the reference data for single tree based evaluation of dataset II, which is directly marked in the colored TLS point cloud. The tree tops in the reference data correspond to bottom arrow of the markers representing the presence of reference trees in TLS data. Please keep in mind that the crowns of dominated high trees are deliberately cut out in Fig.3(a) in order to enhance the visualization effect. Fig.3(b) shows the reference data for area coverage based evaluation of dataset II, which is already converted into a georeferenced raster.

3.4 Results

The procedure for 3D single tree detection was applied to the sample plots in a batch procedure without any manual interaction. A binary SVM classification is applied to automatically detected

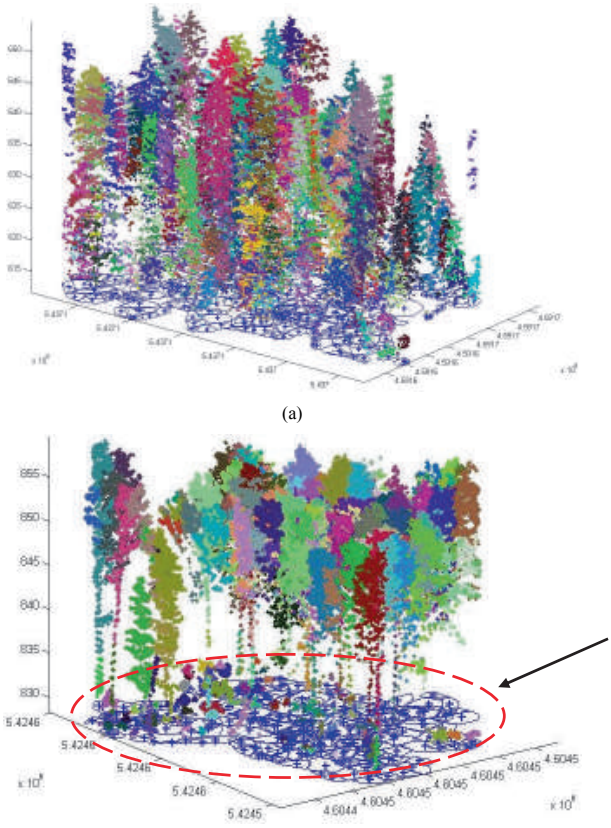


Fig. 4: Single tree detection towards extracting young regenerations in two sample plots, (a): Dataset I, (b): Dataset II.

trees to distinguish between young and mature ones based on various features defined at single tree level. Randomly selecting one fifth of total test data as training data is performed in order to fairly assess the results and minimize the impact of the selection of training data. Fig.4 shows the results of single tree detection from full-waveform LiDAR data towards characterizing tree regenerations, where point clouds of each detected tree are rendered in 3D.

3.4.1 Single tree based evaluation

Tab.6 shows the results of young regeneration detection based on single tree evaluation for the two datasets. The single tree based evaluation showed a better performance of the proposed algorithm with respect to the dataset I, which was acquired in leaf-on condition with less young regeneration trees.

Tab.6: Detection of young regeneration based on single tree evaluation

Dataset	Detection rate [%]	False Pos [%]
I	45.7	25.8
II	40.1	27.8

3.4.2 Area coverage based evaluation

Tab. 7 shows the results of young regeneration detection based on area coverage evaluation for the two datasets. However, compared to former case, the area coverage based evaluation showed a better performance of the proposed approach with respect to the dataset II, which was acquired in leaf-off condition with more young regeneration trees but less high mature trees.

Tab.7: Detection of young regeneration based on area coverage evaluation

Dataset	Detection rate [%]	False Pos [%]
I	39.5	27.5
II	58.2	18.5

4 Discussion

A hybrid strategy by combing 3D segmentation with the object-based classification was developed to extract understory forest regenerations in a heterogeneous temperate forest. The experimental results demonstrate that the moderate capability of single tree detection approach for characterizing the young regenerations when compared to the ground truth extracted from TLS data. When we compare these results to that of tree detection in the low forest layer from REITBERGER et al. (2009), it can be found that we have achieved a more encouraging detection rate, but also associated with a higher false alarm.

When we compare data set I (leaf-on) and data set II (leaf-off) we still cannot address the question whether the foliage condition could affect the detection of young regenerations, since the two test plots do not represent exactly the same forest area. However, the detection rate is in leaf-on situation by approximately 5% worse for single-tree evaluation. But in leaf-on situation

the results are reversed leading to a loss of detection of 20% based on area-coverage evaluation. A low percentage of the young regenerations in plot I lead to a relatively higher stem density of tall overstory trees, which poses an obstacle for the laser pulse penetration to hit understory young trees. That could also be a reason for better detection performance for the plot II based on the area coverage evaluation scheme. Moreover, the higher stem densities of the upper forest layer and leaf-on foliage condition are able to weaken the intensity of echo pulses from the young regenerations. However, the better performance for plot I based on the single tree evaluation could partially be attributed to the few total number of the young regenerations and the relatively greater mean tree height in the test site.

If we take a glance at the results of the two test plots with respect to two different evaluation methods, it can be found that contrary to our expectation the results obtained by area-coverage evaluation is almost by ca. 5-% worse than that obtained by single tree based evaluation for plot I, whereas for plot II the detection rate has much improved by almost 20%. There could be two major reasons to cause this effect. Firstly, the reference data for area-coverage evaluation scheme are directly collected from the TLS data based on the digital terrain model (DTM) of the forested areas. The automated derivation of the DTM from TLS is not always error-free, which can incorporate some ground-layer vegetation such as grass or shrub into reference data of the regeneration cover. Secondly, other than an object-based evaluation scheme which is used by single tree evaluation, the pixel to pixel matching method used by the area-coverage evaluation scheme generates a more strict measure to assess the object distribution. The pixel-based evaluation method is more demanding especially when comparing the detected ALS regeneration pixels directly to TLS regeneration pixels, since the understory forest regenerations can be much better scanned by TLS in a more complete and accurate way. Consequently, the factors mentioned above could lead the detection results evaluated by the area-coverage method to be more susceptible to the quality of reference data, the topography of test sites and forest structures.

5 Conclusion and outlook

We have performed a study on the characterization of understory forest regenerations in a temperate forest using airborne LiDAR data. This study has chosen the TLS data to replace field works and serve as the data source for retrieving the reference. The feasibility and a moderate performance of using a 3D segmentation method combined with object-based classification are achieved for extracting understory regeneration information in forested areas from point clouds acquired by a full-waveform LiDAR. Up to ca.60% of the coverage area of young regenerations being lower than 3 meters can be successfully demonstrated to be detected in our experiments. Furthermore, it is necessary to quantify the influence of the uncertainty associated with markers manually placed in TLS data which represent the entity of single forest regenerations. Moreover, the DBH and tree species information of single regenerations will be attempted to be directly determined from the TLS data by post-editing, in order to enable a species-specific model for estimating the attributes of single regenerations from ALS data.

6 References

CHAPELLE, O., VAPNIK, V., BOUSQUET, O. & MUKHHERJEE, S., 2002: Choosing multiple parameters for support vector machine. Machine Learning, 46 (1-3), 131–159.

COLLINS, B J., RHOADES, C.C., HUBBARD, R M. & BATTAGLIA, M A., 2011: Tree regeneration and future stand development after bark beetle infestation and harvesting in Colorado lodgepole pine stands. Forest Ecology and Management, 261, 2168-2175.

D'ERRICO, J., 2006: Surface fitting using gridfit. http://www.mathworks.com/matlabcentral/fileexchange(accessd on 10.01.2009)

HEURICH, M., 2006: Evaluierung und Entwicklung von Methoden zur automatisierten Erfassung von Waldstrukturen aus Daten flugzeuggetragener Fernerkundungssensoren. Forstliche Forschungsberichte München. Nr. 2002/2006. 328 S. und Anlage.

HEURICH, M., FISCHER, F., KNÖRZER, O. & KRZYSTEK, P., 2008: Assesment of digital terrain Models (DTM) from data gathered with airborne laserscanning in temperate European beech (Fagus sylvatica) and Norway spruce (Picea abies) forests. Photogrammetrie, Fernerkundung, Geoinformation. 6/2008. 473-488.

HEURICH, M., 2008: Automatic recognition and measurement of single trees based on data from airborne laser scanning over the richly structured natural forests of the Bavarian Forest National Park. Forest Ecology and Management. 255 (2008) 2416–2433.

KING, D J., PITT, D G. & POULIOT, D A., 2005: Development and evaluation of an automated tree detection–delineation algorithm for monitoring regenerating coniferous forests. Canadian Journal of Forest Research, 35(10), 2232-2345.

KORPELA, I., ØRKA, H.O., MALTAMO, M., TOKOLA T. & HYYPPÄ, J., 2010: Tree species classification using airborne LiDAR – effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. Silva Fennica, 44(2), 319–339.

KORPELA, I., HOVI, A. & MORSDORF, F., (2012) Understory trees in airborne LiDAR data — Selective mapping due to transmission losses and echo-triggering mechanisms. Remote Sensing of Environment, Volume 119, 92-104.

LI, A., HUANG, C., SUN, G., SHI, H, TONEY, C., ZHU, Z., ROLLINS, M., GOWARD, S. & MASEK J., 2011: Modeling the height of young forests regenerating from recent disturbances in Mississippi using Landsat and ICESat data. Remote Sensing of Environment, 115, 1837-1849.

MARTINUZZI, S., VIERLING, L A., GOULD, W A., FALKOWSKI, M J., EVANS, J S., ANDREW T. HUDAK, A T. & VIERLING, K T., 2009: Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability, Remote Sensing of Environment, Volume 113, Issue 12, 2533-2546.

NÆSSET E., 2011: Estimating above-ground biomass in young forests with airborne laser scanning. International Journal of Remote Sensing, 32(2), 473-501.

POTTER, C., LI, S., HUANG, S. & CRABTREE, R L., 2012: Analysis of sapling density regeneration in Yellowstone National Park with hyperspectral remote sensing data, Remote Sensing of Environment, 121, 61-68.

REITBERGER, J., KRZYSTEK, P. & STILLA, U., 2008: Analysis of full waveform LiDAR data for the classification of deciduous and coniferous trees. International Journal of Remote Sensing, 29(5), 1407-1431.

REITBERGER, J., SCHNÖRR, C., KRZYSTEK, P. & STILLA, U., 2009: 3D segmentation of single trees exploiting full waveform LiDAR data. ISPRS Journal of Photogrammetry and Remote Sensing, 64(6), 561–574.

SHI, J. & MALIK, J., 2000: Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(8), 888 -905.

STILLA, U., YAO, W. & JUTZI, B., 2007: Detection of weak laser pulses by full waveform stacking. In: Stilla U, et al (eds) PIA07 Photogrammetric Image Analysis 2007. International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences, Vol 36(3/W49A): 25-30.

YAO, W. & STILLA, U., 2010: Mutual Enhancement of Weak Laser Pulses for Point Cloud Enrichment Based on Full-Waveform Analysis. IEEE Transactions on Geoscience and Remote Sensing, 48(9), 3571-3579.

YAO, W., KRZYSTEK, P. & HEURICH, M., 2012: Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. Remote Sensing of Environment, Vol.123, 368-380.

YU, X., HYYPPÄ, J., VASTARANTA, M., HOLOPAINEN, M. & VIITALA, R., 2011: Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. ISPRS Journal of Photogrammetry and Remote Sensing, 66(1), 28-37.

UAV-gestützte Datenerfassung für Anwendungen der Landesvermessung – das Hessigheim-Projekt

MICHAEL CRAMER, NORBERT HAALA, MATHIAS ROTHERMEL, BENEDIKT LEINSS & DIETER FRITSCH

Kurzfassung: *Die flexible, aktuelle und hochauflösende Erfassung von Luftbildern kleinräumiger Gebiete ist ein Anwendungsfeld, welches zugeschnitten ist für die Verwendung von unbemannten Fluggeräten (UAV, heute oftmals auch mit RPAS (Remotely Piloted Aircraft System) bezeichnet). Bemannte Bildflüge mit klassischen Luftbildflugzeugen und großformatigen Kamerasystemen lassen sich aus Kostengründen für derartige Anwendungen kaum einsetzen. Diese Lücke kann durch UAVs sinnvoll geschlossen werden. Auch im Rahmen der Aufgaben der Landesvermessung gibt es Anwendungen, bei denen der Einsatz von UAVs sinnvoll erscheint. Hier sind z.B. Flurneuordnungsverfahren zu nennen, die oftmals kleinräumig stattfinden und einen hohen Bedarf an aktuellen Bilddaten haben, auch um die Teilnehmergemeinschaft anschaulich über den Fortgang der Arbeiten zu informieren. Das Landesamt für Geoinformation und Landentwicklung BW (LGL BW) hat daher in Kooperation mit dem Institut für Photogrammetrie (ifp) der Universität Stuttgart das Pilotprojekt UAV@LGL durchgeführt, in dem stellvertretend für alle Landesvermessungsbehörden in Deutschland das Potenzial der UAV-gestützten Datenerfassung im Rahmen eines Flurneuordnungsverfahrens untersucht werden sollte. Im Rahmen des Beitrags werden die Ergebnisse aus den durchgeführten Projektbefliegungen vorgestellt. Dazu liegen zwei UAV-Flüge mit unterschiedlichen Kameras (Canon IXUS 100 und Ricoh GXR Mount A12) vor. Ein Schwerpunkt liegt dabei auf dem Vergleich der Kamerasysteme (u.a. geometrische Auflösungsuntersuchungen) und der Analyse des Genauigkeitspotenzials der 3D-Punktbestimmung im Rahmen der Aerotriangulation (AT). Die orientierten Bilder werden anschließend benutzt, um dichte digitale Oberflächenmodelle (DOM) zu generieren, basierend auf einem modifizierten Semi-Global Matching Ansatz, welcher verwendet wird, um Stereo-Korrespondenzen zu bestimmen. Die rekonstruierten Oberflächenmodelle werden hinsichtlich Genauigkeit und Vollständigkeit evaluiert. Aufgrund der vorliegenden Resultate kann die UAV-gestützte Datenerfassung in operationellen Anwendungen der Landesvermessung als erfolgreich betrachtet werden. Die flexible Verwendung derartiger Plattformen ist jedoch durch das derzeitige Genehmigungsverfahren zur Erlangung einer Aufstiegserlaubnis stark limitiert.*

1 Einleitung

Unbemannte Fluggeräte sind auch in der zivilen Praxis angekommen! Dies zeigt sich nicht nur auf wissenschaftlichen Konferenzen – UAV war neben der 3D-Punktwolkengenerierung „das“ Thema auf dem ISPRS Kongress in Melbourne –, sondern vor allem auch an der starken Präsenz dieser Systeme

Dr.-Ing. Michael Cramer, Universität Stuttgart, Institut für Photogrammetrie (ifp), 70174 Stuttgart, Tel.: +49-711-685-84118, e-mail: michael.cramer@ifp.uni-stuttgart.de .

Prof. Dr.-Ing. Norbert Haala, Universität Stuttgart, Institut für Photogrammetrie (ifp), 70174 Stuttgart, Tel.: +49-711-685-83383, e-mail: norbert.haala@ifp.uni-stuttgart.de .

Dipl.-Ing. Mathias Rothermel, Universität Stuttgart, Institut für Photogrammetrie (ifp), 70174 Stuttgart, Tel.: +49-711-685-84094, e-mail: mathias.rothermel@ifp.uni-stuttgart.de .

cand.geod. Benedikt Leinss, Ödweilerhofstraße 8, 73104 Börtlingen, ben.leinss@web.de .

Prof. Dr.-Ing. Dieter Fritsch, Universität Stuttgart, Institut für Photogrammetrie (ifp), 70174 Stuttgart, Tel.: +49-711-685-83386, e-mail: dieter.fritsch@ifp.uni-stuttgart.de .

auf kommerziellen Ausstellungen und Messen, wie der Intergeo 2012 in Hannover. Obwohl ein großer Teil der präsentierten Anwendungsfelder momentan im Bereich der rein visuellen Inspektion oder Dokumentation liegt (z.B. Unfallsituationen oder dynamische Prozesse wie Menschenansammlungen oder Fortschritte von Bauvorhaben), erlangen zunehmend auch geodätische Anwendungen im Bereich der Photogrammetrie (3D-Objektpunktbestimmung / Oberflächenmodelle) an Bedeutung. Der große Vorteil der UAV-gestützten Datenerfassung ist der in der Regel flexible und kostengünstige Einsatz (im Vergleich zu klassisch bemannten Bildflügen), vor allem wenn es sich um sehr kleinräumige Gebiete handelt. Im Rahmen der Landesvermessung trifft dies z.B. auf Flurneuordnungsverfahren zu, die sich in lokal beschränkten Bereichen abspielen, und über einen längeren Zeitraum ablaufen. Derartige Verfahren werden in enger Abstimmung mit den betroffenen Eigentümern (Teilnehmergemeinschaft) durchgeführt. Der jeweilige Stand der Arbeiten ist der Teilnehmergemeinschaft kontinuierlich und vor allem anschaulich zu präsentieren. Hier bietet sich natürlich vor allem die photogrammetrische Aufnahme an.

Im Rahmen des UAV@LGL-Projekts wurde erstmals in Deutschland das Potenzial der UAV-Datenerfassung für Anwendungen in der Landesvermessung am praktischen Beispiel analysiert. Das Landesamt für Geoinformation und Landentwicklung Baden-Württemberg (LGL BW) wurde im Rahmen der Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) als Pilotanwender definiert. Damit können die im Rahmen dieses Projekts durchgeführten Studien als richtungsweisend für die zukünftige Rolle von UAVs in der deutschen Landesvermessung gesehen werden.
Das LGL BW formuliert die Motivation für die Verwendung von UAV wie folgt: „Vor allem zur Aktualisierung und Qualitätssicherung der Geobasisinformationen, für Flurneuordnungsverfahren (spezielle photogrammetrische 3D-Auswertungen), für die Öffentlichkeitsarbeit sowie zu Dokumentationszwecken (z.B. Schrägluftbilder für Flurneuordnungsverfahren und Touristik, Flugsimulationen) prüft das LGL weitere wirtschaftliche und flexible Erfassungsmethoden als Ergänzung bzw. Alternative zu den […] großflächigen Befliegungen. In einem Pilotprojekt soll deshalb für kleinräumige, lokal begrenzte Interessensgebiete der Einsatz von UAVs untersucht werden. […] Ziel des Pilotprojekts ist, Erkenntnisse zur Methodik der Datenerfassung und -verarbeitung sowie zur Qualität, Wirtschaftlichkeit und Praxistauglichkeit des Einsatzes von UAVs zu gewinnen. (LGL, 2012)“. Zur Beantwortung dieser Fragestellungen wurden auf Vorschlag vom LGL die nachfolgend präsentierten Untersuchungen im Projektgebiet „Hessigheimer Felsengärten“, als Teil des Flurneuordnungsverfahrens Hessigheim (Wurmberg) durchgeführt.

2 UAV-Systemträger

Die UAV-Befliegungen wurden in enger Kooperation mit dem Institut für Flugmechanik und Flugregelung (iFR) der Fakultät Luft- und Raumfahrttechnik & Geodäsie der Universität Stuttgart durchgeführt. Das iFR verfügt über verschiedene Plattformen zur UAV-gestützten Datenerfassung, wobei der Forschungsschwerpunkt im Bereich UAV auf der Entwicklung von Autonomiefunktionen liegt. Durch die Kooperationen mit dem ifp fanden seit Sommer 2011 erste photogrammetrische Bildflüge statt. Diese Erstflüge wurden mit dem Multiplex Twinstar II Träger durchgeführt, einem Starrflügler mit zwei Elektromotoren, einer Spannweite von 1.4 m und einem Abfluggewicht von ca. 1.3 kg. Die Nutzlastkapazität von ca. 300 g erlaubt es, eine kleine Digitalkamera mitzuführen. Die maximale Flugzeit liegt bei ca. 20 min.
Das Flugzeug besteht aus einem Elapor Schaumstoffkörper, der es zu einer leichten aber auch sehr robusten und preisgünstigen Trägerplattform macht. Das Flugzeug landet ohne Fahrwerk

unmittelbar auf dem Flugzeugrumpf. Dies ist wichtig für eine möglichst vielseitige Anwendung, da keine besonderen Anforderungen an eine befestigte Landebahn gemacht werden müssen. Eine normale Wiese reicht in der Regel aus, um das Flugzeug zu landen. Der Start erfolgt aus der Hand.

Das Computersystem im Flugzeug ist eine Eigenentwicklung des iFR. Die wesentlichen Komponenten sind ein Single Board Computer (SBC) auf Linux-Basis, welcher für Navigations- und Regelungsaufgaben genutzt wird, sowie einem Field Programmable Gate Array (FPGA), der in der Standardkonfiguration für das Einlesen von Sensordaten sowie die Schaltlogik zur sicheren Umschaltung zwischen manuellem und rechnergestütztem Modus zuständig ist. Weitere Details können auch IFR (2012) entnommen werden. Die Sensorik im Flugzeug umfasst verschiedene Komponenten zur Echtzeitnavigation. Eine 9-DOF IMU liefert Drehraten, Beschleunigungen und Magnetfeld in jeweils 3 Achsen. Weitere Bestandteile sind ein barometrischer Höhenmesser, ein Differenzdrucksensor und ein L1-C/A-Code GPS-Empfänger (uBlox LEA-6T). Diese Sensordaten werden in einem erweiterten Kalmanfilter fusioniert und daraus die Informationen über den aktuellen Flugzustand geschätzt. Für die spätere post-processing Auswertung werden die Rohdaten der GPS-Messungen ebenfalls an Bord gespeichert. Weitere Details zur Systemkonfiguration sind HAALA et al. (2011) zu entnehmen.

Für die Erstbefliegung im Rahmen des UAV@LGL Projekts wurde die Twinstar II zusammen mit der Canon IXUS 100 IS im März 2012 über dem Projektgebiet geflogen. Die Kamera Ricoh GXR Mount A12 hat inklusive des Objektivs ein Gesamtgewicht von ca. 650 g, das für die Twinstar II zu schwer ist. Daher wurde die Zweitbefliegung mit dem Flugzeug Maja der Firma Bormatec durchgeführt, einer ebenfalls kommerziell verfügbaren Plattform, welche über eine höhere Nutzlastkapazität verfügt und bereits für die Aufnahme von Equipment z.B. für Luftbildaufnahmen ausgelegt ist (BORMATEC, 2012). Der Rumpf kann dazu in fast kompletter Länge aufgeklappt werden. Das einmotorige Flugzeug (Luftschraube am Heck hinter dem Leitwerk angebracht) besteht aus expandiertem Polypropylen (EPP) und kann wie die Twinstar II auf dem Rumpf gelandet werden. Es ist ebenfalls mit dem iFR onboard Computersystem ausgestattet. Die sonstigen technischen Daten sind wie folgt: Spannweite 220 cm, Länge 120 cm, Abfluggewicht: 3,5 kg (inkl. Kamera 650 g, Akku 545 g, Elektronik 300 g). Die Flugzeit liegt bei ca. 30 min.

Die nachfolgende Abbildung 1 zeigt beide Trägersysteme und die jeweilige Kamerainstallation. Während in der Twinstar II die Kamera mit faserverstärktem Klebeband in einer Rumpfaussparung befestigt wurde, ist in der Maja eine Pertinaxplatte im Flugzeugrumpf montiert, an der die Kamera mittels Objektivgewinde fixiert werden kann. Zusätzlich befindet sich vor dem Objektiv eine mittels Servo verschiebbare Klappe, die zur Landung vor die Optik geschoben wird, um diese vor Verschmutzungen und Beschädigungen zu schützen.





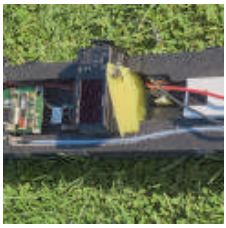



Abb. 1: UAV-Trägersysteme: Twinstar II mit Canon IXUS 100 (links) und Maja mit Ricoh GXR Mount A12 (rechts).

3 Kamerasysteme

Für die Aufnahme der Bilder kamen zwei verschiedene Kamerasysteme zum Einsatz. Für die Erstbefliegungen mit der Twinstar II wurde eine Canon IXUS 100 IS Kompaktdigitalkamera mit 12 MPix (3000 x 4000 pix) eingesetzt. Die Kamera hat eine nominelle Brennweite von 5.9 mm. In Kombination mit einer 1/2.3“ CCD – korrespondierende CCD-Fläche 4.62 x 6.16 mm², Bilddiagonale 7.7 mm – realisiert die Kamera einen maximalen Bildwinkel von ca. 66°, im regulären Kleinbildformat entsprechend einer Weitwinkelbrennweite von ca. 35 mm. Die individuellen Pixel auf der CCD haben nur eine Größe von 1.54 x 1.54 µm², was sich negativ auf die Empfindlichkeit bzw. das Rauschverhalten der Kamera auswirken kann. Wie viele Modelle der Consumer-Klasse verfügt auch diese Kamera über einen optischen Bildstabilisator, der vor allem dazu dienen soll, bei schlechten Lichtverhältnissen ohne Einsatz eines Stativs verwacklungsfreie Aufnahmen mit längeren Verschlusszeiten zu ermöglichen. Zu dem Zeitpunkt, als die Kamera als optischer Sensor für diese Plattform ausgewählt wurde, standen photogrammetrische Belange nicht im Mittelpunkt. Die Kamera wurde in erster Linie ausgewählt, da neben ihrem geringen Gewicht von 140 g (ohne Akku) für Canon-Kameras eine Schnittstelle zur Verfügung steht (sog. Canon Hack Development Kit (CHDK)). Mit dieser Schnittstelle kann direkt in die Funktionalität der Kamera eingegriffen werden. Damit wird über den Flugzeugrechner und somit ferngesteuert die Auslösung der Kamera kontrolliert.

Zwei Maßnahmen wurden ergriffen, um Bildunschärfen bei der Canon-Kamera zu vermeiden, wie sie bei Aufnahmen aus bewegten Trägern auftreten können. Dazu wurde zunächst das innerhalb der Optik befindliche Neutraldichtefilter (ND-, oder Graufilter) mittels der CHDK-Firmware deaktiviert, welches die Original-Kamerafirmware bei hellem Umgebungslicht standardmäßig aktiviert. Ein solches Filter reduziert die einfallende Lichtmenge und wird in der praktischen Photographie üblicherweise *vor* dem Objektiv angebracht, sodass es flexibel eingesetzt werden kann, um bei guter Beleuchtung längere Belichtungszeiten ohne übermäßiges Abblenden zu gestatten. Hier ist das ND-Filter hingegen eher als Ersatz für die fehlende Lamellenblende gedacht. Durch Deaktivierung des ND-Filters können bei gleicher Blende deutlich kürzere Belichtungszeiten realisiert werden.

Als zweite Maßnahme wurde die Bildstabilisierung wirksam deaktiviert: Vereinfacht dargestellt sieht das Prinzip der Stabilisierung vor, die Kamerabewegung mithilfe von Gyrosensoren zu registrieren und sofort mit einer entsprechenden Gegenbewegung einer verschiebbaren Linse und damit einer entsprechenden Ablenkung der Abbildungsstrahlen auszugleichen. Die weiche, bewegliche Lagerung der Bildstabilisator-Linse kann bei den teils heftigen Flugbewegungen zu zusätzlichen Bewegungsunschärfen im Bild führen. Bewegliche Teile innerhalb einer Optik verändern aber auch die innere Geometrie der Kamera und sind aus photogrammetrischer Sicht kritisch zu beurteilen (siehe Kamerakalibrierung). Daher wurde die beweglich-

che Linse durch Klebung an geeigneter Stelle im Innern des Objektivs physikalisch fixiert. Die Bildstabilisation nur in der Kamera-Firmware abzuschalten, würde indes nicht ausreichen, da dies nur den Stellmechanismus (Piezoelemente) deaktivieren würde und die Linse nach wie vor beweglich wäre. Die nachfolgenden Flüge wurden alle mit der derart modifizierten Canon Kamera durchgeführt. Die Modifikationen haben zu signifikanten Qualitätsverbesserungen geführt, was z.B. durch den Vergleich mit den Resultaten in CRAMER & HAALA (2012) ersichtlich ist.
Alternativ zur Canon wurde eine zweite Kamera beschafft, die speziell unter Berücksichtigung der Anforderungen für einen photogrammetrischen Bildflug ausgewählt wurde. Auswahlkriterien waren unter anderem eine hohe Auflösung des Bildsensors bei möglichst großem Bildformat, geringes Gesamtgewicht / Gesamtvolumen bei dennoch stabilem Gehäuse, keine bewegten Teile in der Optik (Bildstabilisierung oder Sensorreinigungssystem), gute (Festbrennweiten-)Optik bevorzugt mit manueller direkter mechanischer Fokussierung und eine kabelgebundene Auslösung des Verschlusses. Ausgewählt wurde letztlich die Kamera Ricoh GXR Mount A12. Die GXR-Modellreihe ist ein modular aufgebautes Kamerasystem mit verschiedenen erhältlichen Modulen, die auf den Kamerakörper (Speicher, Bedienteil, Display) geschoben werden. Bei den meisten Modulen sind Bildsensor und Optik in einem Bauteil fest verbunden. Dies entspricht vom Konzept her einer klassischen Messkamera mit möglichst fest vorgegebener Kamerageometrie. Das gewählte Modul Mount A12 entspricht jedoch nicht diesem Konzept, sondern verfügt als einziges über einen Anschluss für Wechselobjektive (Leica-M-Bajonett) und erlaubt so die Kombination mit hochwertigen Objektiven von Leica bzw. kompatiblen Objektiven (z.B. von Zeiss oder Voigtländer). In unserem Fall wurde ein Zeiss Biogon T* 21mm f/2.8 beschafft. Das nun zur Verfügung stehende Kamerasystem hat demnach folgende wesentliche Kenngrößen: Bildsensor CMOS, 4288 x 2848 Pix, Pixelgröße 5.5 x 5.5 µm², 23.6 x 15.7 mm² (APS-C), Bilddiagonale 2.83 cm, maximaler Bildwinkel ca. 68°. Damit ist die Bildgeometrie nahezu identisch mit der Canon IXUS 100, sodass für beide Kameras identische Bildflüge (identische Blockgeometrie) geplant werden können. Die jedoch im Vergleich zur Canon deutlich größere Pixelgröße und höhere Qualität der Optik sollte sich in Bezug auf Bildqualität und Qualität der Punktmessung bei nachfolgenden Auswertungen positiv bemerkbar machen.

3.1 Auflösungsuntersuchungen

Die Analyse des geometrischen Auflösungspotenzials durch Auswertung von entsprechenden Testmustern ist ein Standardverfahren zur Qualitätsbeurteilung von Kamerasystemen. Anhand von Auflösungsmustern können Modulationstransfer- und Punktverschmierungsfunktion berechnet werden, aus denen quantitative Aussagen zum geometrischen Auflösungsvermögen des Sensorsystems abgeleitet werden. Dieses Standardverfahren für die Beurteilung von optischen Systemen erlaubt einen ersten Qualitätsvergleich der Canon und der Ricoh/Zeiss Kamera. Zur Untersuchung werden verschiedene Bilder von Siemenssternen analysiert, die sowohl terrestrisch aufgenommen wurden als auch aus Bildflügen stammen. Terrestrische Aufnahmen erlauben exakt identische Aufnahmebedingungen und zeitgleiche Aufnahme des identischen Objekts. Für die späteren Projektbefliegungen wurde vor Ort ein Siemensstern im Gelände ausgelegt, der dann beim Überflug mehrfach in unterschiedlicher Lage bezogen auf den Bildraum abgebildet wurde. In beiden Szenarien wurde mit weitestgehend identischen Kameraeinstellungen gearbeitet, um die terrestrischen Ergebnisse mit den Befliegungsresultaten vergleichen zu können. Dies waren folgende Kameraeinstellungen: Canon: Blende (fest vorgegeben) f/3.2, Verschlusszeit 1/1500 s, ISO 80 (terrestrisch) bzw. ISO 100 (Bildflug); Ricoh/Zeiss: Blende (wählbar) f/5.6, Verschlusszeit 1/2000 s, ISO 250.

Die Tabelle 1 und Abbildung 2 zeigen Ergebnisse aus den Auflösungsuntersuchungen, angegeben für die drei Farbkanäle. σ PSF charakterisiert die Standardabweichung der Punktverschmierungsfunktion, MTF10 beschreibt die Ortsfrequenz (Linien pro pix), an der die Auflösungsgrenze erreicht wird. Beide Größen beziehen sich auf den Bildraum und werden in pix bzw. L/pix angegeben. Der GSD FWHM Wert überträgt das Auflösungsvermögen in den Objektraum, indem die Breite der Punktverschmierungsfunktion auf der halben Höhe des Maximums („Full-Width-Half-Maximum“, FWHM) mit dem Bildmaßstab multipliziert und damit in den Objektraum übertragen wird.
Die in Tabelle 1 angegebenen Werte für den Fall *terrestrisch* sind aus Bildern berechnet, die zeitnah aus nahezu gleicher Aufnahmedistanz von einem Stativ aus aufgenommen wurden, sodass von identischen Aufnahmebedingungen ausgegangen werden kann. Die Größe des Siemenssterns wurde so gewählt, dass seine Bildgröße vergleichbar dem Stern für die späteren Befliegungen ist. Die Kamera wurde direkt auf den Stern ausgerichtet, so dass die Auflösungsparameter sich auf die Bildmitte beziehen. Diese Werte können mit dem Fall *Bildflug – Bildmitte* verglichen werden. Die Unterschiede können durch den Einfluss der Atmosphäre (größere Aufnahmeentfernung zwischen Kamera und Objekt) und die Flugbewegung erklärt werden. Weiterhin ist zu beachten, dass die Canon- und die Ricoh/Zeiss-Befliegungen an unterschiedlichen Tagen in unterschiedlichen Monaten (März respek. August) durchgeführt wurden. Im Vergleich zu den Werten aus den Bildern *Bildflug – Bildrand* erkennt man den ausgeprägten Randabfall des Auflösungsvermögens von der Bildmitte zum Bildrand.
Die GSD FWHM Werte müssen in Relation zu der nominellen GSD betrachtet werden. Diese nominellen GSD-Werte sind für den Fall *terrestrisch* 0.18 cm, für den Fall *Bildflug – Bildmitte* 6.91 cm bzw. 4.79 cm für Canon bzw. Ricoh/Zeiss und für den Fall *Bildflug – Bildrand* 4.28 cm bzw. 4.82 cm für Canon bzw. Ricoh/Zeiss. Der Vergleich der nominellen GSD mit den anhand der Siemenssternuntersuchungen abgeleiteten Werte zeigt, dass die Ricoh/Zeiss Kamera für Objekte in der Bildmitte das nominelle Auflösungspotenzial sogar übererfüllt, wohingegen die Canon die nominelle GSD auch in der Bildmitte nicht ganz erreicht. Am Bildrand ist das Auflösungsvermögen für beide Kamerasysteme aber deutlich unter dem nominellen Wert.
Dieser Randabfall wird auch in der Abbildung 2 deutlich. Für diese Grafik wurden die MTF10 Werte aller Siemenssternauswertungen aus den Bildflügen Canon (Befliegung im März) und Ricoh/Zeiss (Befliegung im August) in der Bildebene eingezeichnet und dann die MTF10 flächig für die Bildebene interpoliert. Diese Interpolation ist möglich, da der Siemensstern während der Befliegung mehrfach überflogen wird und sich daher an verschiedenen Stellen bezogen auf die Bildebene abbildet. Für die Canon stehen 46 Siemenssternanalysen, für die Ricoh/Zeiss 28 Analysen zur Verfügung. Die Auswertungen der Canon konzentrieren sich auf den mittleren Bildbereich, da der Streifen, bei dem der Stern in der Bildmitte zu liegen kommt, mehrfach abgeflogen wurde. Auf diese Mehrfachbefliegung wurde bei der Ricoh/Zeiss-Kamera verzichtet. Insgesamt ergibt sich für die Ricoh/Zeiss ein homogeneres Auflösungsbild als für die Canon. Die inhomogenere Punktverteilung für den Canon-Datensatz ist aber zu berücksichtigen.
Alle hier vorgenommenen Auswertungen bestätigen die Annahme, dass die Ricoh/Zeiss-Kamera in ihrer Bildqualität besser als die Canon abschneidet. Dieses sollte sich vor allem auf die spätere Punktübertragung auswirken.

Tab. 1: Indikatoren für die geometrische Auflösung in Bild- und Objektraum.

	Canon			Ricoh/Zeiss		
	Rot	Grün	Blau	Rot	Grün	Blau
σ PSF [pix]						
Terrestrisch	0.476	0.499	0.418	0.330	0.333	0.333
Bildflug - Bildmitte	0.461	0.462	0.436	0.378	0.371	0.372
Bildflug - Bildrand	0.798	0.674	0.615	0.548	0.540	0.558
MTF10 [L/pix]						
Terrestrisch	0.717	0.684	0.817	1.035	1.027	1.025
Bildflug - Bildmitte	0.741	0.739	0.784	0.904	0.919	0.918
Bildflug - Bildrand	0.428	0.506	0.555	0.623	0.632	0.612
GSD FWHM [cm]						
Terrestrisch	0.207	0.216	0.181	0.142	0.143	0.144
Bildflug - Bildmitte	7.500	7.525	7.094	4.263	4.194	4.198
Bildflug - Bildrand	8.042	6.799	6.198	6.230	6.137	6.336

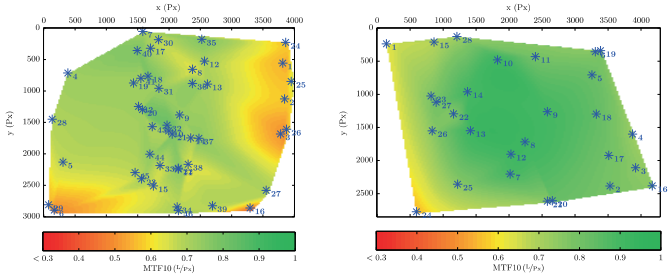


Abb. 2: Variation des geometrischen Auflösungsvermögens (MTF10) im Bildraum für Canon (links) bzw. Ricoh/Zeiss (rechts).

3.2 Kamerakalibrierung

Die Bestimmung der Kamerageometrie ist wesentlicher Bestandteil im Rahmen der photogrammetrischen Auswertung, zumal, wenn Kamerasysteme verwendet werden, die nicht primär für den photogrammetrischen Gebrauch konzipiert wurden, wie die im UAV@LGL-Projekt eingesetzten Systeme. Beide Kameras wurden mehrfachen Kalibrierungen unterzogen, die sowohl im Labor, als auch vor Ort, unmittelbar vor oder nach den Testflügen durchgeführt wurden. Alle Kalibrierungen wurden mithilfe von 2D-Testfeldern (z.T. mit Referenzmaßstab) bewerkstelligt, für die Ricoh/Zeiss erfolgte eine einzige Kalibrierung mit einem 3D-Testfeld. Die nachfolgende Tabelle 2 zeigt die Ergebnisse der verschiedenen Kalibrierläufe, Abbildung 3 visualisiert die Verzeichnungskurven für radiale bzw. tangentiale Verzeichnung aller Läufe. Alle Kalibrierungen erfolgten nach dem 8-Standpunkte-Verfahren (LUHMANN, 2012). Zum Einsatz kam das Australis Programm mit dem bekannten physikalischen Parametermodell nach Brown in der Australis Realisierung.

Für die Canon wurde ein 2D-Indoor-Testfeld verwendet. Die Kalibrierungen wurden nacheinander durchgeführt, wobei zwischen Kalibrierung (i) und (ii) die Kamera ausgeschaltet wurde, zwischen Kalibrierung (ii) und (iii) die Kamera aber eingeschaltet blieb. Damit kann die Stabilität der Kalibrierung nachgewiesen werden, da bei jedem Ein-/Ausschaltvorgang die Optik komplett in das Gehäuse eingefahren wird. Die Längenmessgenauigkeit (Standardabweichung der aus den ausgeglichenen Objektraum-Koordinaten berechneten Strecke des Maßstabs) kann als äußeres Genauigkeitsmerkmal der Kalibrierung herangezogen werden.

Die Ergebnisse in Tabelle 2 zeigen demnach, dass die Ricoh/Zeiss-Kamera eine höhere Objektraumgenauigkeit liefert. Die Genauigkeit der Bildpunktmessung (BPM) liegt für die Canon im Bereich von 1/10 pix. Das Ricoh/Zeiss-System schneidet auch in dieser Kategorie mit Werten unter 5% der Pixelgröße deutlich besser ab, wobei die 3D-Testfeldkalibrierung im Ergebnis schlechter ist. Die Ergebnisse in Tabelle 2 zeigen demnach, dass die Ricoh/Zeiss-Kamera eine höhere Objektraumgenauigkeit liefert. Insgesamt stimmen die Parameter der inneren Orientierung der drei Ricoh/Zeiss-Kalibrierungen nicht gut überein und sollten wegen der unterschiedlichen Testfelder auch nicht direkt verglichen werden. Die Kalibrierungen liegen zeitlich einige Monate auseinander. Ein Test wie bei der Canon, mit drei Kalibrierläufen am gleichen Testfeld unmittelbar hintereinander, wurde für die Ricoh/Zeiss nicht durchgeführt. Die innere und äußere Genauigkeit für die Ricoh/Zeiss-Kalibrierung ist bei der 2D-Laborkalibrierung am besten, für den 3D-Fall ist die interne Genauigkeit deutlich schlechter. Möglicherweise liegt das an der ungünstigeren Schnittkonfiguration im 3D-Fall mit nur 4 „rundum"-Aufnahmen und 8 verkanteten Aufnahmen bzw. an den aufgetretenen Überstrahlungen und Farbsäumen an den abgebildeten Retrotargets. Aufgrund der räumlichen Gegebenheiten konnte zudem für das 3D-Testfeld nur ein geringerer Objektabstand als im 2D-Fall eingehalten werden, weshalb im 3D-Fall unschärfere Aufnahmen resultierten. Es handelt sich jedenfalls nicht um ein prinzipielles Problem, wonach ein 3D-Testfeld schlechtere Ergebnisse als ein 2D-Testfeld liefert.

Um die möglichst aktuelle innere Geometrie der Kamera zu erfassen, wurde direkt vor jedem Bildflug eine Kalibrierung mit einem mobilen 2D-Testfeld vorgenommen. Bei den weiteren Auswertungen wurde dann immer diejenige Kalibrierung verwendet, die zeitlich am nächsten an den Bildflügen lag. Die Vor-Ort-Kalibrierung der Ricoh/Zeiss-Kamera ist in Tabelle 2 dokumentiert. Auch diese Parameter stimmen weder mit dem 2D- noch dem 3D-Fall überein, können aber auch nicht mit diesen verglichen werden, da Unterschiede möglicherweise bereits durch die etwas instabile Bauart des mobilen Testfelds und dessen zu kleine Punktmarken (bedingt durch den Objektabstand) zu erklären sind. Letztlich wurde noch nicht nachgewiesen, ob die unterschiedlichen Ergebnisse der drei Ricoh/Zeiss-Kalibrierungen auf derartige nicht-Kamera-bezogene Ursachen und ggf. auf Effekte aus den individuellen Kalibrierblockgeometrien zurückzuführen sind oder ob dafür physikalische Änderungen in der Kamerageometrie verantwortlich sind. Hierzu wären – wie bei der Canon geschehen – mehrere zeitnahe Kalibrierläufe an einem unveränderten Testfeld erforderlich.

Tab. 2: Resulate der Kamerakalibrierung für Canon (nach Modifikation) bzw. Ricoh/Zeiss.

	Canon			Ricoh/Zeiss		
	(i)	(ii)	(iii)	2D	3D	Vor-Ort
c (mm)	6.0248	6.0237	6.0246	21.6485	21.6408	21.6335
x_0 (mm)	0.1167	0.1167	0.1163	-0.1024	-0.0936	-0.1064
y_0 (mm)	-0.1832	-0.1834	-0.1823	-0.2305	-0.2438	-0.2200
K1	2.30e-3	2.35e-3	2.41e-3	1.13e-4	1.10e-4	1.08e-4
K2	-1.95e-5	-2.69e-5	-4.08e-5	-3.23e-7	-3.10e-7	-2.60e-7
K3	-5.42e-7	-1.67e-7	5.41e-7	2.73e-10	3.06e-10	6.31e-11
P1	-5.86e-4	-5.74e-4	-5.81e-4	1.10e-5	1.72e-5	2.85e-5
P2	8.72e-4	8.72e-4	8.71e-4	1.04e-5	8.31e-6	1.68e-5
B1	8.09e-5	9.69e-5	7.92e-5	-1.17e-4	-7.44e-5	9.91e-5
B2	4.15e-4	3.88e-4	3.65e-4	4.17e-5	5.48e-5	-1.37e-6
RMS BPM (µm)	0.17	0.14	0.14	0.26	0.56	0.27
RMS BPM (% Pixelgr.)	11.04	9.09	9.09	4.72	10.18	4.90
Längenabweichung (mm)	0.0648	0.0619	0.0598	0.0407	0.0360	-

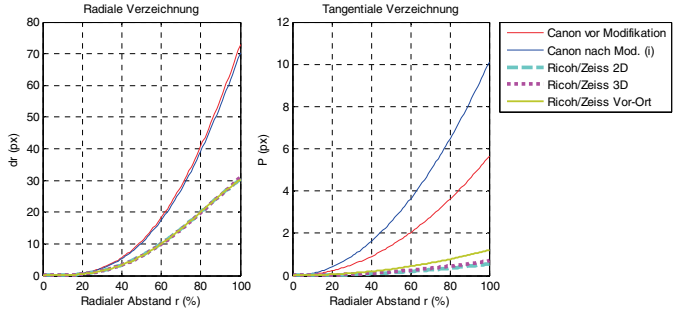


Abb. 3: Verzeichnungskurven aus verschiedenen Kalibrierläufen mit der Canon (vor und nach der Hardware-Modifikation) bzw. Ricoh/Zeiss.

4 Bildflüge Projektgebiet Hessigheim

4.1 Datenerfassung

Die empirischen Testflüge wurden im Gebiet der Hessigheimer Felsengärten durchgeführt, das Gebiet in dem das LGL auch seine Maßnahmen zur Flurneuordnung vorgenommen hat. Das Gebiet liegt unmittelbar am Neckar und zeichnet sich durch steile Weinberglagen aus. Für die UAV-Befliegungen wurde ein ca. 1000 m x 400 m großes Gebiet definiert, welches in einem klassischen photogrammetrischen Bildverband erfasst werden sollte. Dazu wurden insgesamt 5 Längsstreifen und zwei Querstreifen geplant mit hohen Überdeckungsverhältnissen von ca. 80% innerhalb des Streifens und ca. 70% zwischen benachbarten Streifen. Diese hohe Überdeckung dient einerseits dazu, die für einen UAV-Flug deutlich größeren Abweichungen vom Normalfall aufzufangen – vor allem bei geringeren Querüberdeckungen kann es sonst zu ungenügender Verknüpfung zwischen Nachbarstreifen kommen. Der hohe Überlapp ist aber auch für die spätere Generierung dichter Oberflächenmodelle hilfreich. Die Abbildung 4 zeigt die Flugplanung mittels IGIPlan (links). Anhand dieser GoogleEarth-Visualisierung kann auch die anspruchsvolle Geländetopographie erahnt werden. Die Flüge wurden derart geplant, dass eine nominelle Bodenpixelgröße (GSD) von 4.5 – 7.3 cm realisiert wird. Die starke Variation der GSD liegt an den Geländehöhenvariationen von ca. 100 m. Ebenfalls angedeutet in der Abbildung sind die Überdeckung der Flugstreifen im Gelände und die Lage der signalisierten Pass- und Kontrollpunkte. Insgesamt kann im Rahmen der weiteren Auswertungen auf 33 im Projektgebiet durch Farbmarkierungen signalisierte Punkte zugegriffen werden, die im Vorfeld der Flugkampagnen vom LGL vermarkt und mit GPS eingemessen worden sind. Die Objektpunktgenauigkeit dieser Referenzpunkte liegt bei etwa 1 cm (Std.Abw.).

Der rechte Teil der Abbildung 4 zeigt die Flugtrajektorie für den Ricoh/Zeiss-Test. Nach Übergabe vom Piloten an den Autopilot (für Start bzw. Landephase) fliegt das UAV die geplanten Fluglinien gemäß den vorgegebenen Wegpunkten aus der Flugplanung zuverlässig ab, wobei der Pilot am Boden jederzeit den Autopilot deaktivieren kann. Für den Autopilot werden als Wegpunkte nur die Anfangs- bzw. Endpunkte der photogrammetrischen Streifen in entsprechender Reihenfolge übernommen, die einzelnen Projektionszentren sind hier nicht von Bedeutung, da die Kamera während des Bildflugs in festen Zeitabständen ausgelöst wird (Auslöseintervall 1.5 s). Unter Berücksichtigung der üblichen Fluggeschwindigkeit von ca. 20 m/s führt das dann zu der gewünschten Überdeckung in Flugrichtung. Das Auslöseinter-

vall wird in der aktuellen Konfiguration nicht an die momentane Geschwindigkeit angepasst. Dies wirkt sich auf die realisierten Basislängen aus.
Die anschließende Abbildung 5 zeigt die Bildverbände Canon- (207 Bilder) bzw. Ricoh/Zeiss-Befliegung (190 Bilder). Die Bildüberdeckungen entsprechen weitestgehend den Vorgaben, wobei aufgrund des festen Auslöseintervalls und der nicht stabilisierten Kameraaufhängung Abweichungen von den nominellen Überdeckungsverhältnissen auftreten können. Die maximalen Abweichungen aus der Horizontalen variieren beim Canon Flug zwischen -18° und 12° für den Rollwinkel (omega), und zwischen -16° und 15° für den Nickwinkel (phi). Für die Ricoh/Zeiss Befliegung liegen die Variationen zwischen -15° und 12° für Roll- und zwischen -18° und 30° für den Nickwinkel. Bei Standardbildflügen mit bemannten Flugzeugen liegen diese Abweichungen i.d.R. unter ±5°. Diese starken Abweichungen vom Nadirfall ziehen gewisse Modifikationen für den späteren photogrammetrischen Auswerteprozess nach sich. Dies wird im nachfolgenden Abschnitt erläutert.

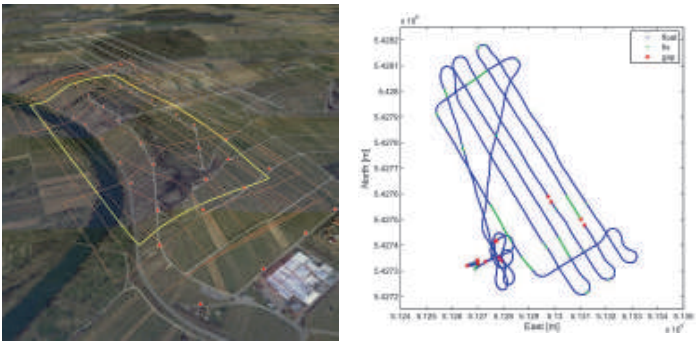


Abb. 4: Projektgebiet Hessigheim mit überlagerter Flugplanung und Streifenabdeckung (links) und die bei der Ricoh/Zeiss-Befliegung realisierte Flugtrajektorie aus der differenziellen GPS-Phasen-Auswertung (rechts).

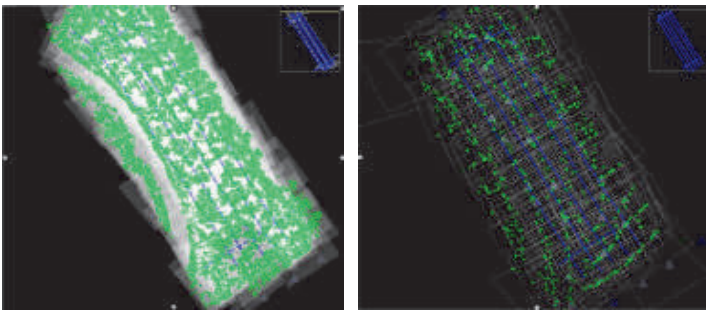


Abb. 5: Bildflüge Hessigheim: Blockgeometrie Canon Flug vom 23.03.2012 (links) und Ricoh/Zeiss Befliegung vom 23.08.2012 (rechts) mit automatischen Verknüpfungspunktmessungen aus Match-AT Auswertung.

4.2 Vorbereitung Aerotriangulation

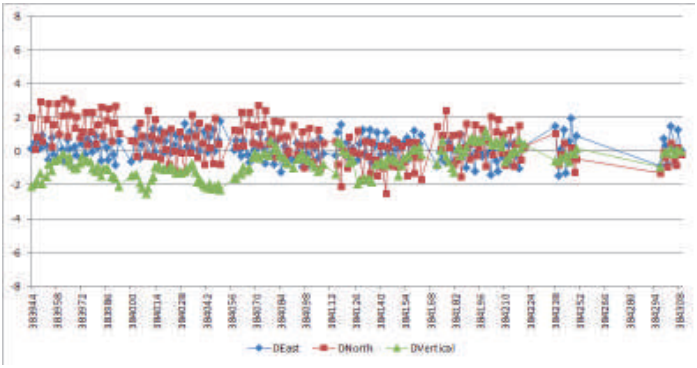
Die photogrammetrische Auswertung der UAV-Bildverbände zerfällt in mehrere Arbeitsschritte. Zunächst werden die im Vorfeld bestimmten Kalibrierparameter der Kamera in den Bildern angebracht (Korrektur Bildhauptpunktlage / Korrektur Objektivverzeichnung), sodass

„quasi verzeichnungsfreie“ Bilder ausgewertet werden. Über ein Structure-from-Motion (SfM) Verfahren werden die Bilder miteinander verknüpft und die Orientierung der Bilder bestimmt. Die Verknüpfung wird über SIFT-Punkte realisiert. Der SfM-Ansatz kommt ursprünglich aus der Computervision und erlaubt die Auswertung von im Prinzip beliebig angeordneten Bildverbänden. Ein Beispiel für einen derartigen SfM-Ansatz ist das Softwaresystem *Bundler*, das innerhalb des Photo Tourism Projektes entwickelt (SNAVELY et al., 2007) und erfolgreich zur automatischen Aerotriangulation von UAV-Bilddaten eingesetzt wurde (KIRCHGÄßNER et al., 2010). Die Auswertung der in diesem Artikel beschriebenen UAV-Bilddatensätze wurde durch die in ABDEL-WAHAB et al. (2011) beschriebene SfM-Implementierung realisiert, die auf einer Modifikation des Ansatzes von FARENZENA et al. (2009) basiert. Als Ergebnis des SfM-Verfahrens liegen die ausgeglichenen 3D-Koordinaten aller Verknüpfungspunkte und die äußere Orientierung aller Bilder in einem lokalen Koordinatensystem vor. Der Datumsübergang vom lokalen Koordinatensystem in das gewünschte Geländekoordinatensystem (hier UTM-Koordinaten) wird über die Projektionszentren aus der GPS-Bahnberechnung realisiert.

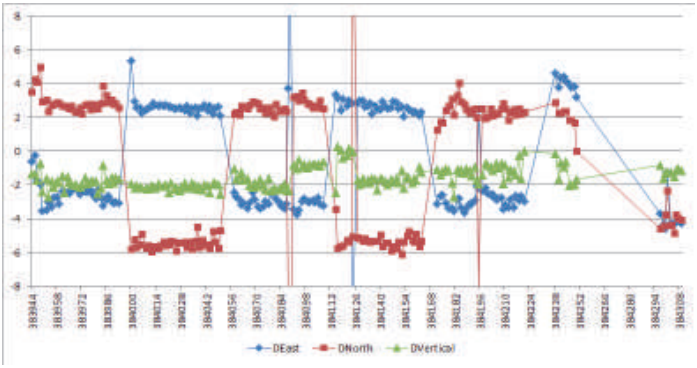
Die Flugführung des UAVs beruht auf den Echtzeit-GPS-Positionsinformationen. Diese werden für den hier verwendeten uBlox LEA-6T Empfänger herstellerseitig mit 2.0 – 2.5 m horizontaler Positionsgenauigkeit angegeben. Die photogrammetrische Auswertung ist aber zunächst nicht auf eine Echtzeitpositionierung angewiesen. Daher wurde für die spätere Aerotriangulation im Postprocessing mit der Software RTKLIB eine differenzielle Code- bzw. Phasenlösung berechnet. Die Vergleiche zwischen den Projektionszentrumskoordinaten aus der Aerotriangulation und den auf die Kameraauslösungen interpolierten GPS-Positionen erlaubt darüber hinaus eine Beurteilung der absoluten Genauigkeit der GPS-Positionierung. Hierbei muss allerdings beachtet werden, dass die in der Aerotriangulation geschätzten Projektionszentren u.U. durch Restsystematiken (z.B. Anteile aus der Kamerakalibrierung) verfälscht sein können, da sie sich mathematisch im Sinne einer optimalen Schätzung ergeben und somit nicht notwendigerweise die physikalischen Projektionszentren repräsentieren. Hinzu kommen ggf. Restfehler in der Zeitsynchronisation zwischen Kamera, Bordrechner und GPS.

In der Abbildung 4 (rechts) ist exemplarisch die Flugtrajektorie aus der Auswertung der differenziellen L1-Phasenmessung für den Ricoh/Zeiss-Flug am 23.08.2012 dargestellt. Die Referenzstation für die differenziellen Auswertungen befand sich an der Uni Stuttgart, ca. 24 km vom Befliegungsgebiet entfernt. Die Farbkodierung zeigt, dass die Mehrdeutigkeiten (Ambiguities) im Regelfall nicht zu Integer-Werten geschätzt werden konnten, offensichtlich im Wesentlichen dadurch verursacht, dass es sich nur um einen 1-Frequenzempfänger handelt. Nur in 12% der Epochen konnten die Mehrdeutigkeiten fixiert werden. Zudem erkennt man kurze Datenlücken, in denen keine GPS-Daten zur Verfügung standen. Für den gezeigten Flug sind es zwar nur 1.7% der Epochen, in denen keine GPS-Daten aufgezeichnet wurden, aber wie in der Abbildung 4 zu erkennen ist, treten diese Datenlücken auch während der photogrammetrischen Streifenflüge auf, was sich negativ auf die spätere Aerotriangulation auswirkt. Für den Ricoh/Zeiss-Flug fallen insgesamt 5 Bilder in derartige GPS-Datenlücken.

Die Abbildung 6 zeigt exemplarisch für den Flug vom 23.08.12 die Differenzen zwischen den unterschiedlichen GPS-Trajektorien. Gegenübergestellt sind im oberen Teil die Differenzen zwischen der GPS-Real-Time-Lösung, die auch für die Flugnavigation verwendet wird, und der differenziellen L1-Phasenlösung aus dem Post-processing. Im unteren Teil werden die Positionen aus der differenziellen L1-Phasenlösung mit den Projektionszentren aus der nachfolgenden Aerotriangulation verglichen. Für beide Fälle wurden die Differenzen nur an den 190 Kameraauslösezeitpunkten berechnet.



Positionsdifferenz [m]: Real-time GPS vs post-processing GPS (differenziell L1-Phase).



Positionsdifferenz [m]: Post-processing GPS (differenziell L1-Phase) vs Projektionszentren aus AT.

Abb. 6: Vergleich der GPS-Trajektorien für die Ricoh/Zeiss Befliegung vom 23.08.2012: Differenz zwischen GPS real-time und post-processing L1-Phasenlösung (oben) und Differenz zwischen GPS post-processing L1-Phasenlösung und geschätzten Projektionszentren aus der AT (unten).

Die Differenzen zwischen real-time und post-processing GPS-Lösung (Abbildung 6, oben) sind sehr konsistent. Die RMS-Werte betragen 0.71 m, 1.24 m, 1.03 m für Ost-, Nord- und Vertikalkomponente und sind damit deutlich besser als die Herstellerangaben zur Positionierungsgenauigkeit. Bei den Differenzen zwischen post-processing L1-Phasenlösung und den geschätzten Projektionszentrumskoordinaten aus der AT zeigt sich ein anderes Bild (Abbildung 6, unten). Hier sind in der Lagekomponente deutliche Systematiken zu erkennen, die klar flugstreifenabhängig sind. Außerdem erkennt man Ausreißer, die durch die GPS-Datenlücken verursacht und von der weiteren Betrachtung ausgeschlossen werden. Die mittlere Standardabweichung der Differenzen innerhalb der Flugstreifen liegt bei 0.52 m, 0.53 m und 0.48 m für Ost-, Nord- und Vertikalkomponente. Im Rahmen der RTKLIB-Prozessierung wird die interne Genauigkeit der differenziellen L1-Phasenlösung mit besser 20 cm abgeschätzt. Dieses Genauigkeitspotenzial kann aus obigen Differenzen in etwa nachgewiesen werden. Die Systematik muss aus der Photogrammetrie bzw. der Zeitsynchronisation zwischen den Komponenten kommen. Die Vorzeichen der Horizontaldifferenzen variieren je nach Flugrichtung der Streifen. Deutlich ist für die fünf Längsflugstreifen zu erkennen, dass

Streifen 2 und 4 und Streifen 1, 3 und 5 identische Systematik aufweisen. Berechnet man den jeweils mittleren horizontalen Positionsoffset dann ergeben sich für die Streifen 2 und 4 die Werte 6.07 m und 5.93 m bzw. für die Streifen 1, 3, und 5 die Werte 3.85 m, 4.03 m und 3.79 m. Ursache für diese Lageoffsets könnten Fehler in der Zeitsynchronisation sein, die Variation des Offsets von ca. 2 m könnte durch Variationen in der Fluggeschwindigkeit über Grund verursacht sein. Zum aktuellen Zeitpunkt ist die genaue Ursache für diese Effekte noch nicht bekannt, die GPS-Projektionszentren können aber durch entsprechende streifenabhängige Offset und ggf. Drift-Korrekturen in der Aerotriangulation verbessert werden.

5 Aerotriangulation

Die Aerotriangulation der beiden Bildverbände wurde mit der Software Match-AT von Trimble/inpho durchgeführt. Das Match-AT-Projekt wurde aufgesetzt mit den Resultaten aus der SfM-Auswertung. Dazu wurden für jedes Bild die Elemente der äußeren Orientierung aus SfM eingelesen, da Match-AT für jedes Bild eine Näherungsorientierung verlangt. Normalerweise reicht die Angabe von Projektionszentrumskoordinaten, z.B. aus GPS-Messungen, während für die Winkel vom Nadirfall ausgegangen wird, was in der UAV-Anwendung aber nicht mehr allgemein vorausgesetzt werden kann. Anschließend erfolgt die manuelle Messung aller signalisierten Punkte. Durch die hohe Überlappung zwischen den Bildern können bis zu 30-fach verknüpfte Punkte bestimmt werden. Anschließend müssen die automatischen Verknüpfungen hergestellt werden. Dazu wurden zunächst die bereits im SfM über SIFT durchgeführten Punktzuordnungen nach Match-AT importiert und eine passpunktbasierte Auswertung durchgeführt. Die Ergebnisse finden sich in Tabelle 3. Die Tabelle zeigt neben der Anzahl der verwendeten Pass- und Kontrollpunkte (PP/KP) und dem geschätzten Wert $\sigma 0$ die mittlere geschätzte Genauigkeit (Std.Abw., innere Genauigkeit) der Geländepunkte und die RMS-Werte aus den Differenzen an den Kontrollpunkten (äußere Genauigkeit). Das $\sigma 0$ der AT mit Verwendung der SfM-Verknüpfung liegt für beide Bildflüge im Bereich von knapp besser 1 pix, was sehr gut mit der zu erwartenden Genauigkeit der SIFT-Punktzuordnung korrespondiert. Wird die Punktübertragung mit Match-AT vorgenommen (Kombination von merkmalsbasierter und grauwertbasierter Zuordnung), so werden zwar deutlich weniger Verknüpfungspunkte gefunden, die Qualität dieser Zuordnung ist aber deutlich besser und liegt im Bereich von 1/3 pix. Dieses gilt sowohl für Canon- als auch Ricoh/Zeiss-Flug. Die höhere Qualität der Punktzuordnung wirkt sich auch positiv auf die innere und äußere Genauigkeit der Objektpunktbestimmung aus. Im Vergleich mit der nominellen Bodenpixelgröße von 4.5 – 7.3 cm (siehe oben) liegen die absoluten Genauigkeiten im deutlichen Sub-Pixel Bereich, wobei sich kaum Unterschiede in den Objektpunktgenauigkeiten aus Canon- bzw. Ricoh/Zeiss-Auswertung erkennen lassen. Dies ist insofern erstaunlich, als dass im Rahmen der Auflösungsuntersuchungen in Abschnitt 3.1 Qualitätsunterschiede beim Auflösungspotenzial des Ricoh/Zeiss-Systems im Vergleich zur Canon-Kamera nachgewiesen wurden. Bei den Vergleichen zwischen beiden Systemen muss aber berücksichtigt werden, dass die realisierte Match-AT-Verknüpfung für den Canon-Flug (Bildflug im Frühjahr, vor Einsetzen der Belaubung) deutlich besser ist, im Vergleich zur Befliegung mit der Ricoh/Zeiss (Bildflug bei voller Belaubung im Hochsommer). Der Einfluss der Vegetation ist auch aus Standard-Photogrammetrieanwendungen bekannt. Die SIFT-Punktzuordnung ist weniger anfällig gegenüber Störungen durch Vegetation, was sich an der in beiden Fällen deutlich größeren Anzahl von Verknüpfungspunkten zeigt. Für die Ricoh/Zeiss-AT wurden die SfM-Punkte auf 27367 Punkte ausgedünnt. Ursprünglich wurden 103787 Punkte im SfM erzeugt, wobei die eigentliche SIFT-Punktzuordnung gar nicht bis zu untersten Pyramidenstufe (max. Auflösung) fortgesetzt wurde, da sonst für den Ricoh/Zeiss-Flug deutlich über

300000 Punkte generiert worden wären. Dies zeigt die bessere radiometrische Qualität der Ricoh/Zeiss Bilder.
Für beide Bildverbände wurde auch die Rolle von Zusatzparametern in der AT untersucht. Zur Anwendung kamen die Polynommodelle von Ebner (12 Parameter) und Grün (44 Parameter), die in Match-AT implementiert sind. Für die Ricoh/Zeiss-Konfiguration zeigen sich im Objektraum kaum Veränderungen, die Absolutdifferenzen an den Kontrollpunkten scheinen eher etwas größer zu werden. Für die Canon ist eine leichte Verbesserung der RMS-Werte, vor allem in der Vertikalen zu erkennen.
Abschließend wurde für die Ricoh/Zeiss-Befliegung eine GPS-gestützte AT bei signifikant reduzierter Passpunktanzahl durchgeführt. Aufgrund der bereits in Abschnitt 4.2 nachgewiesenen streifenabhängigen Systematik wurden je Flugstreifen Offset und Driftkorrekturen für die GPS-Projektionszentren angesetzt. Für die 5 Bilder, die in GPS-Datenlücken (siehe Abb. 4, rechts) fallen, wurden keine direkt gemessenen GPS-Projektionszentren eingeführt. Die für die restlichen 185 GPS-Positionen in der AT geschätzten RMS-Werte nach Offset/Driftkorrektur betragen 0.404 m, 0.418 m und 0.459 m für Ost, Nord und Vertikalkoordinate. Diese Werte korrespondieren mit den RMS-Werten, die schon in Abschnitt 4.2 (siehe Abbildung 6) nachgewiesen wurden. Die erzielten Objektraumgenauigkeiten im Vergleich zur rein passpunktbasierten Auswertung mit vielen Passpunkten sind ca. um den Faktor 1.5 schlechter. Eine GPS-gestützte AT mit nur 4 Passpunkten in den Blockecken ist zwar möglich, bei der hier vorliegenden Qualität der GPS-Trajektorienbestimmung müssen aber Einbußen in der Genauigkeit hingenommen werden.
Für die Berechnung der dichten Oberflächenmodelle (Abschnitt 6) wurden die Orientierungselemente aus der rein passpunktbasierten AT unter Verwendung der Match-AT Verknüpfungspunktmessung ohne Zusatzparameter und ohne GPS verwendet (Version AAT-Pkte). Im Unterschied zu den in Tabelle 3 gezeigten Varianten sind jedoch alle 33 Punkte als Passpunkte eingeführt worden.

Tab. 3: Ergebnisse der Aerotriangulation mit Match-AT.

	PP/KP	σ0 [µm; pix]	Std.Abw. [m]			RMS [m]		
			Ost	Nord	Vert.	Ost	Nord	Vert.
Canon, 202 Bilder								
SfM-Pkte (Anz.: 165970)	22/11	1.0; 0.7	0.036	0.032	0.141	0.050	0.037	0.095
AAT-Pkte (Anz.: 8011)	22/11	0.4; 0.3	0.008	0.007	0.024	0.030	0.023	0.050
AAT-Pkte, SC 12	22/11	0.4; 0.3	0.007	0.007	0.022	0.024	0.021	0.044
AAT-Pkte, SC 44	22/11	0.4; 0.3	0.007	0.007	0.022	0.022	0.021	0.037
Ricoh/Zeiss, 190 Bilder								
SfM-Pkte (Anz.: 27367)	22/11	4.1; 0.7	0.034	0.030	0.110	0.031	0.037	0.058
AAT-Pkte (Anz.: 1439)	22/11	1.8; 0.3	0.018	0.015	0.049	0.029	0.024	0.043
AAT-Pkte, SC 12	22/11	1.8; 0.3	0.018	0.015	0.048	0.028	0.028	0.051
AAT-Pkte, SC 44	22/11	1.8; 0.3	0.017	0.015	0.048	0.027	0.029	0.051
AAT-Pkte, GPS, Off/Drif	4/29	1.8; 0.3	0.024	0.021	0.062	0.040	0.074	0.083

6 Generierung dichter Oberflächenmodelle

Für die Erzeugung dichter 3D-Punktwolken und Oberflächenmodelle wurde ein Ansatz basierend auf dem Semi-Global Matching Verfahren von HIRSCHMÜLLER (2008) genutzt. Dieses Verfahren führt eine Stereozuordnung für jedes einzelne Pixel durch und erzeugt somit sehr dichte 3D-Punktwolken. Darüber hinaus ermöglichen die sich stark überlappenden Bildverbände eine kombinierte pixelweise Zuordnung mehrerer sich überdeckender Stereobildpaare.

Somit stehen pro Pixel mehrere redundante Stereomessungen zur Verfügung, was eine erhebliche Steigerung von Genauigkeit, Vollständigkeit und Zuverlässigkeit der daraus abgeleiteten 3D-Punktwolke ermöglicht (HAALA, 2011). Aus Bildern großformatiger digitaler Luftbildkamerasysteme können so 3D-Punktwolken mit einer Genauigkeit im Subpixelbereich gewonnen werden (HAALA & ROTHERMEL, 2012). Analog zu diesen Untersuchungen wurde zur Auswertung der UAV-Aufnahmen das am Institut für Photogrammetrie entwickelte System SURE (ROTHERMEL et al., 2012) eingesetzt. Dabei wurden aufgrund der ähnlichen Blockkonfiguration für die Aufnahmen mit der Canon IXUS 100 und der Ricoh/Zeiss beide Datensätze mit identischer Parametrisierung für die Auswahl der Stereopaare, die dichte Zuordnung und die Triangulation prozessiert.

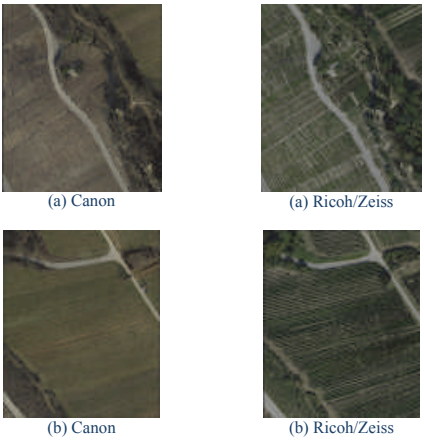
Üblicherweise ermöglichen Befliegungen mit großformatigen Luftbildkamerasystemen vor allem in Kombination mit einer stabilisierten Plattform eine gute Übereinstimmung zwischen geplanter und tatsächlich erfasster Bildüberdeckung. Im Gegensatz dazu verursacht die relativ große Bewegungsdynamik einer UAV-Plattform häufig eine vergleichsweise heterogene Überdeckung der erfassten Bilder. Um dennoch eine automatisierte Auswahl optimaler Bildpaare für die dichte Zuordnung zu ermöglichen, wurde ein Initialisierungsschritt realisiert. Hierfür wurden zunächst alle Bilder des Blocks auf 1/256 der ursprünglichen Auflösung reduziert. Anschließend wurden Bildpaare gebildet, wenn der Abstand der Projektionszentren kleiner als 150 m war und daraus durch Stereozuordnung 3D-Punkte bestimmt. Diese Punktwolken wurden kombiniert und auf einen Punktabstand von 3 m ausgedünnt. Basierend auf dieser groben Punktwolke wurden dann Stereopaare mit ausreichender Überlappung und gutem Schnittwinkel der Bildstrahlen ausgewählt und die Stereozuordnung in der höchsten Auflösung wiederholt.

Für die Ricoh/Zeiss-Befliegung ergaben sich aus dieser Initialisierung für jedes der 190 Bilder bis zu 24 Stereomodelle. Insgesamt wurden so 2003 Bildpaare generiert und zur Oberflächenrekonstruktion genutzt. Analog dazu ergaben sich für den Canon-Bildverband mit 207 Bildern insgesamt 2170 Stereopaare. Da aufgrund der großen Überdeckung für jedes Pixel eines Basisbildes mehrere mögliche Stereopartner zur Verfügung standen, war eine hoch redundante Bestimmung des zugehörigen 3D-Punktes möglich. Dabei wurden fehlerhafte Parallaxenschätzungen eliminiert und das Rauschen der erzeugten 3D-Punktwolke verbessert. Letztlich wurde ein 3D-Objektpunkt nur dann generiert, wenn er aus mindestens zwei Stereomodellen konsistent bestimmt werden konnte. Die Berechnungszeiten zur Generierung des gerasterten Oberflächenmodells der Canon-Befliegung betrug insgesamt 21 h 54 min. Verwendet wurde ein Desktop PC ausgestattet mit einen i7 3.4 GHz Prozessor. Analog wurde für die Prozessierung des Ricoh/Zeiss-Datensatzes 18 h 45 min benötigt.

Aus diesen sehr dichten 3D-Punktwolken wurden dann die in Abbildung 7 dargestellten Digitalen Oberflächenmodelle (DOM) abgeleitet, wobei als Gitterweite des resultierenden Rasters die mittlere GSD der Flüge von 6 cm gewählt wurde. In der Regel standen pro Rasterzelle mehrere 3D-Punkte aus der dichten Stereozuordnung zur Verfügung. Um eine ausreichende Zuverlässigkeit des Gitters zu gewährleisten, wurde Zellen mit weniger als drei Punkten kein Höhenwert zugewiesen. Für alle anderen Zellen wurde aus den darin liegenden 3D-Punkten der Median gewählt, um sowohl die Höhe als auch den Farbwert zu bestimmen.



Abb. 7: Oberflächenmodelle des Testgebiets aus Befliegung Canon (oben) und Ricoh/Zeiss (unten).



(a) Canon (a) Ricoh/Zeiss

(b) Canon (b) Ricoh/Zeiss

Abb. 8: Testregionen zur Vollständigkeitsanalyse: (a) Gebiet mit komplexer Oberflächengeometrie; (b) ebenes Testgebiet.

In einem blockzentralen Testgebiet komplexer Oberflächengeometrie (Abbildung 8(a)) konnten für den Ricoh/Zeiss-Datensatz Höhenwerte für 95.3% der Rasterzellen bestimmt werden. Für dasselbe Gebiet konnte für die Canon-Befliegung eine Abdeckung von 93.7% der Rasterzellen generiert werden. Grund für geringfügig schlechtere Werte sind Bereiche kleiner Strukturen wie Äste, für welche die Rekonstruktion in einigen Bereichen fehlschlug. Des Weiteren macht sich die schlechtere Radiometrie des Canon-Systems in Schattenbereichen

bemerkbar, insgesamt weist dieser Bereich des Testgebiets mit einigen circa 20 m hohen Felsstrukturen eine relativ komplexe Oberflächengeometrie auf. Die Vollständigkeit des DOMs wird am Blockrand maßgeblich durch die Oberflächengeometrie beeinflusst. Für eine ebene Testregion (Abbildung 8(b)) am Rand des Blocks konnten für die Canon-Befliegung 98.6% der Höhenwerte rekonstruiert werden. Analog konnten für die Ricoh/Zeiss-Befliegung 95.6% Vollständigkeit erreicht werden. Wie schon bei der Aerotriangulation (siehe Abschnitt 5) ist die geringere Punktdichte vor allem auf die Belaubung der Rebflächen zurückzuführen. Diese waren bei der Ricoh/Zeiss-Befliegung im August belaubt, so dass speziell am Blockrand durch Abschattungen eine geringere Anzahl von Beobachtungen möglich war. Für den Canon-Flug im März war ein größerer Anteil vom Boden sichtbar, für den in der Regel eine zuverlässige 3D-Erfassung erfolgte. Der Effekt der Belaubung ist auch in Abbildung 9 sichtbar, in der die farbkodierten Differenzbeträge zwischen den beiden erfassten DOM dargestellt sind. Größere Differenzen ergeben sich nur in Bereichen der Weinberge, der Bäume und einzelner Felder, für die im August die Oberfläche der Vegetation und im März die Geländeoberfläche erfasst wurde. In Bereichen, die keiner Änderung durch Bewuchs unterworfen waren, liegen die Differenzen in der Größenordnung von 10 cm. Generell führen Abschattungen und Mehrdeutigkeiten durch Vegetation zu einer leicht reduzierten Dichte des Oberflächenmodells, wie dies auch für Bereiche mit Gestrüpp und Bäumen sichtbar ist.

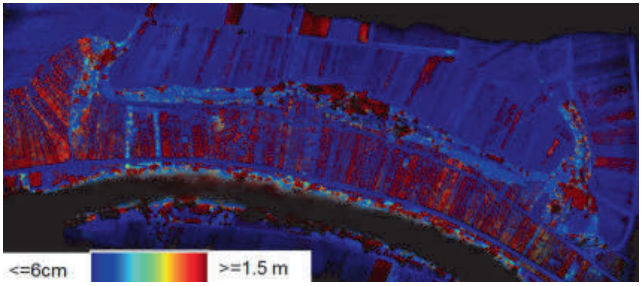


Abb. 9: Differenz der berechneten DOM-Daten Canon – Ricoh/Zeiss.

Zur Bestimmung des Rauschens der rekonstruierten Oberfläche wurden ebene Regionen genutzt. Hierfür wurden die in Abbildung 10 dargestellten 30 Gebiete auf zwei Straßen manuell definiert. Aus den darin befindlichen Rasterpunkten wurde je eine Ebene durch Kleinste-Quadrate-Schätzung bestimmt, wobei der orthogonale Punkt-zu-Ebenen-Abstand minimiert wurde. Aufgrund des variierenden Bildmaßstabs wurden Gebiete von unterer und oberer Straße getrennt untersucht. Für den Canon-Flug ergab sich für die obere Straße mit einer GSD von 6 cm eine Standardabweichung von 30 mm. Für die untere Straße mit einer GSD von 8 cm betrug die Standardabweichung 42 mm. Für den Ricoh/Zeiss-Datensatz ergaben sich für die obere Straße 20 mm und die untere Straße 22 mm. Zur Eliminierung potenzieller Ausreißer wurden alle Punkte mit Abständen 3σ eliminiert und Standardabweichungen aus den verbleibenden Punkten erneut berechnet. Dabei ergaben sich für den Canon-Flug 29 mm und 40 mm für die obere und untere Straße, bei der Ricoh/Zeiss-Befliegung lagen die Werte bei 19 mm und 21 mm.

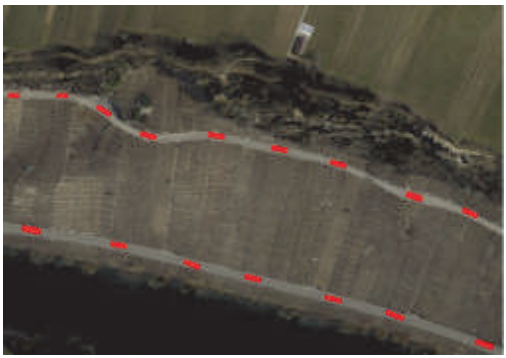


Abb. 10: Ebene Testbereiche zur Bestimmung des Punktrauschens.

Generell führt die Stereozuordnung mittels SGM in Bereichen mit geringer Textur, wie sie in den 30 Testregionen vorlag, eher zu einer Glättung der rekonstruierten Oberfläche, so dass die Werte mit stark variierenden Höhen ein etwas stärkeres Messrauschen aufweisen dürften. Dennoch ist insbesondere die Standardabweichung der Höhendaten für den Ricoh/Zeiss-Flug im Bereich von ½ GSD durchaus beeindruckend und für sehr viele Anwendungen mehr als ausreichend.

7 Zusammenfassung

Im Rahmen des Pilotprojekts UAV@LGL wurde erstmals im Rahmen deutscher Vermessungsverwaltungen die Verwendung von UAVs für Anwendungen der Landesvermessung untersucht. Die erzielten Ergebnisse sind sehr erfolgversprechend. Es wurde der Nachweis erbracht, dass auch mit relativ einfachen Kamera- und Trägersystemen 3D-Objektpunktbestimmungen im Sub-Pixelbereich (absolute Genauigkeit) möglich sind. Für die Oberflächenmodellgenerierung konnten Höhengenauigkeiten im Bereich von ½ Pixel nachgewiesen werden.
Die Bildorientierung wurde mit Standard-AT-Programmen durchgeführt. Lediglich die vor-Ort-Kamerakalibrierung und Vororientierung bedingen einen gewissen Mehraufwand. Für diese sind zusätzliche Software bzw. andere Ansätze notwendig (SfM-Verfahren). Es ist jedoch davon auszugehen, dass klassische AT-Programme in naher Zukunft zumindest um diese SfM-Verfahren ergänzt werden. Die dichte Punktwolkengenerierung wurde ebenfalls sehr zeitnah in photogrammetrische Arbeitsabläufe integriert.
Limitierender Faktor ist derzeit jedoch das noch (zeit-)aufwändige Verfahren, um eine Aufstiegserlaubnis zu erlangen. Für die hier durchgeführten Bildflüge hat es fast 2 Monate gedauert, bis die Aufstiegserlaubnis unter Auflagen erteilt wurde. Allerdings sind auf verschiedenen nationalen und internationalen Ebenen intensive Bestrebungen im Gange, die UAVs als Flugzeuge in den Luftraum zu integrieren. Exemplarisch sollen hier die Initiativen im Rahmen der EU (z.B. Commission Staff Working Document „Towards a European strategy for the development of civil applications of Remotely Piloted Aircraft Systems (RPAS)“, (EU, 2012)) oder die Arbeiten im Rahmen von JARUS (JARUS, 2012) erwähnt werden. Das Bundesministerium für Verkehr, Bau und Stadtentwicklung hat im Juni 2012 gemeinsame Grundsätze des Bundes und der Länder für die Erteilung der Erlaubnis zum Aufstieg von unbemannten Luftfahrtsystemen veröffentlicht (BMVBS, 2012). All das wird dazu führen,

dass sich zukünftig UAV-Anwendungen auch aus genehmigungstechnischer Sicht deutlich einfacher durchführen werden lassen.

Danksagung

Unser besonderer Dank gilt dem LGL Baden-Württemberg für die Finanzierung dieser Studie. Die LGL-Mitarbeiter Manfred Gültlinger und Reinhold Hummel haben uns in der praktischen Durchführung der Studie jederzeit bestmöglich unterstützt. Dieses soll besonders hervorgehoben werden. Dank gilt auch den Kollegen vom iFR, namentlich den Herren Martin Trittler und Florian Weimer. Ohne deren aktive Beteiligung hätten die UAV-Flüge nicht durchgeführt werden können. Herr Maximilian Ehrle stand als fachkundiger Modellflieger für die Durchführung der UAV-Befliegungen zur Verfügung. Weiterhin soll Frau Ye Zhou dankend erwähnt werden, die im Rahmen ihrer Diplomarbeit die Prozessierung der GPS-Flugtrajektorien vorgenommen hat. Sie wurde dabei von Herrn Bernhardt Schäfer vom Institut für Navigation (INS) unterstützt. Auch diesen beiden sei für ihre Kooperation gedankt.

Literatur

ABDEL-WAHAB, M., WENZEL, K. & FRITSCH D. (2011): Reconstruction of Orientation and Geometry from large Unordered Datasets for Low Cost Applications. Proceedings LC3D Workshop, Berlin, Dezember 2011.

CRAMER, M. & HAALA, N. (2012): Genauigkeitspotential der photogrammetrischen Bildauswertung für Daten unbemannter Luftfahrzeuge, in Tagungsband 32. DGPF-Jahrestagung, Potsdam, 14.-17. März 2012, Seiten 428-440.

FARENZENA, M., FUSIELLO, A. & GHERARDI, R. (2009): Structure and motion pipeline on a hierarchical cluster tree, ICCV Workshop on 3-D Digital Imaging and Modeling, pp. 1489–1496.

HAALA, N. & ROTHERMEL, M. (2012): Dense Multi-Stereo Matching for High Quality Digital Elevation Models in PFG Photogrammetrie, Fernerkundung, Geoinformation, Jahrgang 2012 Heft 4 (2012), p. 331-343.

HAALA, N. (2011): Multiray Photogrammetry and Dense Image Matching. Photogrammetric Week '11, Ed. D. Fritsch, Wichmann, Berlin/Offenbach, pp. 185-195.

HAALA, N., CRAMER, M., WEIMER, F. & TRITTLER, M. (2011): Performance Test on UAV-based data collection Proceedings of the International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g). IAPRS, Volume XXXVIII-1/C22, Zurich, Switzerland, 2011, 6p.

HIRSCHMÜLLER, H. (2008): Stereo Processing by Semi-Global Matching and Mutual Information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30 (2), pp. 328 - 341.

KIRCHGÄßNER, U., PUTZE, U., VON SCHÖNERMARK, M. & HAALA, N. (2010): Anforderungen an die Auswertung UAV-gestützter Fernerkundungsdaten DGPF Tagungsband 19/2010 – Dreiländertagung OVG, DGPF und SGPF pp. 597 – 605

LGL (2012): Projektbeschreibung UAV@LGL-Projekt, internes Papier, 2012.

LUHMANN, T. (2010): Nahbereichsphotogrammetrie, Grundlagen, Methoden und Anwendungen, 3. Auflage, Wichmann Verlag.

SNAVELY, N., SEITZ, S. & SZELISKI. R. (2007): Modeling the World from Internet Photo Collections. International Journal of Computer Vision, 2007.

M. ROTHERMEL, M., WENZEL, K., FRITSCH, D. & HAALA, N. (2012): SURE: Photogrammetric Surface Reconstruction from Imagery Proceedings LC3D Workshop, Berlin, Dezember 2012.

Online Quellen

BMVBS (2012): Informationen aus Internetpräsenz, siehe unter http://www.luftverkehr.de/index.php?option=com_content&view=article&catid=63:brd-bekanntmachungen-nfl-1&id=2109:c2012-1-161

BORMATEC (2012): Informationen aus Internetpräsenz, siehe unter http://bormatec.com/index.php?option=com_content&view=article&id=78&Itemid=59&lang=en , letzter Zugriff 20.12.2012.

EU (2012): Informationen aus Internetpräsenz, siehe unter http://register.consilium.europa.eu/pdf/en/12/st13/st13438.en12.pdf, letzter Zugriff 20.12.2012.

IFR (2012): Informationen aus Internetpräsenz, siehe unter http://www.ifr.uni-stuttgart.de/webseiten/computersystem/computersystem.html, letzter Zugriff 20.12.2012.

JARUS (2012): Informationen aus Internetpräsenz, siehe unter http://www.jarus-rpas.org/, letzter Zugriff 20.12.2012.

Bilddatenerfassung mit einem Gyrocopter – Erste Erfahrungen zur „Photogrammetrie nach Bedarf“

DIETER FRITSCH[1], ALBRECHT GRIMM[2] , JENS KREMER[3]
MATHIAS ROTHERMEL[4] & KONRAD WENZEL[5]

Zusammenfassung: Die Bilddatenerfassung nach Bedarf (Photogrammetry on Demand, PoD) hat sich in jüngster Zeit zu einem spannenden Thema entwickelt, nicht zuletzt durch den Einsatz von Unmanned Aerial Vehicles (UAVs). Obgleich die photogrammetrische Bilddatenerfassung mit UAVs eine Lösungsmöglichkeit für den ad-hoc Bedarf an Orthophotos und Oberflächenmodellen zu sein scheint, sind speziell in Deutschland häufig föderale bürokratische Hürden zur Erlangung der Fluggenehmigung zu überwinden.
Um dennoch der Nachfrage nach Produkten der Photogrammetrie zeitnah zu entsprechen, wird in diesem Beitrag ein neues System vorgestellt und in einem Test erprobt. Erste Erfahrungen durch die Auswertungen des Bildflugs Hessigheim, Baden-Württemberg zeigen, dass ein bemanntes Gyrocopter-System einer „Photogrammetrie nach Bedarf“ in idealer Weise entspricht. Das Fluggerät wird von der Firma AutoGyro, Hildesheim für die zivile Nutzung (Rundflüge, Luftbildphotographie, Einsatz Landwirtschaft, usw.) angeboten und ist für das Hessigheim-Projekt von der Firma IGI mbH mit dem Kamerasystem DigiCAM-50 ausgestattet worden. Der Bildflug wurde in Anlehnung an die klassische Luftbildphotogrammetrie mit IGIPlan vorab geplant und auf das CCNS-5 Flugführungs- und Managementsystem des Gyrocopters übertragen. Der Bildflug war unproblematisch und konnte in einer durchschnittlichen Flughöhe von 470 m ü.G. und in nur 2 Stunden zwischen Aufsteigen und Landen auf dem nahegelegenen Flugplatz Backnang durchgeführt werden.
Da die Hessigheim-Befliegung im Rahmen eines Pilotprojekts UAV@LGL BW zwischen dem Landesamt für Geoinformation und Landentwicklung (LGL BW) und dem Institut für Photogrammetrie der Universität Stuttgart (ifp) durchgeführt werden konnte, bei dem es in erster Linie um die Erprobung von zwei UAV-Systemen für Zwecke der Landesvermessung ging, war der Gyrocopter-Flug eine ideale Ergänzung. Dessen Bildauswertung durch Match-AT sowie durch das neuentwickelte Programmpaket SURE des ifp war ohne Schwierigkeiten zu bewerkstelligen. Als Produkte stehen neben hochaufgelösten 3D-Punktwolken auch Oberflächenmodelle und Orthophotos zur Verfügung. Interessant sind insbesondere die Genauigkeitsanalysen zwischen Gyrocopter- und UAV-Datenauswertungen im Bereich der Aerotriangulation wie auch der erzeugten Oberflächenmodelle. Eine Zusammenfassung mit Ausblick rundet den Beitrag ab.

1) Prof. Dr.-Ing. Dieter Fritsch, Universität Stuttgart, Institut für Photogrammetrie (ifp), Geschwister-Scholl-Str. 24D, 70174 Stuttgart, Tel. +49-711-685-83386, E-Mail: dieter.fritsch@ifp.uni-stuttgart.de
2) Prof. Albrecht Grimm, IGI mbH, Langenauer Str. 46, 57223 Kreuztal, Tel. +49-2732-55250, E-Mail: a.grimm@igi-systems.com
3) Dr. Jens Kremer, IGI mbH, Langenauer Str. 46, 57223 Kreuztal, Tel. +49-2732-55250, E-Mail: j.kremer@igi-systems.com
4) Dipl.-Ing. Mathias Rothermel, Universität Stuttgart, Institut für Photogrammetrie (ifp), Geschwister-Scholl-Str. 24D, 70174 Stuttgart, Tel. +49-711-685-84094, E-Mail: mathias.rothermel@ifp.uni-stuttgart.de
5) Dipl.-Ing. Konrad Wenzel, Universität Stuttgart, Institut für Photogrammetrie (ifp), Geschwister-Scholl-Str. 24D, 70174 Stuttgart, Tel. +49-711-685-83384, E-Mail: konrad.wenzel@ifp.uni-stuttgart.de

1 Einleitung

Die Bilddatenerfassung nach Bedarf (Photogrammetry on Demand, PoD) hat sich in kürzester Zeit zu einem spannenden Thema entwickelt, nicht zuletzt stimuliert durch den Einsatz von Unmanned Aerial Vehicles (UAVs). Diese als „Festflügel-Systeme“ oder „Rotor-Systeme“ in verschiedenen Anwendungen erprobten Fluggeräte scheinen in der Praxis angekommen zu sein (CRAMER et al., 2013). Obgleich die photogrammetrische Bilddatenerfassung mit UAVs eine Lösungsmöglichkeit für den ad-hoc Bedarf an Orthophotos und Oberflächenmodellen zu sein scheint, sind speziell in Deutschland häufig Hürden zur Erlangung der Fluggenehmigung zu überwinden. Deren Regelung wird in den Bundesländern unterschiedlich gehandhabt, was den Einsatz von UAVs grundsätzlich erschwert.

Im Rahmen eines Pilotprojekts zwischen den Partnern Landesamt für Geoinformation und Landentwicklung Baden Württemberg (LGL BW) und dem Institut für Photogrammetrie der Universität Stuttgart (ifp) war in einer umfassenden Machbarkeitsstudie der Einsatz von UAVs für Zwecke der Flurbereinigung zu prüfen. Grundsätzlich ging es dabei um die Methodik der Datenerfassung mit UAVs, einer Systembeschreibung der Fluggeräte sowie der eingesetzten Hard- und Software, sowie der Beschreibung des gesamten Workflows mit Arbeitsaufwand für die einzelnen Bearbeitungsschritte von der Flugplanung bis hin zum Endprodukt. Darüber hinaus waren prinzipielle Darstellungen des Datenflusses mit Schnittstellen zu der im LGL BW eingesetzten Standard-Software als auch ein Vergleich unterschiedlicher Flugsysteme (Flächenflugzeug oder Xfach-Copter) mit Aussagen zur Leistungsfähigkeit und Praxistauglichkeit bei gleichzeitiger Validierung der Ergebnisse auf der Grundlage von Referenzdaten durchzuführen. Als konkrete Ergebnisse wurden erwartet: Orientierte Luftbilder mit einer Bodenauflösung von mindestens 10 cm, digitale Orthophotos mit einer Bodenauflösung von 10 cm bei einer Lagegenauigkeit von 1 pixel, 3D-Modelle mit einer Gitterweite von 1 m mit integrierten Strukturelementen (Bruchkanten) und einer Höhen- und Lagegenauigkeit von $\leq$ 10 cm. Ferner waren Luftbildschrägaufnahmen aus vordefinierten Positionen und Aufnahmerichtungen sowie die 3D-Modellierung markanter oder touristischer Objekte (Landmarken) mit Visualisierungen, 3D-Animationen und Videos erwünscht.

Der o.a. Auszug aus dem Lastenheft erforderte eine sehr detaillierte und methodische Vorgehensweise des ifp mit umfassenden Untersuchungen zur grundsätzlichen Verwendung von UAVs sowie weiteren, am Markt verfügbaren Plattformen zur photogrammetrischen Datenerfassung. Im Rahmen einer Zusammenarbeit mit dem Institut für Flugmechanik und Flugregelung (iFR) der Universität Stuttgart konnten zwei UAV-Festflügelsysteme mit Consumer Still Videokameras und GPS ausgestattet und erfolgreich über das Testgebiet Hessigheim geflogen werden. Eine ausführliche Beschreibung der Datenerfassung und -auswertung ist mit dem im Tagungsband verfügbaren Beitrag CRAMER et al. (2013) gegeben. Infolge der föderalen Regelungen zur Erteilung einer Flugerlaubnis für UAVs hat sich herausgestellt, dass es nicht so leicht ist, diese Fluggenehmigung zu erhalten. Hier wurde mit dem Regierungspräsidium Stuttgart über Wochen hinweg verhandelt. Parallel dazu hat das ifp ein weiteres Projekt in Rottenburg (bei Tübingen) bearbeitet, wo ein UAV OctoCopter eingesetzt werden sollte. Auch hier hat die Erteilung der Fluggenehmigung durch das Regierungspräsidium Tübingen einige Wochen in Anspruch genommen.

Um diese bürokratischen Hürden zum Einsatz von UAV-Systemen zu überwinden und der Idee einer Photogrammetrie nach Bedarf sehr zeitnah zu entsprechen, konnte kurzfristig auf ein weiteres Fluggerät zurückgegriffen werden. Im Rahmen der langjährigen Zusammenarbeit zwischen der IGI mbH, Kreuztal und dem ifp wurde beschlossen, einen bemannten Gyrocopter mit IGI Photogrammetrie-Ausrüstung einzusetzen. Hier kam uns der auch die Fa. LutzAir zu Hilfe, ein kleines Unternehmen, welches mit seinem Gyrocopter Rundflüge auf Flugshows organisiert. Der Betrieb eines Tragschraubers unterliegt den generellen Regelungen des Luftfahrtbundesamtes und ist nicht durch föderale Vorgaben für UAVs beeinträchtig. Konkret bedeutet dies, dass ein solches bemanntes Fluggerät jederzeit in Abstimmung mit der Flugsicherung eines benachbarten Flughafens von einem Flugplatz in der Nähe des zu überfliegenden Gebiets aufsteigen darf.

Ein Gyrocopter gehört zur Familie der Tragschrauber, das sind nicht schwebeflugfähige Drehflügler, die eine Art Mittelding zwischen Hubschrauber und Flugzeug darstellen. Im Gegensatz zu einem Flugzeug sind die Tragflächen nur rudimentär vorhanden oder fehlen ganz (http://www.enzyklo.de/lokal/40024). Tragschrauber sind interessant für Anwendungen mit geringen Geschwindigkeiten. Geringes Gewicht, leichte Erlernbarkeit sowie mäßige Anschaffungs- und Betriebskosten zeichnen diese Gruppe von Fluggeräten aus.

Erste Erfahrungen durch die Auswertungen des Bildflugs Hessigheim zeigen, dass das bemannte Gyrocopter-System einer „Photogrammetrie nach Bedarf“ in idealer Weise entspricht. Der von LutzAir betriebene Tragschrauber ist vom Typ AutoGyro MTOsport. Das Fluggerät wurde mit einem IGI Datenerfassungssystem bestehend aus der Kamera *DigiCAM-50* mit 50 mm Objektiv, einem GNSS/IMU System vom Typ *AEROcontrol* und dem *CCNS-5* Flugführungs- und Managementsystem ausgerüstet. Die Ausrüstung wurde an die Stelle des ausgebauten zweiten Sitzes montiert. Die Kamera blickte links am Flugzeugkörper entlang.



Abb. 1: Tragschrauber mit *DigiCAM-50* Kamerasystem

Die ausführliche Beschreibung der Bilddatenerfassung ist Gegenstand von Kapitel 2. Nach der Bilddatenerfassung wurden die Bilder durch klassische Aerotriangulation trianguliert (Abschnitt 3) und der dichten Bildzuordnung zugeführt (Kapitel 4). Im darauffolgenden Abschnitt werden Genauigkeitsanalysen für DOMs zwischen dem Gyrocopter- und einem UAV-Flug ausgeführt.

Das Projekt UAV@LGL BW ist somit erfolgreich abgeschlossen, jedoch führt das ifp noch weitere Untersuchungen speziell im bebauten Teil des Gebiets durch, die sich insbesondere mit der manuellen, halbautomatischen und vollautomatischen 3D-Rekonstruktion von Gebäuden und Gebäudeensembles auseinandersetzen (siehe Schluss und Ausblick).

2 Flugplanung und -durchführung

Die *DigiCAM-50* von IGI besitzt ein 50 Megapixel Flächen-CCD mit einer Größe von 37x49 mm² - bei einer Pixelgröße von 6 µm - und liefert die aus der Luftbildphotogrammetrie bekannte Bildqualität. Der Bildflug Hessigheim wurde klassisch mit IGIplan vorab geplant und auf das *CCNS-5* des Gyrocopters übertragen. Im Hinblick auf eine Bilddatenauswertung durch dichte Bildzuordnungs-Algorithmen des ifp wurde für die Längsüberdeckung 80% und die Querüberdeckung 60% gewählt. Der Bildflug war unproblematisch und konnte in einer durchschnittlichen Flughöhe von 470 m ü.G. und in nur 2 Stunden zwischen Aufsteigen und Landen auf dem nahegelegenen Flugplatz Backnang durchgeführt werden. Die *DigiCAM-50* mit 50 mm Objektiv lieferte eine GSD von ca. 5 cm und entsprach dem zuvor erwähnten Lastenheft.
Der in Abb. 2 dargestellte Testblock wurde am 23. August 2012 zeitgleich mit einem UAV-System (Fluggerät Maja der Firma Bormatec mit einer Ricoh GXR Mount A12) über der Gemeinde Hessigheim im Landkreis Ludwigsburg aufgenommen. Er umfasst ein Gebiet von etwa 3 Quadratkilometern und beinhaltet neben dem Ort Hessigheim auch landwirtschaftlich genutztes Gebiet – das UAV-Untersuchungsgebiet - und eine Schleife des Neckars. Der *DigiCAM* Block besteht aus drei Flugstreifen mit je 33 Bildern, einer Vorwärtsüberdeckung von 80 %, einer seitlichen Überdeckung von 60 % und einem Pixelabstand am Boden zwischen 6.3 und 6.8 cm. Um in der Auswertung der Bilder verschiedene Optionen zu haben, wurden auch drei kurze Querstreifen hinzugefügt. Im Gebiet treten Höhenunterschiede von etwa 100 Metern auf. Die Bilder wurden mit Blende 4.5 und einer Belichtungszeit von 1/800 sec aus der o.a. Höhe ü.G. aufgenommen. Die Fluggeschwindigkeit betrug etwa 50 Knoten (ca. 26 m/s). Der Testblock mit seinen drei Längs- und drei Querstreifen wurde in etwa 15 min. aufgezeichnet. Vor der weiteren Verarbeitung der Bilder wurden die von *AEROcontrol* direkt gemessenen Orientierungen und Bildpositionen durch eine sogenannte Integrierte Sensor-Orientierung optimiert. Dabei wurden zehn lokale Passpunkte verwendet. Nach der Aufnahme des Testblocks wurde noch ein weiterer photogrammetrischer Block mit 19 Flugstreifen und 1093 Bildern aufgenommen. Der gesamte Flug inklusive des zweiten Blocks dauerte etwa 125 min.
Der An- und Abtransport des Fluggeräts erfolgte durch einen Pkw-Anhänger und war daher einfach durchzuführen. Während des Starts und der Landung war ein Flugleiter am Flugplatz Backnang verfügbar.
Die in der Machbarkeitsstudie festgelegte Gebietsausdehnung befindet sich nordwestlich der Ortschaft Hessigheim (siehe Abb. 2). Nur für diesen Bereich konnten die beiden UAV-Bildflüge durchgeführt werden, da ein Überflug über bebautem Gebiet nicht genehmigt worden war. Die für diesen Beitrag durchgeführten Genauigkeitsanalysen und -vergleiche können sich daher auch nur auf Abschnitte in offenem Gelände beziehen.
Der Gyrocopterflug direkt über der Ortschaft erlaubt daher auch eine 3D-Rekonstruktion von Gebäuden und Gebäudeensembles, auf das noch im Schluss und Ausblick näher eingegangen wird. Damit soll einmal mehr demonstriert werden, dass der Einsatz eines bemannten Fluggeräts in Form eines Tragschraubers keinerlei Restriktionen unterworfen ist und daher alle Freiheitsgrade bietet, um hochwertige photogrammetrische Produkte und Geoinformationen bereitzustellen. Grundsätzlich kann der Einsatz eines solchen Fluggeräts für alle Arten photogrammetrischer Datenerfassung, auch des Laser-Scannings, empfohlen werden. Die

Betriebskosten und damit Flugkosten sind gering bei gewohnt garantierter Qualität der Datenerfassung aus der Luft. Hinzu kommt, dass infolge der geringen Flughöhen die Anforderungen an die Wetterbedingungen geringer sein können als bei einem klassischen Bildflug.
Ein erster Genauigkeitsvergleich mit den Ergebnissen der Aerotriangulation zwischen den beiden UAV-Flügen sowie dem Gyrocopter-Flug wird im nächsten Kapitel durchgeführt. Dabei zeichnet sich der Gyrocopter-Flug mit seiner gewohnt klassischen Blockgeometrie durch eine einfache Auswertestrategie aus.



a) b)

Abb. 2: a) Flugplanung auf Basis eines digitalen Geländemodells mit dem Programm *IGIplan*
b) Darstellung der tatsächlichen Auslösepositionen in Google Earth

3 Aerotriangulation

Die Aerotriangulation des Bildverbands wurde mit der am ifp verfügbaren Software Match-AT von Trimble/inpho durchgeführt. Infolge der Standardblockgeometrie waren keinerlei Vorarbeiten notwendig, um den *DigiCAM-50* Block mit seinen insgesamt 131 Bildern zu georeferenzieren. Für die in diesem Testprojekt vorhandene und für Auswertungen der UAV-Datensätze zwingend erforderliche umfangreiche Passpunktkonfiguration hätte die Angabe der Projektionszentrumskoordinaten durch die mitregistrierten GPS-Messungen ausgereicht. Trotzdem wurden Lage und Orientierung durch die von *AEROcontrol* gelieferte integrierte Sensororientierung berücksichtigt, weil diese für den anderen, im gleichen Flug aufgenommenen Block notwendig waren und deshalb mit erfasst wurden. Die Ergebnisse sind im fettumrandeten Teil von Tabelle 1 ausgewiesen. Dabei wurde eine Selbstkalibration mit den bei Match-AT verfügbaren 44 Parametern angewandt. Die Standardabweichungen liegen bei ca. 5 mm in der Lage und 20 mm in der Höhe. An den Kontrollpunkten ergaben sich mittlere Abweichungen von ca. 14 mm in der Lage und 24 mm in der Höhe.

Die Vorarbeiten zur Aerotriangulation der beiden UAV-Flüge sind in CRAMER et al. (2013) detailliert dargestellt. Aus Gründen der Vergleichbarkeit der Ergebnisse der Aerotriangulation werden diese synoptisch in Tabelle 1 integriert.

Tab. 1: Ergebnisse der Aerotriangulation mit Match-AT und Vergleich

	PP/KP	σ0	Std.Abw. [m]			RMS [m]		
		[µm; ix]	Ost	Nord	Vert.	Ost	Nord	Vert.
Canon 202 Bilder AAT-Pkte, keine SC	22/11	0.4; 0.26	0.008	0.008	0.024	0.030	0.023	0.050
AAT-Pkte, SC 44	22/11	0.4; 0.3	0.007	0.007	0.022	0.022	0.021	0.037
Ricoh/Zeiss 190 Bilder AAT-Pkte, keine SC	22/11	1.8; 0.3	0.018	0.015	0.049	0.029	0.024	0.043
AAT-Pkte, SC 12	22/11	1.8; 0.3	0.018	0.015	0.048	0.028	0.028	0.051
DigiCAM-50 131 Bilder AAT-Pkte, SC 44	50/17	0.6; 0.1	0.006	0.005	0.020	0.014	0.013	0.024

Wie in Tabelle 1 nachgewiesen, erreicht die AAT für den Gyrocopter-Flug mit *DigiCAM-50* sehr gute Ergebnisse, die auch als Referenz für die beiden UAV-Flüge gelten können. Insgesamt zeigt dieser Vergleich, dass alle drei eingesetzten Systeme die Vorgaben des in der Einleitung kurz skizzierten Lastenhefts erfüllen.

4 DOM-Ableitung

Digitale Höhenmodelle wurden mit dem am ifp implementierten Softwarepacket *SURE* (ROTHERMEL et al., 2012) erzeugt. Neben einem auf SGM-basierten Verfahren zur dichten Bildzuordnung umfasst dieses Paket eine effiziente Fusion von redundant geschätzten Parallaxen mehrerer Stereomodelle. Alle Datensätze wurden dabei mit identischen Parametrisierungen für die Epipolarbildgenerierung, dichte Zuordnung und Fusion prozessiert.

Die Selektion der zu berechnenden Stereomodelle wurde wie folgt realisiert. Für jedes Bild der *DigiCAM-50* Befliegung wurde eine dichte Zuordnung mit jeweils sechs benachbarten Bildern ausgeführt. Dabei ist jedes Bild mit den zwei vorhergehenden und zwei nachfolgenden Bildern desselben Flugstreifens und den direkt benachbarten Bildern der angrenzenden Flugstreifen zugeordnet worden. Das mittelformatige *DigiCAM-50* Kamerasystem in Kombination mit der vergleichsweisen stabilen Trägerplattform des Gyrocopters resultiert in gleichmäßiger Abdeckung des Geländes und strukturierter Überlappung der Footprints. Sowohl die Instabilität einer UAV-Plattform hinsichtlich Lage und Orientierung als auch ein geringerer Bildmaßstab bei

stark variierenden Geländehöhen resultieren in vergleichbar unstrukturierten Überlappungen der Footprints. Dadurch wird die manuelle Auswahl geeigneter Bildpaare für die dichte Zuordnung eines UAV-Datensatzes erheblich erschwert. Aufgrund dessen wurde ein zusätzlicher Initialisierungsschritt zur automatischen Selektion implementiert. Die Bilder des kompletten UAV-Projekts *Ricoh/Zeiss* wurden auf 1/256 der originalen Auflösung reduziert. Nachfolgend sind alle Bildpaare, für welche der Abstand der Projektionszentren kleiner 150 m bemaß, rektifiziert, zugeordnet und trianguliert worden. Dieser Prozess kann infolge der geringen Auflösung sehr schnell rechentechnisch umgesetzt werden. So erhaltene Punktwolken wurden zusammengeführt und auf einen Punktabstand von 3 m ausgedünnt. Basierend auf diesen Oberflächenpunkten wurde die Überlappung aller Bildpaare berechnet und nur Bildpaare mit einer Überlappung grösser als 40 % für die endgültige DOM-Generierung herangezogen.
Aufgrund hoher Überdeckung ist die Tiefeninformation zwischen Bildpaaren redundant bestimmt. Diese Redundanz kann genutzt werden, um fehlerhafte Parallaxenschätzungen zu filtern und das Rauschverhalten der Oberflächenrekonstruktion zu verbessern (HAALA & ROTHERMEL, 2012). Für beide Datensätze wurden redundante Parallaxenschätzungen gemäß ROTHERMEL et al. (2012) fusioniert und die zugehörige Punktwolke extrahiert. Dabei wurde gefordert, dass redundante Parallaxen in mindestens zwei Stereomodellen räumlich konsistent geschätzt werden konnten, anderenfalls waren die Beobachtungen zu eliminieren. Die so erzeugten Punktwolken dienten als Grundlage für die Extraktion des finalen Oberflächenmodels. Hierfür wurden die erzeugten Punkte jeweils einer Zelle eines 2-dimensionalen, horizontalen Gitters zugewiesen. Als Gitterweite wurde die mittlere GSD des *Ricoh/Zeiss* Datensatzes von 6 cm gewählt. Finale Höhenwerte einer Zelle wurden dann als Median aller Z-Komponenten einer Zelle bestimmt. Analog wurden die zugehörigen RGB-Werte berechnet. Höhenwerte von Zellen, welchen weniger als 3 Messungen zugeordnet werden konnten, wurden als nicht rekonstruierbar gewertet.



Abb. 3: DOM Punktwolke der *Ricoh/Zeiss* Befliegung

Für die *Ricoh/Zeiss* Befliegung (CRAMER et al., 2013) wurden insgesamt 2003 Bildpaare zur Oberflächenrekonstruktion herangezogen. Dabei wurde für jedes der 190 Bilder bis zu 24 Stereomodelle berechnet. Für den *DigiAM-50* Bildverband mit 131 Bildern wurden 255 Stereopaare ausgewertet. Dabei ist zu beachten, dass die rekonstruierte Fläche des Gyrocopter-

Datensatzes um ca. Faktor 6 grösser ist. Abb. 3 und 4 zeigen die resultierenden Oberflächenmodelle.



Abb. 4: DOM Punktwolke der *DigiCAM-50* Befliegung, Bereich des UAV-DOMs mit rot markiert

5 Genauigkeitsanalysen

Um die Qualität der generierten DOMs zu analysieren, wurden basierend auf 30 im Testgebiet gelegen ebenen Bereichen (siehe Abb. 5) das Rauschen der Oberfläche untersucht. Für die Rasterpunkte der jeweiligen Bereiche konnten ausgleichende Ebenen geschätzt werden. Daraufhin wurden Standardabweichungen der Punkt-Ebenen-Residuen berechnet. Um etwaige Ausreißer zu eliminieren, sind alle Punkte mit Residuen >3σ eliminiert, Ebenen erneut geschätzt und die zugehörigen Standardabweichungen berechnet worden. Nachfolgend wurden die so erhaltenen σ-Werte der Testbereiche für „obere“ und „untere Straße“ gemittelt. Die endgültigen Standardabweichungen für den UAV-Datensatz *Ricoh/Zeiss* betrugen 19 mm für die obere Straße und 21 mm für die untere Straße. Analog lieferte die Genauigkeitsanalyse des Gyrocopter-Flugs Werte in einem ähnlichen Genauigkeitsbereich (19 mm obere Straße bzw. 17 mm untere Straße). In Bereichen komplexerer Geometrie muss generell von einem stärkeren Oberflächenrauschen ausgegangen werden.

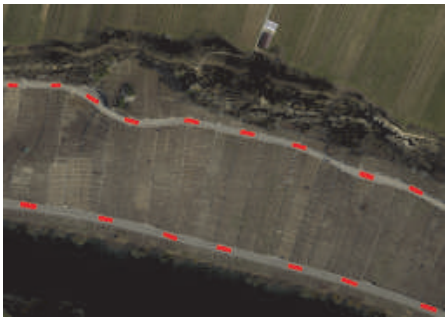


Abb. 5: Ebene Testgebiete entlang der Straßen für die Evaluation des Oberflächenrauschens



Abb. 6: True Orthophotos eines zentralen Ausschnittes des Weinbaugebiets. Größere Vollständigkeit des UAV *Ricoh/Zeiss* DOM (links) gegenüber Gyrocopter *DigiCAM-50* DOM (rechts) aufgrund einer höheren Querüberlappung vor allem in Bereichen der Abbruchkante und Vegetation.

Des Weiteren wurde in einem zentralen Teil des Testgebiets (Abb.6) die Vollständigkeit der generierten Oberfläche untersucht. Für dieses Gebiet konnten für die *Ricoh/Zeiss* Befliegung 95.2% aller Höhenwerte rekonstruiert, für die Befliegung mit der *DigiCam-50* hingegen nur 82.9% der Höhenwerte ermittelt werden. Grund für die geringere Vollständigkeit ist die geringere Querüberlappung des Datensatzes. Dies macht sich vor allem in Bereichen großer Sichtschatten (ca. 20 m hoher Felsabbruch) bemerkbar, ferner wurde eine geringere Abdeckung in Vegetationsbereichen festgestellt. In diesen Gebieten sind im Allgemeinen weniger korrekte Parallaxenschätzungen der dichten Zuordnung zu erwarten und die Forderung der Mindestanzahl konsistenter Beobachtungen häufig nicht erfüllt. Durch Abschwächung der Konsistenzbedingung könnte jedoch eine gesteigerte Vollständigkeit mit nur geringfügiger Zunahme der Ausreißer erzielt werden.



Abb 7: Visualisierung der absoluten Differenz im Testgebiet (links), Vermaschte Oberflächenmodelle des Gyrocopter-Datensatzes (Mitte) und des UAV-Datensatzes *Ricoh/Zeiss* (rechts)

In Abb. 7 (linkes Bild) ist die absolute Differenz der beiden DOM visualisiert. Die größten Abweichungen treten vor allem in Bereichen der belaubten Rebstöcke auf. Für den *Ricoh/Zeiss* Datensatz konnten sehr feine Strukturen (z.B. schmale Rebstöcke) nur hinreichend rekonstruiert

werden. Wie in den vermaschten DOMs in Abb. 5 zu erkennen, konnten für den *DigiCAM-50* Datensatz bessere Ergebnisse erzielt werden. Grund hierfür dürfte ein vorteilhafteres Signal-zu-Rauschverhältnis sein.

Vergleichbar geringe Fluggeschwindigkeiten des Gyrocopter-Systems ermöglichen eine Datenaufnahme mit hoher Bodenauflösung. Dies und die großzügige Überlappung resultiert in einer bemerkenswerten Qualität der erzeugten DOM. Abb. 8 zeigt rekonstruierte Oberflächen eines bebauten Gebiets. Dachkanten der Häuser werden im Allgemeinen sauber extrahiert und auch kleine Strukturen wie Kamine oder Masten von Straßenbeleuchtung können korrekt rekonstruiert werden.



Abb 8: Visualisierung des Gyrocopter DOM in bebautem Gebiet

6 Schluss und Ausblick

Der Einsatz eines Gyrocopters mit einer *DigiCAM-50* von IGI hat die Erwartungen mehr als übertroffen. Die Datenerfassung ist unproblematisch, da das Fluggerät einfach im Pkw-Anhänger zu einem nahegelegenen Flugplatz gebracht werden kann. Die beim Einsatz von UAV-Systemen bekannten Schwierigkeiten der Beantragung einer Fluggenehmigung entfallen.
Bei der Datenauswertung können Standardprogramme zur Aerotriangulation eingesetzt werden. Es bedarf keiner aufwendigen Vorprozessierung durch SfM-Algorithmen wie bei UAV-Systemen. Zusammengefasst scheint ein Tragschraubersystem – wie hier in Form eines Gyrocopters nachgewiesen – eine ideale Plattform für eine „Photogrammetrie nach Bedarf (Photogrammetry on Demand)“ zu sein.
Die zu Beginn des Beitrags zitierte Machbarkeitsstudie ist mit dem Vergleich der zwei UAV-Systeme und dem Gyrocopter-System abgeschlossen. Grundsätzlich hat sich dabei der Einsatz der getesteten Systeme für bestimmte, vorwiegend kleinräumige Zwecke als geeignet herausgestellt – die erzielten Ergebnisse sprechen für sich. Aus operationellen Gründen zeigen sich gewisse Einschränkungen beim praktischen Einsatz der UAVs. So ist zur Orientierung der großen Zahl unregelmäßig angeordneter Bilder eine ausreichende Anzahl von Passpunkten und eine relativ aufwendige Vorauswahl der Aufnahmen notwendig. Weiterhin stellen die zeitaufwendige Beschaffung der Fluggenehmigung sowie die Einschränkungen beim Überflug bebauter Gebiete eine Limitierung der praktischen Einsetzbarkeit dar.

Diese Schwierigkeiten treten bei Verwendung eines Gyrocopters nicht auf. Mit dem gleichen System können zusätzlich sogar auch großflächigere Projekte realisiert werden.
Das Konzept einer „Photogrammetry on Demand“ kann mit dem hier verwendeten Gyrocopter-basierten System optimal umgesetzt werden.
Als weitere Schritte wird das ifp Teile der Ortschaft Hessigheim aus den hochqualitativen *DigiCAM-50* DOMs manuell, halbautomatisch und vollautomatisch in ein virtuelles Stadtmodell transformieren. Hierzu müssen vor Ort noch die Texturen der Gebäudefassaden erfasst werden, um anspruchsvolle LoD2-Modelle wie auch Visualisierungen zu ermöglichen. Das geringe Rauschen zeigt bereits schon jetzt, dass solche Folgeprodukte mit hoher Qualität aus den *DigiCAM-50* Daten abgeleitet werden können.

7 Danksagung

Unser besonderer Dank gilt dem LGL Baden-Württemberg (vertreten durch seinen Präsidenten Hansjörg Schönherr) für die Beauftragung und Finanzierung dieser Machbarkeitsstudie. Die LGL-Mitarbeiter Manfred Gültlinger und Reinhold Hummel haben uns in der praktischen Durchführung der Studie jederzeit best möglichst unterstützt. Dank gebührt ebenso Herrn Lutz Weigand, der mit seinem Gyrocopter ad hoc den Hessigheim-Flug kurzfristig und kostenlos durchgeführt hat.

8 Literatur

ABDEL-WAHAB, M., WENZEL, K. & FRITSCH, D., (2011): Reconstruction of Orientation and Geometry from large Unordered Datasets for Low Cost Applications. Proceedings LC3D Workshop, Berlin, Dezember 2011.

CRAMER, M., HAALA, N., ROTHERMEL, M., LEINSS, B & FRITSCH, D., (2012): UAV-gestützte Datenerfassung für Anwendungen der Landesvermessung. In Tagungsband 33. DGPF-Jahrestagung/Dreiländertagung, Freiburg, 27.-28. Februar 2013

HAALA, N. & ROTHERMEL, M., (2012): Dense Multi-Stereo Matching for High Quality Digital Elevation Models. PFG Photogrammetrie, Fernerkundung, Geoinformation, Jahrgang 2012, Heft 4, p. 331-343.

ROTHERMEL, M., WENZEL, K., FRITSCH, D. & HAALA, N., (2012): SURE - Photogrammetric Surface Reconstruction from Imagery. Proceedings LC3D Workshop, Berlin, Dezember 2012.

Online-Referenzen
(http://www.enzyklo.de/lokal/40024).

Co-registration of Time-of-Flight (TOF) camera generated 3d point clouds and thermal infrared images (IR)

LUDWIG HOEGNER [1], MARTIN WEINMANN [2], BORIS JUTZI [2], STEFAN HINZ [2] & UWE STILLA [1]

Abstract: This article presents an investigation on the co-registration of a 3d point cloud and intensity image of a near-infrared (NIR) time-of-flight (TOF) camera system and a thermal infrared (TIR) camera. The TIR-camera is relatively oriented to the TOF-camera. As the radiometric behavior of the TIR-camera and the NIR-image of the TOF-camera is different for most pixels, corners are detected in both images and co-registered. The intensity values of TIR-camera are projected into the TOF point cloud and from there in the TOF-camera image plane. The quality of the projection is evaluated comparing the projected TIR image from the view of the TOF-camera and the TOF NIR image.

Zusammenfassung: Gegenstand dieses Artikels ist die Koregistrierung einer 3d Punktwolke und eines Nah-Infrarot (NIR) Bildes einer Time-of-Flight (TOF) Kamera mit einer thermischen Infrarot (TIR) Kamera. Die TIR Kamera wird relativ zur TOF Kamera orientiert. Da sich die radiometrischen Eigenschaften der TIR Kamera und des NIR Bildes der TOF Kamera für die meisten Pixel unterscheiden, werden Ecken in beiden Bildern detektiert und koregistriert. Die Intensitätswerte der TIR Kamera werden in die Punktwolke der TOF Kamera projiziert und von dort in die Bildebene der TOF Kamera. Die Qualität der Projektion wird durch Überlagerung des in die TOF Kamera projizierten TIR Bildes und des NIR Bildes der TOF Kamera überprüft.

1 Motivation

Deriving an appropriate 3D description of man-made and natural environments is of great interest in Computer Vision, Photogrammetry and Remote Sensing. Most of the current approaches are based on the use of image and/or range data (HARTLEY & ZISSERMAN, 2008). Sets of images and image sequences can be used to simultaneously calculate the relative orientation of the images and generate 3d points clouds (POLLEFEYS et al, 2008; MAYER, 2007). Image-based methods of deriving 3d point clouds allow the observation and reconstruction of moving objects and include color information for the 3d points. These methods are nevertheless limited to textures scenes where corresponding homologues image points can be found. Range measuring sensors like laserscanners overcome this limitation but are in general not capable of capturing moving objects. The combination of laserscanners and a camera allows the generation of a colored 3d point cloud.
Simultaneously capturing intensity information of high quality as well as range information by images with a single measurement, new active time-of-flight (TOF) sensors seem to be well-suited for combining the advantages of range measurement and images (WEINMANN et al., 2011). However, the acquired intensity typically represents information of the visual domain and hence, only radiometric and geometric surface properties of observed objects are captured

1) Photogrammetrie & Fernerkundung,Technische Universität München (TUM)
Arcisstraße 21, 80333 München, www.pf.bv.tum.de
2) Institut für Photogrammetrie und Fernerkundung, Karlsruhe Institute of Technology (KIT)
Englerstr. 7, 76131 Karlsruhe, http://www.ipf.kit.edu

which might not always be sufficient. In contrast to this, infrared (IR) cameras sense thermal radiation in the infrared spectrum which is emitted by objects in the scene and not visible in visual images. Thus, the captured images offer a different look on objects and the extraction of additional information like temperature and material of observed objects. Especially in building observation (IWASZCZUK et al., 2011; HOEGNER et al., 2007), the extraction of textures on facades of buildings allows a reconstruction of the surface temperature and a look into the interior behavior of a wall. Due to different radiometric behavior in different bands, well-known features points like SIFT are not applicable for the co-registration of thermal infrared and visible band images, but a combination of these different bands can enhance features in both kinds of images and even reveal new features that might not be present in either IR or visual images (CHEN & LEUNG, 2009; BAI et al., 2011). The registration of such image data representing information in different spectral bands has for instance been investigated using a segment based approach (COIRAS et al., 2000) or an approach involving normalized mutual information (PARK et al., 2008). This includes that an image transformation is required for mapping the IR texture onto intensity images. For almost planar scenes, the transformation model of a homography was investigated (WEINMANN et al., 2012). Another solution for co-registration is the use of ground control points, that can be identified in both images that are to be matched. In case of a TOF camera, the ground control points are directly related to image coordinates of the TOF camera and with this coordinates directly to 3d points of the TOF point cloud. The alignment of such point clouds is commonly referred to as point cloud registration, where standard techniques such as the Iterative Closest Point (ICP) algorithm (BESL et al., 1992) or Least Squares 3D Surface Matching (LS3D) (GRUEN & AKCA, 2005) only exploit information of the spatial 3D geometry. The ICP algorithm relies on iteratively minimizing the difference between two point clouds which can be very time-consuming for large point clouds and requires a good initial alignment of the respective 3D point clouds. In contrast, the LS3D approach minimizes the distance between matched surfaces which also requires a good initial alignment of the point clouds. If unreliable 3D/3D correspondences might be present in the set of all 3D/3D correspondences, a RANSAC-based filtering (FISCHLER & BOLLES, 1981) scheme can be applied in order to increase the robustness of the estimation (SEO et al., 2005).
In this paper, the matching of both a TOF camera with intensity image and 3d point cloud and a thermal infrared camera with an intensity image is investigated for scenes without the restriction of planarity. At first, a relative orientation of the IR camera to the TOF camera is done using homologues points. From these points, 3d coordinates in a relative coordinate system are calculated. These homologues points are directly related to 3d points of the TOF camera which can now be enriched with an additional thermal intensity value. Additionally, ground control points in an object coordinate system are used for verification of the results.

2 Methodology

The method investigated in this paper deals with two main aspects: the relative orientation of the cameras and the co-registration of the TOF 3d point cloud and the thermal infrared intensity image. The relative orientation is based on 5 parameters, namely the displacement in axis Y and Z for a fixed displacement in X and the 3 angles around the axes. It is assumed that the cameras

are already calibrated geometrically (ALBERTZ & WIGGENHAGEN, 2009). These 5 parameters are classically determined using homologues feature points in two images like Förstner or SIFT. Due to the different radiometry of the TOF camera intensity (near infrared) and the thermal camera (thermal infrared), radiometric features cannot be compared with sufficient certainty. Thus, in both images, only the coordinates of features are used. If the relative orientation of the cameras is roughly known, the assignment of the features only from their image coordinates is possible. From this set of homologues points, the relative orientation of the thermal camera to the TOF camera is calculated.

In the second step, the model coordinate system of the intensity images of the two cameras with the origin in the TOF camera has to be transferred into the model coordinate system of the TOF camera 3d point cloud. This is necessary because the relative orientation of the intensity images does not include a scalar factor and additionally refines the orientation of the cameras. This transformation is done by a 3D/3D matching. From the relative orientation, 3d points in the model coordinate system of the camera are calculated for the homologues points. In the TOF camera, these points are connected to a depth value of the point cloud with the same image coordinate. So corresponding 3D coordinates in both model coordinate systems are known.
If reliable 3D/3D correspondences are already known, there is no need to estimate the closest points as should be done before an ICP matching. Further assuming that the 3D/3D correspondences represent physically almost identical 3D points, there is no need to estimate and align the respective object surfaces. Hence, the straight forward approach consists of applying a standard rigid transformation which minimizes the difference between the point clouds in the standard Least Squares sense.
For aligning 3D point clouds in a common coordinate frame, the respective transformation between them has to be estimated. This transformation consists of a rotation described with the rotation matrix **R** and a translation represented as translation vector **t**. The standard rigid transformation minimizes the error between the point clouds by minimizing the energy function

$$E = \sum_i \left\| \mathbf{X}'_i - \left(\mathbf{R}\,\mathbf{X}_i + \mathbf{t} \right) \right\|^2$$

and thus the sum of squared distances between respective 3D points. For estimating such a standard rigid transformation, at least three 3D/3D correspondences are required, but the presence of more correspondences increases the accuracy of the estimated transformation.

The 3d point cloud of the TOF camera is now projected into the thermal infared image to interpolated thermal intensities for every TOF 3d point.

3 Experimental results

3.1 Camera system

For obtaining an infrared-textured 3D model of a scene, thermal information about the local environment has to be captured with a thermal infrared device (InfraTec VarioCAM hr) and the respective 3D information has to be captured with a time-of-flight (ToF) camera (PMD[vision] Camcube 2.0). Furthermore, the used scanning device should also be suited for recording

intensity information which is co-registered to the spatial measurements. These devices have been mounted on a sensor platform as shown in Figure 1. Thus, the devices are coupled and a fixed relative orientation between them is preserved.



Figure 1: sensor configuration: left: TOF camera, right: thermal infrared camera

A PMD[vision] CamCube 2.0 simultaneously captures geometric information as well as radiometric information and thus various types of data by images with a single shot and a frame-rate of 25 frames per second. In contrast to the classical stereo observation techniques, the monostatic sensor configuration of the PMD[vision] CamCube 2.0 preserves information on a discrete 2D grid without the need of a co-registration of the captured data. For each point on this discrete grid, three features are measured, namely the respective range, the active intensity and the passive intensity. The active intensity depends on the illumination emitted by the sensor, whereas the passive intensity depends on the background illumination arising from the sun or other external light sources. Using the grid information, the different types of data can be represented as images. These images have a size of 204x204 pixels which corresponds to a field of view of 40°x40° and hence, the device provides measurements with an angular resolution of approximately 0.2°. However, the non-ambiguity range which is also called unique range is less than 10 m and depends on the tunable modulation frequency. In order to overcome this range measurement restriction, image- or hardware-based unwrapping procedures have recently been introduced (JUTZI, 2009; JUTZI, 2012).
The used infrared camera is a bolometer-based VarioCAM hr from InfraTec. Its sensor records in the wavelength interval from 7.5 to 14 µm with a radiometric resolution of 0.05 K. The image representation of the captured thermal information has a size of 384x288 pixels and, considering an angular resolution of approximately 0.16°, this corresponds to a field of view of approximately 61° to 46°. As the frame rate is 25 fps, this device can also be applied for observing dynamic scenes. Both cameras have been calibrated geometrically before to obtain their interior orientation and distortion.

3.2 Scenario

The test scenario is an indoor scene with different depth layers of a static scene and moving objects and persons. Both cameras record synchronized images with 25 fps. Figure 2 shows an exemplary image pair of the scene. In the color-coded thermal image of the infrared camera, the cube that is held by the person is rarely visible, but the person can easily be separated from the

cold background. The edges and corners of the shelves of the board and the door can be seen. In the TOF passive intensity image, the borders and edges of the shelves and the door also appear clearly, but show different contrast and behavior in their local image neighborhood. In contrast to the thermal image, where all the books show their temperature, in the TOF intensity image, they show almost their visible spectrum intensity.

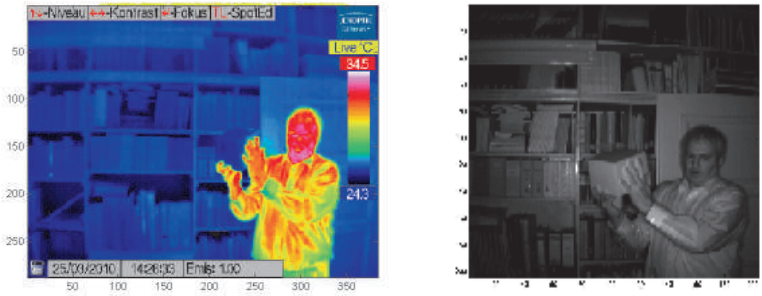


Figure 2: synchronized image pair: left: thermal infrared image with color-coded temperature, right: active intensity image of the TOF camera. Both images before distortion correction.

The distance image of the TOF camera (Fig. 3 left) shows that the cube and the person can easily be separated from the background, but the cube and person can hardly be separated. This is eary to see in the TOF point cloud generated from the depth image (Fig. 3 right). A combination of both the thermal and depth information allows to segment foreground and background (TOF depth image) and cube and person (thermal image).

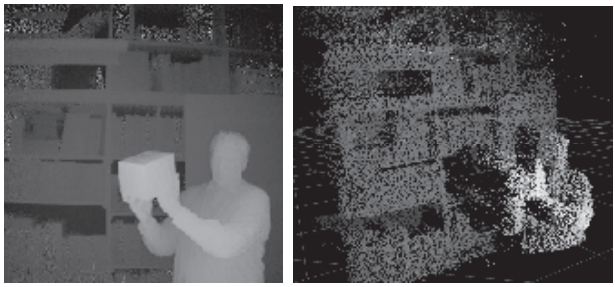


Figure 3: Left: Distance image of the TOF camera. Right: Point cloud generated from the depth image. Bright values have a smaller distance to the sensor than dark values.

3.3 Relative Orientation

Due to the similarity of the observed scene it is assumed, that geometric features like edges and corners appear in both intensity images in the same geometric order. Features are automatically selected from their reliability. This takes into account the position and orientation of the feature

in both images, the certainty of the depth value, and the distribution in the images (Fig. 4). Points are found on the shelves, in the shelves and at the door. The cube and the person are ignored here. The person shows no straight lines and the outline of the person has unsure depth values. The cube shows no edges in the thermal image.

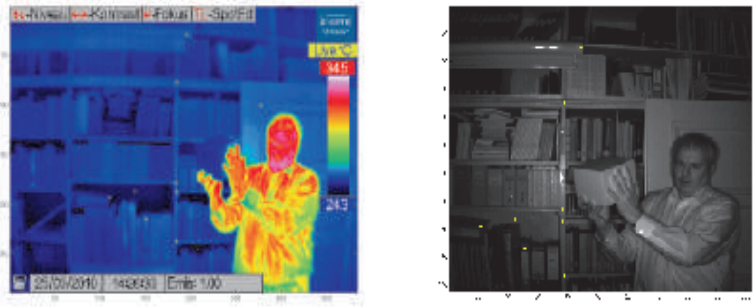


Figure 4: Same image pair as Fig. 2. Corresponding homologues points are extracted for the determination of the relative orientation.

The 5 parameters of the relative orientation are no determined in a bundle adjustment with the TOF camera in the origin and the baseline b_x with length 1 fixed. Using the relative orientation, 3d points for the homologues points are computed and matched with the corresponding TOF depth points to calculated the scale factor of the relative orientation to the coordinate system of the TOF camera. The baseline b_x was manually measured with about 18 cm. After the matching of the relative oriented point cloud onto the TOF point cloud, the calculated baseline is given with 16,6 cm.

In the last step, the TOF 3d points can now be projected into the image plane of the thermal camera to interpolate thermal intensity values for every TOF 3d point (Fig. 5). In the point cloud, there is now a cold background representing the shelves with a big gap, where the shelves are hidden by the person. The person and the cube are the second group of 3d points that is clearly before the shelves. The shape of the person is clearly visible in the 3d point coordinates and the thermal colors of the points which indicate a good fit of the TOF point cloud and the thermal image. In front of the gap in the shelves, there are a few single 3d points with warmer color. These are points at the outline of the hand of the person, where the depth value is unsure. At the right edge of the scene, a few points show a temperature color distribution. This is the projection of the caption of the original infrared image.

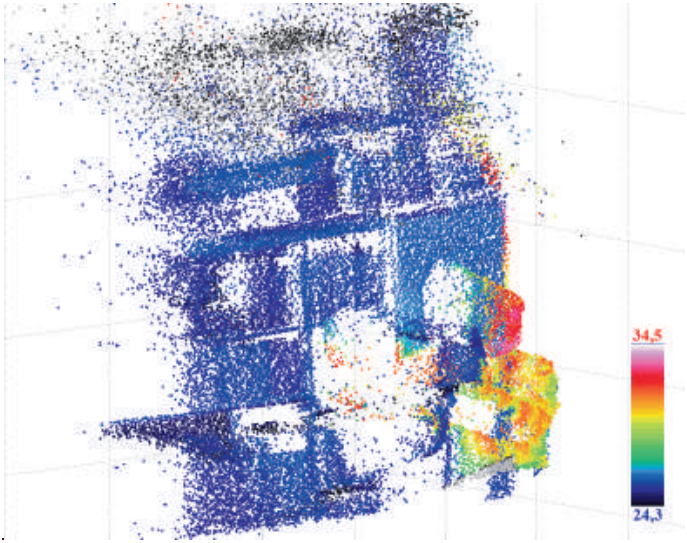


Figure 5: TOF 3d points with thermal intensities.

4 Conclusion

The co-registration of the TOF and thermal IR camera using relative orientation and the TOF 3d point cloud shows more flexibility and accuracy compared to the homography approach (WEINMANN et al., 2012). For scenes with different depth layers of moving objects, the new approach allows a better matching of TOF and thermal images. The approach depends on the possibility to find corresponding, non moving geometric elements in both images. For scenes without structures elements, further investigation are necessary on the extraction of feature pairs in different spectral domains. The experiments so far, are working also on image sequences, but there are so far no investigations on how the sequence itself could be used to improve the orientation. The relative orientation is determined once in one image pair and then used for the hole sequence.

5 Literature

ALBERTZ, J. & WIGGENHAGEN, M., 2009:Taschenbuch der Photogrammetrie und Fernerkundung, Wichmann, Heidelberg. ISBN 978-3-87907-384-9

BAI, X., ZHOU, F. & XUE, B., 2011: Fusion of infrared and visual images through region extraction by using multi scale centersurround top-hat transform. Optics Express 19(9), pp. 8444– 8457.

BESL, P. J. & MCKAY, N. D., 1992: A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence 14(2), pp. 239-256.

CHEN, S. & LEUNG, H., 2009: An EM-CI based approach to fusion of IR and visual Images. 12th International Conference on Information Fusion, pp. 1325–1330.

COIRAS, E., SANTAMARIA, J. & MIRAVET, C., 2000: A segmentbased registration technique for visual-IR images. Optical Engineering 39, pp. 282–289.

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

GRUEN, A. & AKCA, D., 2005: Least squares 3D surface and curve matching. ISPRS Journal of Photogrammetry and Remote Sensing 59(3), pp. 151-174.

HARTLEY, R. I. & ZISSERMAN, A., 2008: Multiple view geometry in computer vision. University Press, Cambridge.

HOEGNER, L., KUMKE, H., MENG, L. & STILLA, U., 2007a: Automatic extraction of textures from infrared image sequences and database integration for 3D building models. PFG Photogrammetrie, Fernerkundung, Geoinformation 6 / 2007, pp. 459–468.

IWASZCZUK, D., HOEGNER, L. & STILLA, U., 2011: Detection of windows in IR building textures using masked correlation. In: Stilla, U., Rottensteiner, F., Mayer, H., Jutzi, B., Butenuth, M. (Eds.), Photogrammetric Image Analysis, ISPRS Conference - Proceedings. Lecture Notes in Computer Science, Vol. 6952, Springer, pp. 133–146.

JUTZI, B., 2009: Investigations on ambiguity unwrapping of range images. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3 / W8), pp. 265–270.

JUTZI, B., 2012: Extending the range measurement capabilities of modulated range imaging devices by time-frequency multiplexing. AVN - Allgemeine Vermessungs-Nachrichten 2 / 2012.

MAYER, H., 2007: 3D Reconstruction and Visualization of Urban Scenes from Uncalibrated Wide-Baseline Image Sequences. In: Photogrammetrie – Fernerkundung – Geoinformation 2007(3), pp. 167–176.

PARK, C., BAE, K.-H., CHOI, S. & JUNG, J.-H., 2008: Image fusion in infrared image and visual image using normalized mutual information. Proceedings of SPIE, Vol. 6968, pp. 69681Q–1–9.

POLLEFEYS, M., NISTÉR, D., FRAHM, J. M., AKBARZADEH, A., MORDOHAI, P., CLIPP, B., ENGELS, C., GALLUP, D., KIM, S. J., MERRELL, P., SALMI, C., SINHA, S., TALTON, B., WANG, L., YANG, Q., STEWÉNIUS, H., YANG, R., WELCH, G. & TOWLES, H., 2008: Detailed real-time urban 3D reconstruction from video. In: International Journal of Computer Vision (IJCV), 78(2-3), 143-167

SEO, J. K., SHARP, G. C. & LEE, S. W., 2005: Range data registration using photometric features. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2, 1140-1145.

WEINMANN, M., WURSTHORN, S. & JUTZI, B., 2011: Semi-automatic image-based co-registration of range imaging data with different characteristics. PIA11 - Photogrammetric Image Analysis. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science. 38(3) / W22, S. 119-124.

WEINMANN, M., HOEGNER, L., LEITLOFF, J., STILLA, U., HINZ, S. & JUTZI, B., 2012: Fusing Passive and Active Sensed Images to Gain Infrared-textured 3D Models. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., 39(B1), 71-76

Erstellung großflächiger Höhenmodelle und Vergleich mit anderen Daten

KARSTEN JACOBSEN[1]

Zusammenfassung: Es wird ein Überblick über die Möglichkeiten zur Erstellung von Höhenmodellen mit optischen und Radaraufnahmen aus dem Weltraum gegeben. Die sensorbedingten Eigenschaften der Höhenmodelle werden erläutert und es erfolgt ein Vergleich mit großflächigen bis weltweit vorhandenen und angekündigten digitalen Oberflächenmodellen (DOM).
Mit den hochauflösenden optischen Weltraumbildern lassen sich DOM für flache Gebiete mit Systemgenauigkeiten (relative Standardabweichung) in der Größenordnung von einem Objektpixel erstellen. Mit flächenbasierten Bildzuordnungen ist ein Punktabstand der DOM in der Größenordnung von drei Objektpixeln sinnvoll; bei Semi-Global-Matching kann wegen der pixelbasierten Zuordnung ein Objektpixelabstand von einem Pixel gewählt werden. Radaraufnahmen haben den Vorteil der Unabhängigkeit von der Wolkenbedeckung weswegen sie sich besonders für großflächige Höhenmodelle eignen , sind aber bei abrupten Höhenunterschieden, wie sie in Städten und Gebirgen auftreten, durch layover beeinträchtigt.

1 Einleitung

Digitale Höhenmodelle werden für viele Anwendungen benötigt und sind ein Grundbestandteil der meisten Geoinformationssysteme. Ihre Bestimmung mittels Luftbildphotogrammetrie oder Laserscanneraufnahme ist zeitaufwendig und damit teuer. Nicht in jedem Fall ist die höchste Genauigkeit und Auflösung erforderlich, so dass unter Umständen auf die nahezu weltweiten, kostenlosen Höhenmodelle zurückgegriffen werden kann. Eine höhere Auflösung und Genauigkeit wird mit dem TanDEM-X Höhenmodell ab 2014 vorhanden sein oder bereits jetzt mit dem TanDEM-X intermediate DEM. Von diesem DHM zu den Ergebnissen der Luftbildphotogrammetrie gibt es einen gleitenden Übergang durch hochauflösende optische Satellitenbildstereomodelle. Die erzeugten Höhenmodelle unterscheiden sich nicht nur durch den Punktabstand und die allgemeine Genauigkeit, die verschiedenen Sensoren haben unterschiedliche Eigenschaften, die spezielle Charakteristika der erstellten DHM hervorrufen.

2 Höhenmodelle aus optischen, hochauflösenden Satellitenbildern

Mit den Bildern von WorldView-1, WorldView-2 und GeoEye-1 stehen zurzeit Satellitenbilder mit einer Bodenauflösung von bis zu 0,5 m zur Verfügung. Mit den angekündigten Satelliten GeoEye-2, WorldView-3 und Cartosat-3 wird die Bodenauflösung in absehbarer Zeit auf 31 cm bis 34 cm verbessert werden. Nicht nur die Bodenauflösung, auch die Aufnahmekapazität hat sich erheblich gesteigert (Abb. 1). Heute haben die hochauflösenden Satelliten insgesamt

1) Dr. Karsten Jacobsen, Leibniz Universität Hannover, Nienburger Str.1, 30167 Hannover; E-Mail: jacobsen@ipi.uni-hannover.de

nicht nur theoretisch, sondern auch praktisch eine um mehr als den Faktor 10 gesteigerte Aufnahmekapazität als vor vier Jahren, wodurch sich die Zeit von der Beauftragung bis zur Aufnahme der Satellitenbilder deutlich verringert hat.

Alle hochauflösenden optischen Satelliten sind heute flexibel – durch Drehung des ganzen Satelliten wird die Aufnahmerichtung geändert, womit die Erstellung von Stereomodellen innerhalb von etwa 90 Sekunden möglich ist. Die früheren Probleme der Objektänderung von einer zur nächsten Aufnahme bestehen damit nicht mehr, was die automatische Bildzuordnung erheblich verbessert. Die beiden WorldView-Satelliten und Pleiades sind mit control moment Kreiseln statt mit reaction Kreiseln ausgestattet, wodurch sie sich erheblich schneller drehen können als die anderen Satelliten (Abb. 2), damit ist die Voraussetzung zur Erstellung von Stereo- oder auch Multistereoaufnahmen erheblich verbessert. Aus diesem Grund wurde mit WorldView-1 und -2 im Jahr 2011 eine sehr viel größere Fläche stereoskopisch aufgenommen als mit IKONOS und GeoEye-1 (GOURDINE 2012). Wegen der langsameren Drehung von GeoEye-1 und IKONOS ist die Erstellung von Stereoaufnahmen nicht sehr wirtschaftlich. Die künftigen sehr hoch auflösenden optischen Satelliten werden aus diesem Grund mit control moment Kreiseln ausgestattet sein.

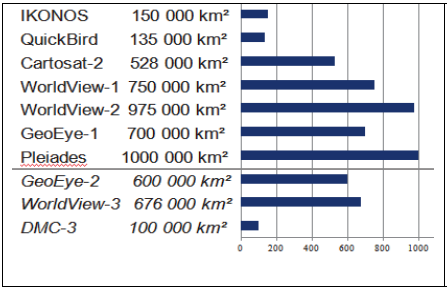


Abb. 1: Theoretische Aufnahmekapazität hochauflösender optischer Satelliten je Tag

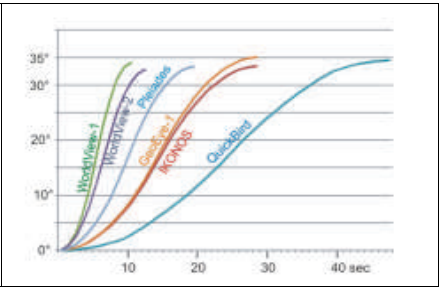


Abb. 2: Zeit für die Aufnahmerichtungsänderung optischer Satelliten

Mit WorldView-2 wurden am 6. Juni 2011 innerhalb von 92 Sekunden drei sich teilweise überlappende Stereobilder des Schwarzmeerküstenbereichs in der Nähe von Istanbul aufgenommen (Abb. 3) um damit Höhenmodelle zu erstellen da die amtlichen Höhenmodelle wegen großflächigen Bodenabbaus nicht mehr aktuell waren (BÜYÜKSALIH et al. 2012). Die Szenenorientierung mit bias-korrigierten rationalen Polynomkoeffizienten basierte auf Passpunkten, die für Luftbildaufnahmen mit 10 cm Objektpixelgröße erstellt wurden. Dementsprechend war die Identifizierung der Passpunkte in den WorldView-Aufnahmen schwierig und es wurde im Durchschnitt eine Standardabweichung der Orientierung von etwa einem Objektpixel (50 cm) erreicht. Durch Bildzuordnung nach der Methode der kleinsten Quadrate basierend auf einer Kombination von Bildpyramiden mit Regionswachstum wurden drei sich teilweise überlappende Höhenmodelle erstellt. In den Waldgebieten musste die Toleranzgrenze für die Bildzuordnung auf 0,4 gesenkt werden um eine ausreichende Anzahl von Objektpunkten zu erhalten, das wirkte sich kaum auf die Genauigkeit gegenüber einer Toleranzgrenze von 0,4 aus. Wie das Beispiel eines Qualitätsbildes (Abb. 4) zeigt, sind die

Korrelationskoeffizienten der Bildzuordnung stark von der Bodenbedeckung abhängig. Für offenes Gelände und dementsprechend auch die Straßen ist der Korrelationskoeffizient mit 0,95 bis 1,0 am größten.



Abb. 3: links: überlappende WorldView-2 Aufnahmen rechts: Aufnahmekonfiguration der 3 Stereokombinationen

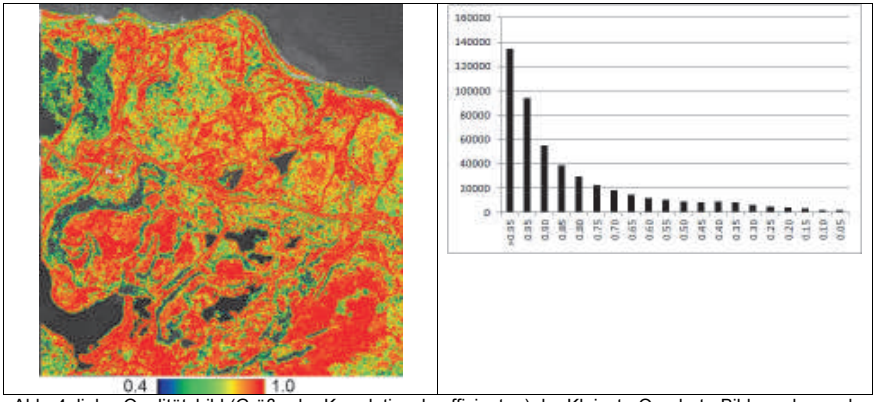


Abb. 4: links: Qualitätsbild (Größe der Korrelationskoeffizienten) der Kleinste-Quadrate-Bildzuordnung der WV-2-Aufnahme überlagert dem WV-2-Bild rechts: entsprechende Häufigkeitsverteilung der Korrelationskoeffizienten

Die in Abbildung 4 nicht durch zugeordnete Punkte mit einem Korrelationskoeffizienten größer als 0,4 (grauer Hintergrund) erfassten Flächen sind überwiegend Wasserflächen, teilweise aber auch Waldgebiete.

Die DOM in den Überlappungsbereichen der Stereomodelle wurden unabhängig voneinander erstellt – sie basieren auf unterschiedlichen Passpunkten und sind in den jeweils beiden Modellen nicht lageidentisch, erfordern für einen Vergleich somit eine Interpolation. Das Quadratmittel von 411.354 Widersprüchen der überlappenden DOM ergibt: $SZ = 0.74m + 0.55m * \tan$

(Geländeneigung). Da die Fehleranteile beider überlappender Höhenmodelle gleich sind, folgt daraus als Standardabweichung der Höhenpunktbestimmung durch ein WV-2 Modell in flachem Gelände von SZ = 0,52 m, was etwa einer Objektpixelgröße entspricht. Die Widersprüche gegenüber dem Referenzhöhenmodell sind deutlich größer, was auf den Einfluss von Vegetation, Bodenabbau, Gebäuden, aber auch eine begrenzte Genauigkeit des Referenzhöhenmodells zurückzuführen ist. Für offenes Gelände ergibt sich im Quadratmittel ein Widerspruch von SZ = 1,72m + 0,93∗ tan (Geländeneigung) und nach Filterung von Punkten, die nicht der Geländeoberfläche entsprechen, SZ = 1,44m + 0,85∗ tan (Geländeneigung). Ein Teil der Widersprüche ist auf verbliebene Einflüsse der Vegetation und von Gebäuden zurückzuführen, das Referenzhöhenmodell entspricht aber auch nicht der für ein Höhenmodell aus Bildern mit 0,5 m Objektpixelgröße erforderlichen Qualität eines Referenzhöhenmodells.
Eine relative Standardabweichung der DOM in flachen Bereichen von etwa einer Objektpixelgröße ist typisch für die automatische Bildzuordnung optischer Satellitenstereobilder, sie wurde auch mit SPOT HRS (5 m Objektpixelgröße) (JACOBSEN 2004), Cartosat-1 (2,5 m Objektpixelgröße) (JACOBSEN 2006) und IKONOS (1 m Objektpixelgröße) sowie GeoEye-1 Stereomodellen (0,5 m Objektpixelgröße) (ALOBEID et al. 2010) erreicht. Die relative Genauigkeit ist identisch mit der absoluten wenn ausreichend genaue Passpunkte vorhanden sind. Ohne Passpunkte ist die absolute Lagerung der Genauigkeit von der direkten Sensororientierung abhängig. Die in Tabelle 1 angegebenen Herstellergenauigkeiten wurden durch eigene Untersuchungen bestätigt. Allgemein wird die Genauigkeit mit CE90 (circular error 90% Wahrscheinlichkeit) angegeben. Bei Normalverteilung der Ergebnisse ist die Standardabweichung der Objektkoordinatenkomponente SX oder SY gleich CE90 / 2,146. Die begrenzte Orientierungsgenauigkeit des Stereosatelliten Cartosat-1 kann durch In-Flug-Kalibrierung auf etwa 30 m verbessert werden, für operationelle Anwendungen werden die Cartosat-1-Modelle allerdings meist auf das SRTM-Höhenmodell durch Ausgleichung verschoben, womit die absolute Genauigkeit der SRTM-Höhenmodelllagerung in der Größenordnung von 3 m für die Höhe und 10 m für die Lage erreicht wird. Eine Korrektur der absoluten Höhenlage ist mit Profilpunkten des Lasersatelliten ICESat möglich, die eine Standardabweichung von 0,2 m bis 0,3 m haben. Wegen des ICESat Fußabdruckdurchmessers von 170 m erfordert eine derartige Höhenkorrektur in den Bereichen der ICESat-Profilpunkte ein ausreichend ebenes Gebiet. Eine Blockausgleichung überlappender Modelle verbessert die Orientierungsgenauigkeit (D' ANGELO, REINARTZ 2012).

Tabelle 1: Genauigkeit der direkten Sensororientierung optischer hochauflösender Satelliten

Sensor	SX = SY	CE90
IKONOS	7m	15m
QuickBird	9m	19m
Orbview-3	12m	25m
WorldView-1	2m	5m
WorldView-2	2m	5m
GeoEye-1	2m	5m
Pleiades	2m	3m
Cartosat-1 / mit in-flight Kalibrierung	93m / 30m	200m
KOMPSAT-2	37m	80m

3 Höhenmodelle aus Satellitenradarbildern

Synthetische Apertur Radaraufnahmen (SAR) sind wie die optischen Bilder ebenfalls bildhafte Darstellungen der Rückstrahlungsintensität, allerdings in einem deutlich anderen Wellenlängenbereich (Tabelle 2). Optische Bilder separieren die Objektpunkte nach der Richtung aus der die Strahlung kommt, während bei SAR-Aufnahmen nur die Komponente in der Flugrichtung richtungsabhängig ist, ist die Komponente quer zur Flugrichtung von der Entfernung Antenne zum Objekt abhängig.

Tabelle 2: Für Höhenbestimmung bedeutsame Radarwellenlängenbereiche

Band	X	C	L	P
Wellenlänge	2,4 – 3,75 cm	3,75 – 7,5 cm	15 – 30 cm	77 – 136 cm
Höhe im Wald	DOM	DOM	nahezu DHM	DHM

Für SAR-Bilder und die Erstellung von Höhenmodellen werden hauptsächlich X-, C-, L- und P-Band eingesetzt (Tabelle 2). Je kürzer die Wellenlänge und die Pulsdauer, desto besser ist die Detailerkennbarkeit. Umgekehrt kann mit den längeren Wellenlängen L- und P-Band Vegetation weitgehend durchdrungen werden, während mit dem X- und C-Band wie bei den optischen Aufnahmen digitale Oberflächenmodelle erstellt werden.

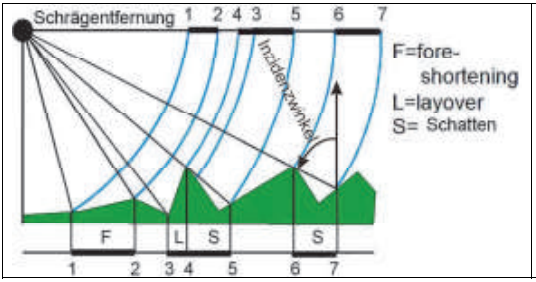


Abb. 5: Grundprinzip und Probleme der Radarabbildung quer zur Flugrichtung

SAR-Aufnahmen werden mit Inzidenzwinkeln (Abb. 5) zwischen 20° und 65° aufgenommen. Ein großer Inzidenzwinkel hat den Vorteil einer besseren Bodenauflösung, vergrößert aber Sichtschattenbereiche und die Signalüberlagerung des sogenannten Layover (Bereich L in Abb. 5). Im Gebirge und in städtischen Bereichen ist ein Layover nicht vermeidbar und führt zu Auswertelücken, die bei Gebieten, die mehrfach erfasst werden, durch unterschiedliche Inzidenzwinkel und die unterschiedliche Aufnahmerichtung des aufsteigenden und des absteigenden Orbits zumindest verkleinert werden.
Wie bei optischen Bildern werden zur Höhenbestimmung mit SAR-Bildern zwei Bilder benötigt, da aus den 2 Dimensionen eines Bildes keine 3 Dimensionen des Objektkoordinatensystems ermittelt werden können. Eine Höhenbestimmung ist durch interferrometrisches SAR (InSAR) mit einer kurzen Basis durch Vergleich der Phasendifferenz der von zwei Antennen aufgenommenen Rückstrahlung oder durch Radargrammetrie mittels Bildzuordnung zweier mit großer Basis aufgenommener Bilder möglich. InSAR mit X- und C-Band benötigt zeitgleich erstellte Aufnahmen da bei Zeitunterschieden die Rückstrahlung durch die Vegetation zu

unterschiedlich ist, so dass es zu Dekorrelationen kommt. Bei L- und P-Band kommt die Rückstrahlung hauptsächlich vom Boden, womit auch SAR-Aufnahmen mit Zeitunterschieden durch InSAR ausgewertet werden können. Zurzeit ist aber kein zivil verfügbarer Satellit mit L- oder P-Band im All, was sich erst mit dem für 2013 vorgesehenen Start von ALOS-2 ändern wird.
Für die Erstellung großflächiger Höhenmodelle ist eine InSAR-Aufnahmekonfiguration von erheblichem Vorteil, wie es bei der nur 11 Tage dauernden SRTM-Mission war, mit dem ein nahezu weltweites Höhenmodell erstellt wurde. Zurzeit wird mit TanDEM-X, das aus den in naher Konfiguration fliegenden zwei SAR-Satelliten des DLR besteht, ein weltweites Höhenmodell, das TanDEM-X Global DEM aufgebaut, das ab 2014 verfügbar sein soll.
Das TanDEM-X Global DEM wird von ASTRIUM vertrieben werden. Es wird eine Punktdichte von 0,4 arcsec (~12 m am Äquator) haben. Das Deutsche Fernerkundungsdatenzentrum (DFD) des DLR gibt folgende Genauigkeiten des TanDEM-X-Höhenmodells an (HUBER et al 2012):

absolute Lagegenauigkeit: CE90 < 10 m, entsprechend SX, SY < 4.7 m
absolute Höhengenauigkeit: LE90 < 10 m (SZ < 6 m)
relative Höhengenauigkeit innerhalb einer 1° x 1° Masche für Geländeneigungen von kleiner gleich 20%: LE90 < 2 m, entsprechend SZ < 1,2 m und für Geländeneigungen mit einer Neigung größer als 20%: LE90 < 4 m, entsprechend SZ < 2,4 m

Wobei die absolute Genauigkeit durch Blockausgleichung und Einpassung auf ICESat-Höhenpunkte verbessert wird, womit die absolute Höhengenauigkeit nur als Spezifikation angesehen werden muss und in der Regel deutlich besser sein wird.
Die Probleme der SAR-Bilder mit Layover sollen in den Gebirgsregionen durch Mehrfachaufnahme verringert werden.
Die italienische COSMO SkyMed Satellitenkonfiguration hat keinen Tandem-Flug, weswegen mit ihnen Höhenmodelle durch Radargrammetrie erstellt werden, was weltweite Höhenmodelle weitgehend ausschließt (CAPALDO et al. 2010).

4 Vergleich der Höhenmodelle

Zurzeit ist nur das GMTED 2010 als vollständig weltweites, kostenloses Höhenmodell im Internet verfügbar. Das GMTED 2010 ist mit 30, 15 und 7,5 arcsec Punktabstand (~230 m am Äquator) vorhanden. Es basiert zu 70% auf SRTM-Daten. Als Nachfolgemodell des GTOPO 30 ist es durch die deutlich bessere Genauigkeit und Homogenität des SRTM-Höhenmodells dominiert und weist deutlich bessere morphologische Details auf.
Nahezu weltweite, kostenlose Höhenmodelle sind das InSAR SRTM-Höhenmodell, das auf der Shuttle Radar Topography Mission vom Februar 2000 beruht und das ASTER GDEM, das aus allen verfügbaren Höhenmodellen des optischen Stereosatelliten ASTER zusammengestellt wurde.
Das SRTM Höhenmodell wurde gegenüber der ersten Version um Ausreißer verbessert und es wurden Lücken gefüllt. Lücken tauchen, wie bei InSAR-Höhenmodellen üblich, durch Layover im Gebirge und auch in trockenen Sandwüsten auf in denen es keine Rückstrahlung gibt. Meist wurden Höhen aus dem Reference 3D (SPOT High Resolution Stereo) zur Lückenfüllung benutzt. Außerhalb der USA sind die SRTM Höhenmodelle nur mit einer Rasterweite von 3 arcsec (~92 m am Äquator) vorhanden, es gibt allerdings eine Diskussion die Höhenmodelle auch mit 1 arcsec verfügbar zu machen.

Die erste Version des ASTER GDEM hatte vor Mittelung der Ergebnisse die einzelnen ASTER Höhenmodelle nicht optimal einander zugeordnet, wodurch morphologische Details verloren gingen. Dieses wurde mit der zweiten Version GDEM2, die seit Oktober 2011 verfügbar ist, verbessert, außerdem wurden weitere Höhenmodelle zur Genauigkeitssteigerung hinzugefügt. Eine weitere Verbesserung durch zusätzliche Höhenmodelle ist in Arbeit und soll als GDEM3 im Laufe des Jahres 2013 in das Internet gestellt werden (URAI et al. 2013). Trotzdem sind die Genauigkeiten des ASTER GDEM etwas heterogen (Abb. 6 links), da die Anzahl der benutzten Bilder je Objektpunkt stark variiert. Als Mittelwert von 12 Testgebieten ergab sich eine Standardabweichung der Höhe von SZ = 12,4m – 0,15 ∗ Bilder/Punkt. Dementsprechend weist ein Einzelmodell eine Genauigkeit von 12,1 m auf, was bei den 15 m Bodenauflösung der ASTER-Bilder 0,8 Objektpixeln entspricht. ASTER GDEM hat den Vorteil einer Rasterweite von 1 arcsec (~30 m am Äquator).

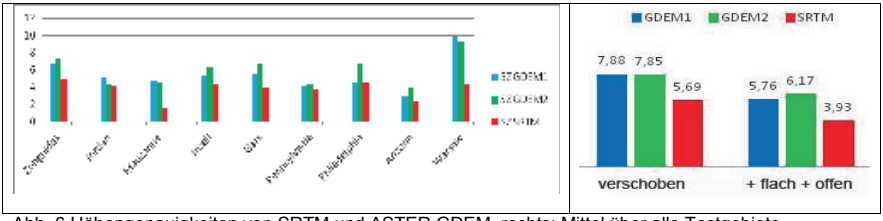


Abb. 6 Höhengenauigkeiten von SRTM und ASTER GDEM, rechts: Mittel über alle Testgebiete

In den in Abb. 6 dargestellten, weltweit verstreut liegenden Untersuchungsgebieten, die stark unterschiedlichen Charakter aufweisen, zeigten sich Lagefehler der Höhenmodelle in der Größenordnung von 10 m bis 30 m, außerdem ist wie bei praktisch allen Höhenmodellen eine Abhängigkeit der Genauigkeit von der Geländeneigung vorhanden, die mit der Koppe'schen Formel SZ=A + B∗tan (Geländeneigung) gut modelliert werden kann. Für flaches Gebiet, das nicht durch Vegetation gestört ist, zeigt das ASTER GDEM eine Standardabweichung von etwa 6m und das SRTM Höhenmodell eine von etwa 4m.

Nicht nur die Genauigkeit und Rasterweite der Höhenmodelle, auch die morphologischen Details sind von Bedeutung. In Abbildung 7 ist deutlich der Einfluss der Rasterweite der zugrunde liegenden Höhenmodelle zu sehen. Mit den durchschnittlichen 80 m Rasterweite des SRTM C-Band Höhenmodells sind Details der Höhenlinien nicht erfassbar. Das ist bei dem SRTM X-Band Höhenmodell, das vom DLR mit 1 arcsec Rasterweite verfügbar ist, allerdings große Lücken zwischen den Aufnahmestreifen hat, erheblich besser. Deutlich ist die morphologische Verbesserung vom ASTER GDEM1 zum GDEM2 erkennbar. Selbstverständlich sind mit den durchschnittlich 201 m des GMTED2010 nicht die Details erkennbar, wie mit den anderen Höhenmodellen.

Weitere großflächige, allerdings kommerzielle, Höhenmodelle gibt es mit NEXTMap World 30, PlanetObserver Global 90 m DEM, Reference 3D, NEXTMap und EURO-MAPS 3D. NEXTMap World 30 fusioniert SRTM mit 3 arcsec Punktabstand und ASTER GDEM zu einer Rasterweite von 1 arcsec. Außerdem werden systematische Höhenfehler durch ICESat Höhenpunkte verbessert. Das PlanetObserver Global 90 m DEM bietet keinen besonderen Vorteil gegenüber dem SRTM Höhenmodell nachdem dessen Lücken gefüllt wurden. Basierend auf Bildern des High Resolution Stereo (HRS) auf SPOT5 (5 m Objektpixelgröße in

Orbitrichtung) wurden von ASTRIUM von etwa 40% der Landfläche Höhenmodelle mit einer Rasterweite von 1 arcsec durch Bildzuordnung erstellt. Die Standardabweichung der Höhe wird für Geländeneigungen bis 20% mit 6 m, für Geländeneigungen >20% bis 40% mit 11 m und für Geländeneigungen >40% mit 18 m angegeben. Damit liegt die Genauigkeit im flachen Gelände bei 1,2 Objektpixelgrößen.
Intermap Technologies erstellte auf eigene Rechnung von Westeuropa, den USA und Teilen anderer Länder ein lückenloses Höhenmodell durch Flugzeug-InSAR mit X-Band aus 10,4 km Höhe mit einem Inzidenzwinkel zwischen 35° und 55°. Diese Höhenmodelle mit einer Rasterweite von 5 m weisen die SAR-typischen Probleme in Städten und Gebirge auf, weswegen die Standardabweichung der Höhe allgemein für 40% mit <0,6 m, für weitere 40% mit 0,6 bis 1,8 m und für 20% mit größer als 1,8 m angegeben wird.

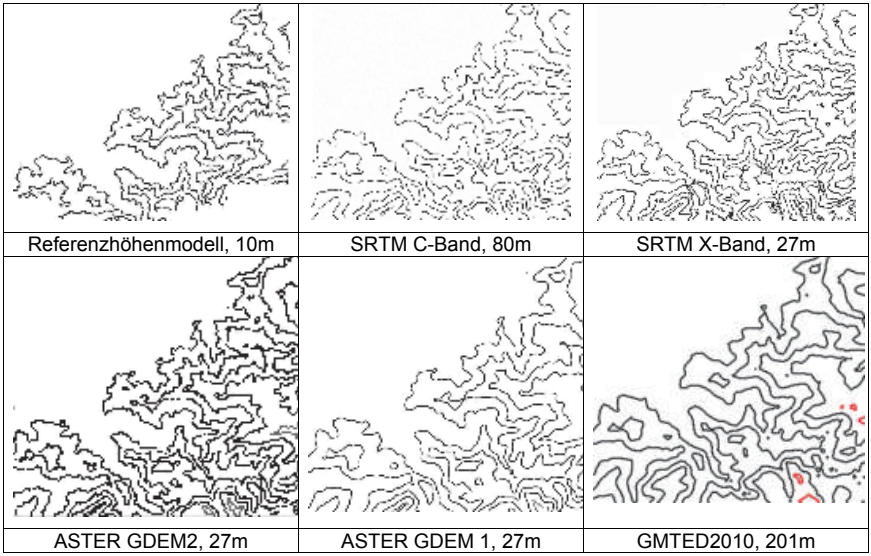


Abb.7: Höhenlinien basierend auf den verschiedenen Höhenmodellen mit Angabe der Rasterweite

Seit 2007 betreibt Euromap in Kooperation mit dem DLR eine Empfangsstation in Neustrelitz mit der Bilder des Stereosatelliten Cartosat-1 aufgenommen werden. Bis jetzt ist der gesamte Einzugsbereich von Neustrelitz fast lückenlos mit ein- bis dreifacher Überdeckung durch Cartosat-1 Stereomodelle aufgenommen. In Zusammenarbeit mit dem DLR wurde das EUROMAPS 3D aus Cartosat-1 Stereobildern erstellt. Da Cartosat-1 eine eingeschränkte Genauigkeit der direkten Sensororientierung von etwa 200 m hat, wurden die Cartosat-1 Höhenmodelle auf das SRTM-DOM eingepasst, wobei eine intern im DLR verbesserte Version des SRTM-Höhenmodells verwendet wurde, das die systematischen Höhenfehler von bis zu 4 m mittels ICESat-Daten verbessert hat. Das DLR (D'ANGELO, REINARTZ 2012) ist jetzt dazu übergegangen die Cartosat-1-Szenenorientierungen durch sehr große Blöcke gemeinsam auf das SRTM-

Höhenmodell zu orientieren um systematische Orientierungsfehler zu minimieren und durch die Szenenverknüpfung Modellanpassungsprobleme zu vermeiden. Die automatische Bildzuordnung erfolgt durch Semi-Global-Matching (SGM) (HIRSCHMÜLLER 2008) um auch feine Objektdetails zu erfassen. SGM kann plötzliche Höhenunterschiede, wie sie in Städten und Gebirgen auftreten, pixelgenau erfassen. Der Vorteil liegt vorwiegend in den morphologischen Details, wogegen die Höhengenauigkeit durch SGM nur geringfügig verbessert wird (ALOBEID et al. 2010). Damit ist eine Rasterweite von 5 m möglich. Euromap gibt die Lagegenauigkeit allgemein mit SX=SY=4,6 m und die Höhengenauigkeit mit SZ<<6 m an, wobei UTTENTHALER et al. 2010 die absolute Genauigkeit innerhalb Europas für die Lage mit SX=SY= 3,1 m und für die Höhe mit SZ= 3,1 m angibt. Als relative Standardabweichung werden 2,5 m genannt, was mit eigenen Untersuchungen (JACOBSEN 2006) übereinstimmt.

5 Schlussfolgerungen

Sowohl mit optischen als auch mit Radaraufnahmen im X- und C-Band werden Höhenmodelle der sichtbaren Oberfläche und nicht des Bodens erstellt. Nur in teilweise offenen Gebieten kann durch Filterung ein DHM erzeugt werden. Nur mittels länger welliger SAR-Aufnahmen kann die Vegetation durchdrungen werden und somit die Höhe des Bodens bestimmt werden – mit Ausnahme der Gebäude. Durch flächenbasierte Zuordnung optischer Weltraumbilder, aber auch durch SGM, kann eine relative Höhengenauigkeit von etwa einer Objektpixelgröße erreicht werden. Diese Genauigkeit bezieht sich allerdings auf die sichtbare Oberfläche und nicht den Boden und gilt nur für flaches Gebiet da die Höhengenauigkeit deutlich von der Geländeneigung abhängig ist. Die absolute Genauigkeit ist von der Sensororientierung abhängig. Sind Passpunkte vorhanden, kann die absolute Genauigkeit mit der relativen Genauigkeit übereinstimmen. Dieses bezieht sich auf die Höhe, aber auch die Lage. Mit Lageverschiebungen der Höhenmodelle ist ohne Verwendung von Passpunkten zu rechnen. Die erreichte hohe Qualität der direkten Sensororientierung erlaubt für viele Anwendungen deren direkte Verwendung. Ist die direkte Sensororientierung nicht genau genug, kann durch Blockausgleichung für einen Verband von Höhenmodellen die relative Genauigkeit auf ein größeres Gebiet ausgedehnt werden und es ist eine Einpassung auf vorhandene Daten, wie das SRTM DOM und ICESat Höhenpunkte möglich. Dieses gilt sowohl für optische als auch SAR-Daten.

Durch optische Stereosatelliten oder InSAR-Aufnahmekonstellationen können großflächige Höhenmodelle erstellt werden, wie das ASTER GDEM, das SPOT HRS Reference 3D oder die großflächigen Cartosat-1 DOM, SRTM und das kommende TanDEM-X Höhenmodell zeigen. SAR-Bilder haben den Vorteil, dass sie nicht von der Wolkenbedeckung beeinträchtigt werden, somit können in kürzerer Zeit als mit den optischen Bildern große Flächen erfasst werden. Die Genauigkeit der Höhenmodelle ist von der Bodenauflösung abhängig, was sowohl für SAR als auch die optischen Daten gilt. Durch Mehrfachaufnahme kann die Genauigkeit gesteigert werden, was besonders für das ASTER GDEM gilt, aber auch bei TanDEM-X zutrifft. Auch die morphologischen Details sind von der Bodenauflösung abhängig, hier haben für Feinstrukturen die optischen Stereomodelle Vorteile, besonders wenn die Auswertung auf dem pixelbasierten SGM beruht.

Die kostenlos verfügbaren Höhenmodelle GMTED 2010, SRTM und ASTER GDEM haben alle eine beschränkte Genauigkeit und Auflösung, auch wenn dieses durch verbesserte Versionen gesteigert wurde und noch wird. Eine höhere Genauigkeit und Auflösung ist nicht kostenlos verfügbar.

6 Literaturverzeichnis

ALOBEID, A., JACOBSEN, K., HEIPKE, C., 2010: Comparison of Matching Algorithms for DSM Generation in Urban Areas from IKONOS Imagery, PERS 76(9). pp. 1041-1050

D' ANGELO, P., REINARTZ, P., 2012: DSM based Orientation of Large Stereo Satellite Image Blocks, IntArchPhRS Vol XXXVIII-B1, pp 209 - 214

BÜYÜKSALIH, G., BAZ, I., ALKAN, M., JACOBSEN, K., 2012: DEM Generation with WorldView-2 Images, IntArchPhRS Vol XXXVIII-B1

CAPALDO, P., CRESPI, M., FRATARCANGELI, F., PIERALICE, F., 2010: Radargrammetric application with COSMO-SkyMed imagery: definition and implementation of an orientation model for Spotlight stereo pairs, EARSeL Symposium Paris, 2010

GOURDINE,R., 2012: Competitive Overview, WorldView User Meeting, München 2012, http://www.spaceimagingme.com/LinkClick.aspx?fileticket=g4Z2t9lCNZM%3D&tabid=172

HIRSCHMÜLLER, H., 2008. Stereo Processing by Semiglobal Matching and Mutual Information, IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(2): 328-341

HUBER M., GRUBER, A., WENDLEDER, A. WESSEL, B., ROTH, A. ,SCHMITT, A., 2012: The Global TanDEM-X DEM: Production Status and first Validation Results, IntArchPhRS Vol XXXVIII-B7

JACOBSEN, K., 2004: DEM Generation by SPOT HRS, ISPRS Congress, Istanbul 2004, Int. Archive of the ISPRS, Vol XXXV, B1, Com1, pp 439-444

JACOBSEN K., 2006: ISPRS-ISRO Cartosat-1 Scientific Assessment Programme (C-SAP) Technical report - test areas Mausanne and Warsaw, ISPRS Com IV, Goa 2006, IntArchPhRS Vol. 36 Part 4, pp. 1052-1056

URAI, M., TACHIKAWA,T., FUJISADA, H., 2012: Data Acquisition Strategies for ASTER Gobal DEM Generation, IntArchPhRS Vol XXXVIII-B4, pp 199-202

UTTENTHALER, A., ANGELO, P., REINARTZ, P., HASS, T., CARL, S., BARNER, F., 2010: A Concept for a Standardized DSM Product Automatically Derived from IRS-P5 Cartosat-1, http://www.euromap.de/pdf/A_Concept_for_a_standardized_DSM_product_automatically_derived_from_IRS-P5_Cartosat-1.pdf

Integrierte Georeferenzierung von luft- und bodengestützten Stereobilddaten im Straßenbereich – Strategien und Ergebnisse

S. NEBIKER[1], K. KAMER[2] & H. EUGSTER[1,2]

Zusammenfassung: In diesem Beitrag werden Lösungsansätze zur integrierten Georeferenzierung von mobil erfassten Stereobilddaten und hoch aufgelösten Luftbilddaten aufgezeigt und an einem konkreten Anwendungsfall untersucht. Es wird gezeigt, dass sowohl unter Einbezug von terrestrisch bestimmten Passpunkten als auch von Passpunkten aus Luftbildern eine markante Steigerung der Georeferenzierungsgenauigkeit gegenüber einer direkten Georeferenzierung möglich ist. Für bildbasierte Updates zur integrierten Schätzung der Fahrzeug-Trajektorie zeigte sich der Ansatz der Einzelpasspunkt-Updates robuster und im konkreten Fall auch genauer als derjenige der Bündelblock-Updates basierend auf Passpunktgruppen. Die Kombination von boden- und luftgestützten Bilddaten ermöglicht mit den vorgestellten Verfahren eine genaue und ökonomische Georeferenzierung bzw. Co-Registrierung. Zudem ergänzen sich die beiden Bilddatensätze und die daraus abgeleiteten Produkte, wie die sehr dichten, texturierten 3D-Puntwolken, in idealer Weise.

1 Einleitung

Modernes Infrastrukturmanagement im Straßen- und Schienenbereich erfordert genaue, zuverlässige und aktuelle dreidimensionale Geodaten. Aus Gründen der Effizienz und Sicherheit werden diese Daten zunehmend mit mobilen Plattformen und Geosensoren erfasst. Obwohl die ersten experimentellen Mobile Mapping Systeme vor über 20 Jahren auf Stereobilddaten basierten (NOVAK, 1993; SCHWARZ et al., 1993), wurde die mobile 3D-Geodatenerfassung in den letzten 10 Jahren durch mobiles Laserscanning dominiert. Dank großer Fortschritte in der Bildsensorik, der Datenspeicherungs- und Übertragungskapazitäten und der Algorithmik für das dichte Stereobildmatching, erleben Bilddaten seit einigen Jahren ein eigentliches Revival. Das Bewusstsein für die Leistungsfähigkeit von boden- und luftgestützten Geobilddatendiensten wird durch bekannte Publikumsdienste wie Google Street View und deren immer engere Integration in Kartendienste wie Google Maps zusätzlich befördert. Während dem die Genauigkeit und Aktualität in den erwähnten Geodiensten für das breite Publikum in der Vergangenheit eine eher untergeordnete Rolle spielte, ist sie für Anwendungen im Infrastrukturmanagement zentral.

2 Kombination von boden- und luftgestützten Stereobilddaten

Eine Kombination von bodengestützt mobil erfassten Bilddaten mit Luftbilddaten ist insbesondere für die Erfassung von Verkehrskorridoren aus unterschiedlichen Gründen interessant. Ein

[1] Stephan Nebiker, Hannes Eugster, Fachhochschule Nordwestschweiz, Institut Vermessung und Geoinformation, 4132 Muttenz, Schweiz; E-Mail: stephan.nebiker@fhnw.ch, hannes.eugster@fhnw.ch

[2] Kay Kamer, Hannes Eugster, iNovitas AG – Mobile Mapping Solutions, 4132 Muttenz, Schweiz; E-Mail: kay.kamer@inovitas.ch, hannes.eugster@inovitas.ch

wichtiger Grund ist die Schwierigkeit einer genauen und zuverlässigen direkten Georeferenzierung mobiler Sensoren und Plattformen im urbanen Raum, die gerade im Strassen- und Schienenraum auf Grund der oft stark eingeschränkten Satellitenabdeckung eine besondere Herausforderung darstellt. Im Falle des mobilen Laserscannings wurde diese schon verschiedentlich durch eine Stützung bzw. Integration der mobilen Laserscans in luftgestützte Laserscanningdaten angepackt. Bei mobil erfassten Stereobilddaten bieten sich indirekte bzw. integrierte Georeferenzierungsansätze mit Einbezug von Bildmessungen (vision updates) auf Passpunkte aus terrestrischer Bestimmung sowie auf solche aus Luftbildern an. Ein zweiter wichtiger Grund für den Einbezug von Luftbildern ist die dadurch ermöglichte komplementäre Sicht auf den Verkehrskorridor, u.a. zur Erfassung von Geoobjekten außerhalb des direkt einsehbaren Fahrbahnbereichs. Diese kombinierte Nutzung von boden- und luftgestützten 3D-Bilddaten erfordert wiederum eine exakte Co-Registrierung beider Bilddatensätze. Dieser Beitrag adressiert die folgenden Hauptfragen: wie und bis zu welchem Grad lässt sich die Georeferenzierungsgenauigkeit von mobil erfassten Stereobilddaten auch unter ungünstigen GNSS-Bedingungen steigern bzw. sicherstellen und bietet der Einbezug von hochaufgelösten Luftbildern aus einer zeitnahen Befliegung eine ökonomische Alternative zur Verwendung von terrestrisch bestimmten Passpunkten im Fahrbahnbereich?

3 Erfassungs- und Auswertesysteme

3.1 Bodengestützte Bilddaten: Stereovision Mobile Mapping System

Für die nachfolgenden Untersuchungen wurde das stereobildbasierte Mobile Mapping System des Instituts Vermessung und Geoinformation (IVGI) eingesetzt, welches seit 2009 entwickelt wird (vgl. Abb. 1). Das System besteht aus einem INS/GNSS-Navigationssystem (zum Zeitpunkt der Untersuchungen einem Applanix POS LV 210), welches zur direkten Georeferenzierung von bis zu fünf Stereo-Kamerasystemen genutzt wird. Als Bildsensoren kommen Industrie-Kameras mit wahlweise 11 MP oder 2 MP (FullHD) mit einer radiometrischen Auflösung von 12 Bits und einer Pixelgröße von mind. 7 µm zum Einsatz. Die Stereokameras können in unterschiedlichen Konfigurationen (Stereobasen, Blickrichtungen) auf eine stabile Plattform montiert werden. Je nach Projektanforderungen können die Sensoren mit bis zu 5 bzw. 30 fps betrieben werden, was zu Bilddatenmengen bis zu 1 TB pro Stunde führt.



Abb. 1: IVGI Mobile Mapping System



Abb. 2: Mittelformat-Luftbildkamerasysteme Leica RCD30

Für die Prozessierung und Nutzung der Stereobilddaten wurde im Rahmen des Forschungsprojekts SmartMobileMapping eine umfangreiche Prozesspipeline entwickelt und umgesetzt (NEBIKER et al., 2012). Diese umfasst die folgenden Teilprozesse: *Systemkalibrierung* mit spezifischen Verfahren für Multi-Stereokonfigurationen; *direkte oder integrierte Georeferenzierung* der Stereobildsequenzen (EUGSTER et. al., 2012); Stereobilddatenaufbereitung (BURKHARD et al., 2012); *Tiefenkartengenerierung mittels Dense Stereo Matching*, momentan unter Nutzung des Semi-Global Block Matching Algorithmus aus der OpenCV Library; *automatische Feature-Extraktion* wie bspw. automatische Kartierung von Strassenverkehrszeichen (CAVEGN & NEBIKER, 2012). Für die Nutzung der Stereobilddaten wurde ebenfalls im Rahmen des SmartMobileMapping-Projekts die vollständig webbasierte Architektur 3DCityTV® entwickelt (EUGSTER et al., 2013). Diese kombiniert eine Cloud-basierte Datenhaltung mit einem WebSDK, einem Software Development Kit auf HTML5- und WebGL-Basis. Die 3DCityTV® WebSDK erlaubt eine denkbar einfache und gleichzeitig genaue Nutzung der Bilddaten mittels 3D-Monoplotting in den gestreamten Bild- und Tiefenkarten.

3.2 Luftbilddaten: Leica RCD30

Die Luftbilddaten für die folgenden Untersuchungen wurden mit einer Mittelformatkamera Leica RCD30 erfasst (Abb. 2). Die RCD30 weist vier exakt co-registrierte Spektralkanäle (R,G,B und NIR) auf, welche durch ein gemeinsames Objekt erfasst und durch einen dichroitischen Strahlteiler (WAGNER, 2011) auf zwei Flächensensoren aufzeichnet werden. Die folgenden Merkmale machen die RCD30 besonders interessant für diese Art von hochaufgelösten Korridorvermessungen: ein einzelner Kamerakopf mit einer 60 MP-Auflösung für RGBN; eine zweiachsige mechanische Bildwanderungskompensation (FMC) und eine hohe Bildwiederholrate von min. 1 s pro Aufnahme, welche hohe Bildüberdeckungen bei geringen Flughöhen erlauben. Und zu guter Letzt: ein co-registrierter NIR-Kanal, der eine Reihe von automatischen Klassifikations- und Objektextraktionsaufgaben ermöglicht. Die Prozessierung der RCD30-Bilddaten umfasst im Wesentlichen die folgenden Hauptschritte: die *Sensorkalibrierung* – wahlweise mittels Kalibrierflug und Bündelausgleichung oder mittels Laborkalibrierung (TEMPELMANN et al., 2012); die *Generierung von verzerrungsfreien Multispektralbildern* mittels Leica FramePro; eine direkte oder integrierte *Georeferenzierung* sowie die *Extraktion dichter Tiefenkarten bzw. 3D-Punkwolken* mittels XPro DSM. Letztere basiert auf einer Weiterentwicklung bzw. Adaptierung des Semi-Global Matching-Ansatz von (HIRSCHMÜLLER, 2008).

4 Georeferenzierungsstrategien

Zur Georeferenzierung bzw. Co-Registrierung der luftgestützten und terrestrischen Bilddaten wurden verschiedene Strategien entwickelt und untersucht. Die wichtigsten zwei Strategien sind:

A) Integrierte Georeferenzierung der bodengestützten Stereobilddaten auf den Luftbildblock. Als Stützpunkte dienen dabei ausgewählte natürliche Punkte, die in beiden Bilddatensätzen sichtbar sind. Die 3D-Stützpunktkoordinaten werden dabei aus dem Luftbildblock bestimmt und anschliessend für die integrierte Georeferenzierung der bodengestützten Bilddaten verwendet.

B) Unabhängige indirekte und integrierte Georeferenzierung der luft- und bodengestützten Bilddaten auf einen gemeinsamen Bezugsrahmen. Dieser Bezugsrahmen kann durch unterschiedliche terrestrisch eingemessene Passpunktszenarien realisiert werden.

Für die Bestimmung der bildbasierten Beobachtungen in der integrierten GNSS/INS-Georeferenzierung der bodengestützten Plattform-Trajektorie wurden wiederum zwei unterschiedliche Lösungsansätze implementiert und untersucht:

1) *Bündelblock-basierte Updates:* Bestimmung der Projektionszentren der bodengestützten Sensoren mittels einer auf einer Passpunktgruppe gelagerten Bündelblockausgleichung (Abb. 3).

2) *Einzelpasspunkt-Updates:* Unabhängige Bestimmung von 3D-Koordinaten der Stützpunkte aus den Stereobilddaten (Abb. 4) und Einführung der Residuen als Koordinatenupdates in der integrierten Georeferenzierung.

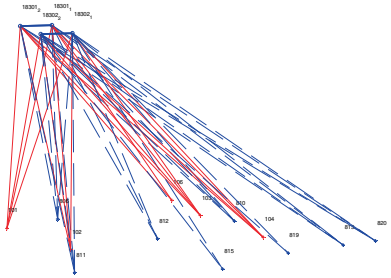


Abb. 3: Bündelblock-basierte Updates zur Bestimmung der äußeren Orientierung der MMS-Sensoren (blau fett ausgezogen: Stereobasis, rot ausgezogen: Passpunktbeobachtungen, blau gestrichelt: Verknüpfungspunktbeobachtungen)

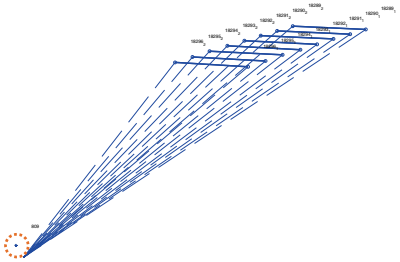


Abb. 4: Einzelpasspunkt-Update: 3D-Koordinatenbestimmung durch Multi-Image-Matching im Vergleich zu Passpunkt-Koordinaten (Punkt in gestricheltem Kreis)

5 Testprojekt und Datenerfassung

5.1 Untersuchungsgebiet und Referenzdaten

Die folgenden Untersuchungen zur integrierten Georeferenzierung von luft- und bodengestützten Stereobilddaten erfolgten am Beispiel eines Abschnitts der Autobahn A1 in einem urbanen Umfeld am westlichen Rand von Zürich. Das gesamte Projektgebiet umfasst einen Abschnitt von ca. 22 km Länge mit zwischen 3 und 5 Fahrspuren pro Fahrtrichtung. Das eigentliche Untersuchungsgebiet (Abb. 5) hat eine Ausdehnung von ca. 1.9 km x 0.4 km.

Als Referenzdaten wurden mittels GNSS-RTK-Vermessung und Tachymetrie verschiedene Passpunktgruppen eingemessen. Die Passpunkte für die integrierte Georeferenzierung der Luftbilder (Punkte Δ in Abb. 5) konnten optimal platziert und mittels GNSS-RTK doppelt bestimmt werden. Die mittl. Koordinatenfehler aus Doppelbestimmung bewegen sich in der Größenordnung von 0.5-1.0 cm in der Lage und von 1.0-1.7 cm in der Höhe (KAMER, 2012). Die Bestimmung der Passpunkte im Fahrbahnbereich (Punkte O in Abb. 5) erfolgte mittels

reflektorloser Tachymetrie und GNSS-Anschlussmessungen und gestaltete sich deutlich schwieriger, da im Projekt der Fahrbahnbereich nicht betreten werden konnte. Die resultierenden Passpunktgenauigkeiten liegen daher in einer Größenordnung von ca. 4 cm in der Lage und ca. 2-4 cm in der Höhe (KAMER, 2012).



Abb. 5: Untersuchungsgebiet entlang der Autobahn A1 am Westrand Zürichs (Perimeter: gelbe ausgezogene Linie; Δ Passpunkte Luftbilder; O Passpunkte MMS-Bilddaten) (Orthophoto © Swisstopo)

5.2 Bodengestützte Bilddatenerfassung und -aufbereitung

Die bodengestützte Bilddatenerfassung in beide Fahrtrichtungen erfolgte am 24. September 2011 mit dem eingangs beschriebenen Mobile Mapping System des IVGI. Zum Einsatz kamen ein vorwärtsblickendes 11-MP-Stereosystem und ein seitwärts blickendes FullHD-Stereosystem. Bei einer Bilddatenrate von 5 fps und einer Messgeschwindigkeit von ca. 80 km/h resultierte eine Bilddatendichte von einem Stereoframe pro 5-6 m. Die Stereobilddatenverarbeitung erfolgte mit der in Kap. 3.1 beschriebenen Prozessierungskette.



Abb. 6: Autobahnbrücke aus Perspektive MMS



Abb. 7: Identische Situation in Luftbilddaten (Aufnahmegeom. MMS: Dreieck Bildrand unten)

5.3 Luftgestützte Bilddatenerfassung

Derselbe Autobahnabschnitt wurde nur vier Tage später am 28. Sept. 2011 mit einer Leica RCD30-Kamera auf einem Pilatus Porter PC-6 beflogen. Die Befliegung erfolgte mit einem

NAG-D-Objektiv mit einer Brennweite von 53 mm auf einer Flughöhe von 400 m und mit einer Längsüberdeckung von 80%. Die durchschnittliche Bodenauflösung der RGBN-Bilddaten beträgt 5 cm. Die Bilddaten wurden mit Leica FramePro und Leica IPAS TC prozessiert und direkt georeferenziert. Die Luftbilder wurden zusätzlich mittels Bündelblockausgleichung in ORIMA unter Verwendung der oben aufgeführten Passpunkte indirekt georeferenziert. Die empirische Standardabweichung dieser Bündelblockausgleichung betrug 2.3 cm in X, 1.8 cm in Y und 3.4 cm in Z (Höhe) mit max. Restklaffen von je 5 cm in X und Y sowie 7 cm in Z (KAMER, 2012). Der Nutzen der komplementären Perspektiven der vorliegenden boden- und luftgestützten Bildaufnahmen auf den Straßenraum wird durch Abb. 6 und Abb. 7, die einen lageidentischen Ausschnitt derselben Brücke zeigen, eindrücklich illustriert.

6 Untersuchungen und Ergebnisse

6.1 Genauigkeit der direkt georeferenzierten MMS-Daten

Vorgängig zur Untersuchung unterschiedlicher integrierter Georeferenzierungsansätze wurde die Genauigkeit der direkt georeferenzierten (DG) Trajektorie bzw. der darauf aufbauenden Stereobilddaten untersucht. Dazu wurde auf der Basis der direkt georeferenzierten bodengestützten Stereobilddaten eine unabhängige 3D-Koordinatenbestimmung für eine Reihe von Punkten mit Hilfe eines Multi-Image-Matching Punktmesstools (HUBER, NEBIKER, & EUGSTER, 2011) in Analogie zu Abb. 4 durchgeführt. Als Referenzdaten dienten einerseits die oben aufgeführten terrestrisch bestimmten Passpunktkoordinaten und andererseits eine interaktive 3D-Koordinatenbestimmung aus den indirekt georeferenzierten Stereoluftbildern.

Tab. 1: Genauigkeit der direkt georeferenzierten (DG) MMS-Daten

Genauigkeit der DG MMS-Lösung aus Vergleich mit …	**Mittelwert** [cm]			**Standardabweichung** [cm]		
	ΔX	ΔY	ΔZ	X	Y	Z
… Kontrollpunkten aus Luftbildern	7.9	-8.2	5.2	11.1	8.3	5.6
… terrestrisch best. Kontrollpunkten	8.1	-3.8	5.8	10.4	4.1	4.1

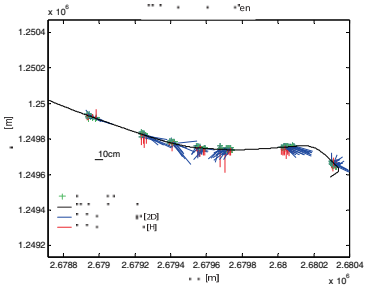


Abb. 8: Vergleich der DG MMS-Lösung mit Passpunkten aus Luftbildern

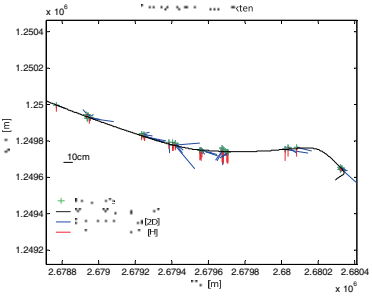


Abb. 9: Vergleich der DG MMS-Lösung mit terrestrisch bestimmten Passpunkten

Die Resultate in Tabelle 1 zeigen eine gute Übereinstimmung zwischen den zwei unabhängigen Vergleichen (Passpunkte Luftbild und terrestrisch) sowohl bezüglich systematischer als auch zufälliger Anteile. Die systematischen Lageoffsets betragen 8 cm in X, 4 cm bzw. 8 cm in Y und der Höhenoffset 5.5 cm. Die systematischen Abweichungen sind in Abb. 8 und Abb. 9 gut zu erkennen. Die direkt georeferenzierte Trajektorie weist somit einen systematischen Fehler von ca. 10 cm in der Lage und 5.5 cm in der Höhe auf.

6.2 Genauigkeit der integriert georeferenzierten MMS-Daten

In der Folge wurden die MMS-Trajektorie und damit auch die absolute Orientierung der Stereobilddaten mittels der in Kap. 4 eingeführten integrierten Georeferenzierungsansätze neu geschätzt. Für den Genauigkeitsnachweis in Tabelle 2 wurden auf der Basis dieser verschiedenen integriert geschätzten Trajektorien bzw. äußeren Bildorientierungen durch photogrammetrische Messungen in den mobil erfassten Stereobilddaten die 3D-Koordinaten der Passpunkte neu bestimmt und mit den Sollwerten verglichen. Eine grafische Darstellung der lage- und höhenmäßigen Koordinatendifferenzen findet sich in Abb. 10 und Abb. 11.

Tab. 2: Genauigkeit der integriert georeferenzierten MMS-Daten

Genauigkeit der integriert georeferenzierten MMS-Daten	**Mittelwert** [cm]			**Standardabweichung** [cm]		
	ΔX	ΔY	ΔZ	X	Y	Z
A) Passpunkte aus Luftbildblock						
A1) Bündelblock-basierte Updates	-2.3	-6.3	5.6	13.0	9.7	6.9
A2) Einzelpasspunkt-Updates	0.0	-2.5	-1.1	4.8	4.5	7.6
B) Terrestrische Passpunkte						
B1) Bündelblock-basierte Updates	5.4	-4.4	3.2	7.0	5.3	3.0
B2) Einzelpasspunkt-Updates	-1.9	-2.4	0.0	5.4	4.4	1.7

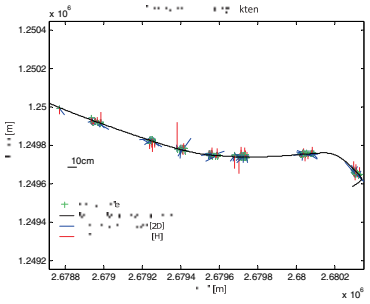


Abb. 10: Ergebnisse der Strategie A) Integrierte Georeferenzierung der bodengestützten Stereobilddaten mit Passpunkten aus Luftbildern (Restklaffen in Passpunkten)

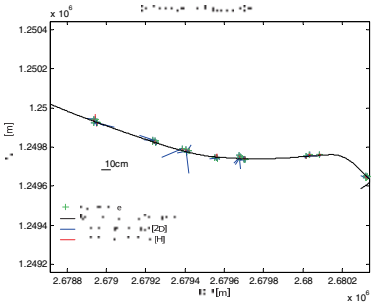


Abb. 11: Ergebnisse der Strategie B) Integrierte Georeferenzierung der bodengestützten Stereobilddaten mit terrestrisch bestimmten Passpunkten (Restklaffen in Passpunkten)

Die Ergebnisse in Tabelle 2 und in Abb. 10 und Abb. 11 zeigen Folgendes auf:

- Eine massive Reduktion der systematischen Lage- oder Höhenfehler bei Verwendung des Lösungsansatzes der Einzelpasspunkt-Updates – im vorliegenden Fall um etwa einen Faktor 4 – auf max. 2-2.5 cm in der Lage und 1 cm in der Höhe sowohl für Strategie A) als auch B).
- Deutlich höhere Standardabweichungen bei Trajektorienupdates mittels Bündelblockausgleichung, welche sich vor allem aus dem größeren Rauschen der Passpunktmessungen aus den Luftbildern (Strategie A) ergeben, und was sogar in einzelnen Fällen zu einer nicht konvergierenden Bündelblocklösung führte.
- Ein deutlich reduziertes Rauschen der Koordinatendifferenzen aus Strategie B) gegenüber derjenigen aus Strategie A) (vgl. Abb. 10 und Abb. 11). Der Verbleib einiger größerer Restklaffungen in der Projektmitte unterstreicht die Vermutung von suboptimal bestimmten terrestrischen Passpunkten im Fahrbahnbereich. Diese Vermutung wird durch die Ergebnisse von EUGSTER et al. (2012) in Projekten mit besseren Passpunktkoordinaten bekräftigt.

6.3 Untersuchung der abgeleiteten 3D-Punktwolken: Dichten und Genauigkeiten

Obwohl der Schwerpunkt der vorliegenden Untersuchungen in den erzielbaren Georeferenzierungs- und den damit verbundenen bildbasierten 3D-Punktmessgenauigkeiten lag, sollte auch die Qualität der abgeleiteten 3D-Punktwolken untersucht werden, da diese eine wertvolle Basis für verschiedene Auswertungen bieten. Dichte 3D-Punktwolken wurden sowohl aus den luftgestützten als auch aus den bodengestützten Bilddaten extrahiert (siehe Kap. 3). Abb. 12 zeigt den linken Stereopartner mit einer überlagerten rohen Tiefenkarte in der linken Bildhälfte. Eine nachbearbeitete Version dieser Tiefenkarte dient zudem als Basis für das 3D-Monoplotting mit Hilfe der 3DCityTV WebSDK. Für die folgenden Vergleiche wurden die texturierten Tiefenkarten aus den bodengestützten Stereobilddaten in den Objektraum projiziert.

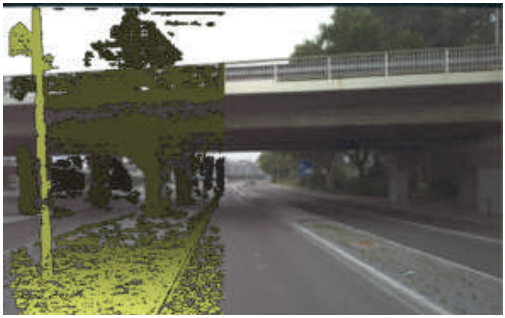


Abb. 12: Linkes Stereo-Normalbild mit überlagerter Tiefenkarte (linke Bildhälfte)

Für eine erste Genauigkeitsabschätzung der extrahierten luft- und bodengestützten 3D-Punktwolken wurden vier möglichst ebene Ausschnitte der Straßenoberfläche mit einer Fläche von je ca. 22 m^2 als Referenzebenen ausgewählt. Die Ergebnisse in Tabelle 3 zeigen die enorm hohe Punktdichte der bodengestützten Punktwolken mit über 1000 Punkten pro m^2, welche eine ideale Ergänzung zu den bereits dichten Oberflächenmodellen aus der Luftbildphotogrammetrie bieten. Interessant ist auch, dass das Rauschen der Oberflächemodelle aus beiden Verfahren in der

Größenordnung von ◘ 1 Pixel liegen, d.h. < 1 cm für die bodengestützten und < 5 cm für die luftgestützten Bilddaten (NEBIKER et al., 2012).

Tab. 3: Punktdichten und Abweichungen von Referenzebenen im Fahrbahnbereich

	Punktwolken aus boden-gestützten Bildaufnahmen	Punktwolken aus luft-gestützten Bildaufnahmen
Mittl. Punktdichte [Punkte / m^2]	1297	109
Mittlere Standardabweichung [m]	0.9 cm	4.5 cm
Max. Abweichung [m]	5.2 cm	16.7 cm

7 Fazit und Ausblick

In diesem Beitrag haben wir verschiedene integrierte Georeferenzierungsansätze vorgestellt, welche – insbesondere unter ungünstigen GNSS-Bedingungen – eine deutliche Genauigkeitssteigerung gegenüber der verbreiteten direkten Georeferenzierung ermöglichen und ganz allgemein die Sicherstellung vorgegebener Genauigkeitsanforderungen in bildbasierten Mobile Mapping Projekten gewährleisten sollen. Mit Blick auf die eingangs aufgeführten Forschungsfragen ergaben die Untersuchungen folgendes Ergebnis: Die systematischen Fehleranteile der direkten Georeferenzierung konnten mit beiden Lösungsstrategien, d.h. sowohl Strategie A mit Einbezug von Passpunkten aus Luftbildern als auch Strategie B mit Einbezug von terrestrisch bestimmten Passpunkten, deutlich reduziert werden auf ca. 2-2.5 cm in der Lage und 1 cm in der Höhe. Von den untersuchten Lösungsvarianten Bündelblock-basierten Updates und Einzelpasspunkt-Updates, erwies sich die Variante der Einzelpasspunkt-Updates als genauer und vor allem auch robuster. Unter Einbezug von Passpunkten aus Luftbildern konnten ebenfalls gute Georeferenzierungsgenauigkeiten von < 1 Pixel in der Lage und ca. 1.5 Pixel in der Höhe erzielt werden. Im Verlauf der Untersuchungen hat sich die limitierte Genauigkeit der verwendeten, terrestrisch bestimmten Passpunktkoordinaten im Fahrbahnbereich als Schwachpunkt herauskristallisiert. So liegen die Standardabweichungen der genausten Lösungsvariante B2 mit 4-5 cm in den Lagekoordinaten und 1.7 cm in der Höhe in der Größenordnung der Genauigkeit der Passpunkte. Laufende Untersuchungen mit einer exakten Bestimmung von Passpunkten im Fahrbahnbereich zeigen denn auch, dass mit dem vorliegenden Stereovision Mobile Mapping System und den vorgestellten integrierten Georeferenzierungsstrategien Lage- und Höhengenauigkeiten von je ca. 1 cm im praktischen Einsatz erreichbar sind. Falls beide Datensätze vorliegen, empfiehlt sich die Registrierung der bodengestützten Stereobilddaten auf Passpunkte aus dem Luftbildblock, da mit dieser Strategie auf kostspielige Sperrmaßnahmen zur Bestimmung von terrestrischen Passpunkten im Fahrbahnbereich verzichtet werden kann.

8 Dank

Wir bedanken uns bei der Kommission für Technologie und Innovation (KTI), der Förderagentur für Innovation des Schweizerischen Bundes, für die finanzielle Unterstützung des Forschungsprojekts SmartMobileMapping (Ref. 12561), beim Bundesamt für Strassen (ASTRA) für die Ermöglichung der Testfahrten im Rahmen der Machbarkeitsstudie '3D-Video' und bei der Firma Leica Geosystems für die Durchführung des RCD30-Bildfluges.

9 Literaturverzeichnis

BURKHARD, J., CAVEGN, S., BARMETTLER, A. & NEBIKER, S., 2012: Stereovision mobile mapping: system design and performance evaluation. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXIX-B5, 453-458.

CAVEGN, S. & NEBIKER, S., 2012: Automated 3D Road Sign Mapping with Stereovision-based Mobile Mapping exploiting Depth Information from Dense Stereo Matching. Photogrammetrie - Fernerkundung - Geoinformation, 2012(5), 631–645.

EUGSTER, H., GOTTSMANN, F., KÄSER, C., NEBIKER, S. & KOCH, R., 2013: 3DRoadTV-Service – Georeferenzierter 3D-Bilddatendienst für die Infrastrukturpflege und -verwaltung. Dreiländertagung der DGPF, OVG und SGPF. DGPF Tagungsband 22 / 2013.

EUGSTER, H., HUBER, F., NEBIKER, S. & GISI, A., 2012: Integrated georeferencing of stereo image sequences captured with a stereovision mobile mapping system – approaches and practical results. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXIX-B1, 309–314.

HIRSCHMÜLLER, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30, 328–341.

HUBER, F., NEBIKER, S. & EUGSTER, H., 2011: Image Sequence Processing in Stereovision Mobile Mapping – Steps towards Robust and Accurate Monoscopic 3D Measurements and Image-Based Georeferencing. In U. Stilla, F. Rottensteiner, H. Mayer, B. Jutzi, & M. Butenuth (Eds.), Photogrammetric Image Analysis (Vol. 6952, pp. 85–95). Springer Berlin / Heidelberg.

KAMER, K., 2012: Untersuchungen zur Integration von mobil erfassten terrestrischen Stereobilddaten und Luftbilddaten. Bachelor Thesis. FHNW University of Applied Sciences and Arts Northwestern Switzerland.

NEBIKER, S., CAVEGN, S., EUGSTER, H., LAEMMER, K., MARKRAM, J. & WAGNER, R., 2012: Fusion of airborne and terrestrial image-based 3d modelling for road infrastructure management – vision and first experiments. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXIX-B4, 79–84.

NOVAK, K., 1993: Real-Time Mapping Technology. International Archives of Photogrammetry and Remote Sensing (pp. 569–575).

SCHWARZ, K. P., MARTELL, H. E., EL-SHEIMY, N., LI, R., CHAPMAN, M. A. & COSANDIER, D., 1993: VIASAT - A mobile highway survey system of high accuracy. Vehicle Navigation and Information Systems Conference, 1993., Proceedings of the IEEE-IEE (pp. 476-481).

TEMPELMANN, U., BEISL, U., BOEHRER, N., HINSKEN, L. & RECKE, U., 2012: Creating virtually distortion-free multi-band frame images: laboratory and flight calibration techniques for the Leica RCD30. European Calibration and Orientation Workshop. Barcelona.

WAGNER, R., 2011: The Leica RCD30 Medium Format Camera: Imaging Revolution. In D. Fritsch (Ed.), Photogrammetric Week 2011 (pp. 89–95). Stuttgart: Wichmann.

Dynamische Erfassung linearer Infrastrukturen unter und über der Erde

MARTIN RUB[1], EMESE FARKAS[2] & ALBERT WIEDEMANN[3]

Zusammenfassung: Lineare Infrastrukturen wie Strassen, Eisenbahnen, Hochspannungsleitungen, etc. stellen einen beträchtlichen Wert dar. Ihr Umfang steigt, somit auch der Bedarf an Informationen für eine optimale Planung von Werterhaltungsmaßnahmen. Eine statische Erfassung ist wegen der Länge der Objekte sehr aufwändig, und sollte sich somit auf Verdachtsstellen begrenzen. Dynamische Erfassungsmethoden können diese Verdachtsstellen ermitteln und bereits mit wichtigen Kennzahlen versehen. Eine Übersicht dynamischen Erfassungsmethoden enthält die unterschiedlichen Blickwinkel, Technologien und Plattformen sowie deren Synergiepotential werden erläutert und anhand von Projekten illustriert. Vom luftgestützten Powerlinemapping bis zur vollständigen Erfassung der Tunnelinnenwände sind den Produkten fast keine Grenzen gesetzt.

1 Einleitung

Die Vermessung von Infrastruktur-Korridoren hat sich in den letzten 5 Jahren maßgebend gewandelt. Ein Trend zur flächendeckenden Erfassungstechnologie führt zur Ablösung von manuellen Einzelpunktaufnahmen. Dieser Wandel ist der Entwicklung von Mobile Mapping, Schrägbildkamerasystemen als auch der Verbesserung des luftgestützten Laserscannings und der Photogrammetrie zuzuschreiben.
Die Technologien wie Mobile Mapping werden heute bereits während der Ausschreibung vorgeschrieben. Dieser Trend gründet auf den Erkenntnissen und Vorteilen der verschiedenen Erfassungsmöglichkeiten. Es zeichnet sich aber auch bereits heute ab, dass sich während der Planungsphase von Unterhalts- und Bauprojekten höhere Auflösung, Genauigkeit und vollständigere Geodaten durchsetzten.

2 Sensoren, Navigation & Trägersysteme

Zur dynamischen Erfassung kommen unterschiedliche Sensoren zum Einsatz um das Objekt zu erfassen, weitere zur Bestimmung der Orientierung der Sensoren im Raum. Die Systeme werden auf unterschiedlichen Trägersystemen eingesetzt, die sich parallel zur Ausdehnungsrichtung der linearen Infrastruktur fortbewegen.

[1] Martin Rub, BSF Swissphoto, Dorfstr. 53, CH-8105 Regensdorf-Watt, martin.rub@bsf-swissphoto.com
[2] Emese Farkas, BSF Swissphoto, Dorfstr. 53, CH-8105 Regensdorf-Watt, emese.farkas@bsf-swissphoto.com
[3] Albert Wiedemann, BSF Swissphoto, Mittelstr. 7, D-12529 Schönefeld, albert.wiedemann@bsf-swissphoto.com

2.1 Sensoren

2.1.1 Kameras

Eine Basis-Technologie ist die Verwendung großformatiger digitaler Kamerasysteme. BSF Swissphoto verwendet hier verschiedene Digitale Kameras der UltraCam Serie, die hochauflösende digitale Luftbilder in einer Auflösung bis zu 3 cm erstellen. Neben hochauflösenden Systemen kommen auch Mittelformat- oder andere Digitalkameras zum Einsatz. Abhängig von der Distanz zum Objekt bzw. der Soll-Messfrequenz werden in dynamischen Anwendungen kalibrierte Kameras mit einer Auflösung von 2 bis 50 Mega Pixel eingesetzt. Diese haben Messfrequenzen von 5 – 20 Hz.

2.1.2 AOS – Aerial Oblique System

Das System besteht aus drei Mittelformat Kameras. Eine stellt Senkrechtaufnahmen her, die beiden anderen schräg eingesetzten rotieren um 90°. Somit werden mit einem Aufnahmezyklus vier Schrägbilder und zwei vertikale Aufnahmen erstellt (WIEDEMANN 2010).

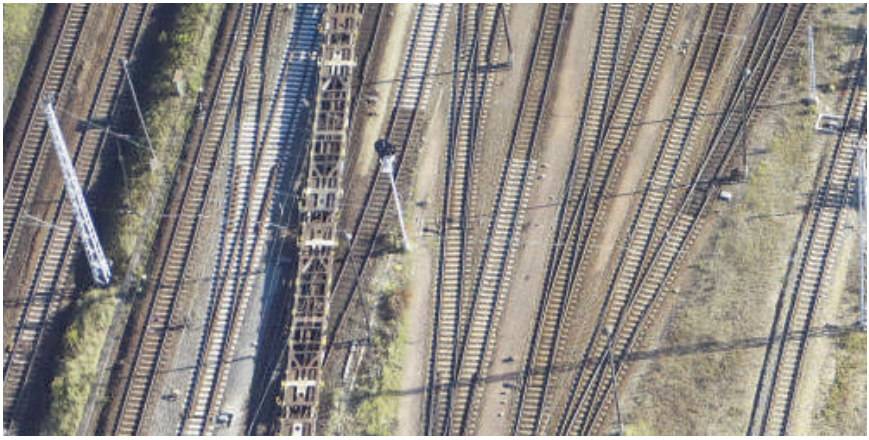


Abb.1: Bildausschnitt Schrägbildsystem AOS

2.1.3 Laserscanning

Laserscanning ist eine aktive Technologie zur Objektdokumentation durch Erfassung von Punktwolken mit Hilfe von Distanzmessung. Das Airborne Laserscanning ist eine effiziente Methode, um die Erdoberfläche aus der Luft schnell, genau und detailliert zu erfassen. Das Resultat der Lasermessung ist eine dreidimensionale Punktwolke mit Punktdichten zwischen 0.5 bis 200 Punkten/m². Die Systeme werden aus Flächenflugzeugen oder Helikoptern eingesetzt.
Das terrestrische Laserscanning besteht je nach Trägerplattform und deren Distanz zum Objekt aus Phasen- oder Pulsscannern. Im mobilen Einsatz werden Punktwolken von 20 – 50'000 Punkten / m² erzeugt.

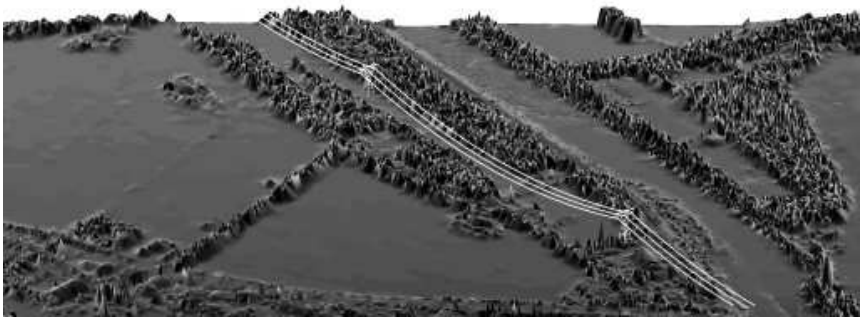


Abb. 2: Laserscanning Datensatz Powerline Mapping

Tab 1: Übersicht von Sensoren die auf dynamischen Plattformen für Korridormapping eingesetzt werden können.

AP Aerial Photogr.	AOS Aerial Oblique S.	ALS Airborne Laserscanning	HALS Handheld Airborne Laserscanning	TLS Terrestrial Lasersc.	DC Digital Cameras
					
Aufnahmerichtung/ Öffnungswinkel					
nadir	nadir, schräg	nadir	nadir/60°	beliebig/ 360°	beliebieg/
Reichweite					
			Reichweite	<100 m	
Auflösung/Punktdichte am Objekt					
3 cm @ 500 m Flughöhe	5 cm @ 350 m Flughöhe	45 cm Raster (4.5 Punkte/m^2)			
Genauigkeit					
cm-dm	cm	cm	cm	mm	cm
Messfrequenz					
0,5 Hz	0,3 Hz	166 kHz	100 kHz	1016 kHz	10-30 Hz

2.1.4 Sonstige

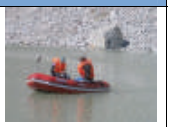
Weitere Sensoren können Infrarot bzw. Kameras anderer Aufnahmespektren sein. Beim ALS oder TLS bieten sich blaue Laser an um Bathymetrie zu betreiben. Um Korridorvermessung entlang von Flüssen zu betreiben kommen nach wie vor Echolote zum Einsatz. Thermalkameras können zur Überwachung von Fernwärmesystemen verwendet werden.

2.2 Navigation

Unter Navigation wird die Lösung der eindeutigen Trajektorie im Raum verstanden. Diese führt zu georeferenzierten Daten, die im Korridormapping in 2 Konzepte unterteilt werden können: A) Navigation basierend auf GNSS, B) Systeme die ohne Satellitennavigation funktionieren. Systeme der Gruppe A) stehen oftmals in Kombination mit GNSS- und einer IMU-Sensorik im Einsatz. Diese werden vorwiegend auf luftgestützten Plattformen verbaut. Der Satellitenempfang erlaubt diesen Systemen eine direkte Georeferenzierung und bedingt lediglich einer GNSS-Referenzstation und Kontrollmessungen. Der zweite Fall gilt insbesondere für terrestrische Systeme, die auf einem Auto, Boot oder Gleiswagen mitgeführt werden. Diese müssen mit Signalstörungen, Abschattungen (Bsp. städtische Gebiete, Berge, etc.) bis hin zum totalen Ausbleiben von Satellitensignalen (Tunnel) auskommen. Systeme der Gruppe B) haben sehr unterschiedliche Navigationskonzepte. So genannte Mobile Mapping Systeme umfassen oftmals eine GNSS-, IMU- und Wegsensoren. Die Georeferenzierung im Bereich von wenigen mm bis hin zu 1 m kann jedoch nur mittels bekannten Passpunkten erreicht werden.

2.3 Trägersysteme

Tab 2: Plattformen zur dynamischen Korridorerfassung.

Flieger	Helikopter	Auto	Bahn	Boot
				
Aufzunehmende Plattformen				
AP, AOS, ALS VK	AP, AOS, ALS HALS, TLS, DC	TLS, DC	TLS, DC	TLS, DC
Erfassungsgeschwindigkeit [km/h]				
180-300	50 - 200	< 80	< 80	< 20
Max. Missionsszeit am Stück [h]				
6.5	3	2	2	2

Auf Unmanned Aerial Vehicles (UAVs) wird in dieser Übersicht bewusst nicht eingegangen, da sie für Korridor Mapping von mehreren Kilometern Länge noch nicht geeignet und zugelassen sind.

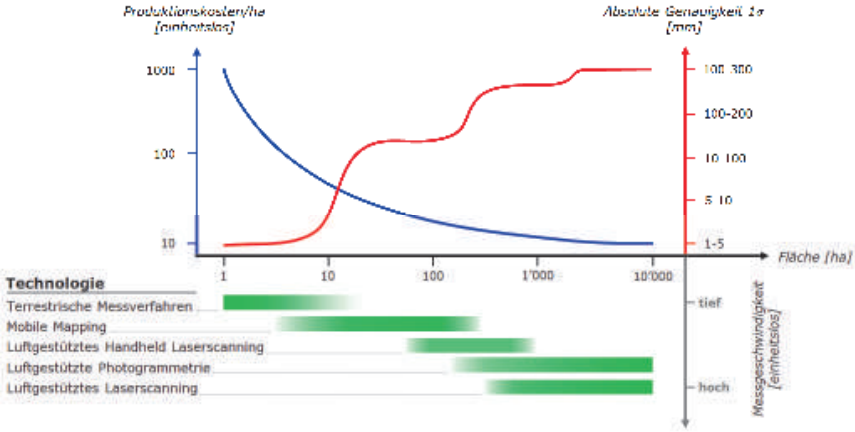


Abb. 3: Vergleich verschiedener mobilen Erfassungsmethoden in Hinblick auf deren Produktionskosten, Genauigkeit und Messgeschwindigkeit.

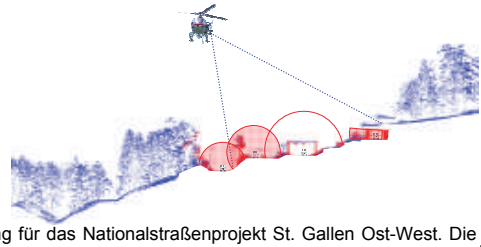


Abb. 4: Korridormapping für das Nationalstraßenprojekt St. Gallen Ost-West. Die erhobenen Laser- und Bilddaten wurden mittels AP und ALS erfasst. Sämtliche Strassen, Tunnel und Überführungen wurden weiter mit Mobile Mapping aufgenommen. Für steile Böschungen würde sich auch ein HALS anbieten.

2.4 Vergleich

Dank der Weiterentwicklung hochauflösender Sensorik und Navigationssystemen konnten die meisten Technologien ihr Anwendungsfeld erweitern. Die luftgestützten Systeme bewähren sich nach wie vor um in kurzer Zeit große Flächen zu erfassen. Diese Aufnahmeverfahren können

weitgehend auf Standardprozesse zurückgreifen. Terrestrische Mobile Mapping Systeme haben in den letzten 5 Jahren auf Seiten der Hardware viele Fortschritte gemacht. Dies betrifft vor allem Genauigkeit und Prozessierung. Dadurch findet immer mehr eine Ablösung von stationären Messungen statt, was aber nicht zwingend eine Reduktion der erforderlichen Arbeitszeit bedeutet. Die Arbeiten haben sich deutlich vom Feld ins Büro verlagert und bedürfen noch großer Fortschritte in den standardisierten Prozessen. Aus heutiger Sicht konnte Mobile Mapping die größten Fortschritte erzielen und ist eine wichtige Ergänzung bei der Erfassung von linearer Infrastruktur.

3 Synergie

Die Synergie der verschiedenen Technologien liegt vorwiegend in der Flexibilität verschiedener Blickwinkel aus der Luft oder vom Boden. Weiter können mit der hohen Punktdichte aus einer Straßenbefahrung alle wichtigen Elemente vektorisiert werden. Diese genügen höchsten Genauigkeitsansprüchen. Nicht alle Korridore benötigen über den gesamten Querschnitt dieselbe Punktauflösung. Die vorgestellten Technologien können von Zentimeter- bis zur 1 m-Punktdichte erzeugen. Gebiete weiter entfernt zum Korridorobjekt haben oftmals geringere Ansprüche. Aus diesem Grund bieten AP, ALS und HALS die richtigen Technologien. Geringere Punktdichte und absolute Genauigkeit sind somit eine kostengünstige Methode um große Flächen zu erfassen. Zudem ergänzen sich Bild- und Laserdaten vor allem im Straßenbereich wenn es um die Objektattributierung geht. Die Korridorausdehnung und die Objekte von Interesse beschreiben schließlich den Einsatz der verschiedenen Technologien. Bei Korridoren mit einer Länge über 5 km kann ein Einsatz von verschiedenen Technologien eine sinnvolle Ergänzung und Effizienzsteigerung bedeuten.

Tab. 3: Übersicht zur Eignung der Erfassungsmethoden und – Kombinationen für Korridorprojekte.

					
Anforderungen					
Korridorbreite [m]	20-50	< 20	<50	100-1000	100-1000
Detaillierungsgrad	hoch	hoch	mittel	hoch	hoch
Genauigkeit	hoch	hoch	hoch	mittel	gering
Geschwindigkeit	hoch	hoch	gering	mittel	mittel
Technologie					
AP	✓✓✓		✓✓✓	✓✓✓	✓✓✓
AOS	✓			✓	
Handheld ALS	✓✓✓		✓✓✓	✓✓✓	✓✓✓
MMS	✓✓✓	✓✓✓	✓✓✓	✓	✓

4 Projekte

Anhand von zwei Projektbeispielen zeigen wir Technologien und deren erweiterten Produkte. Im Straßenprojekt werden die Stärken der unterschiedlichen Technologien unter die Lupe genommen.

4.1 Gotthard Basistunnel, AlpTransit Gotthard AG (ATG)

Die ATG beauftragte die BSF Swissphoto als Teil des Vermessungskonsortiums mit der Befahrung der Tunnelabschnitte Erstfeld West und Sedrun Süd. Als Vorgabe ist ein hochauflösendes Laserscanning vor Übergabe des Rohbaus an die Bahntechnik mit einer Punktdichte von 5 mm und einer Genauigkeit (1 σ) von 4 mm bzw. 10 mm Toleranz durchzuführen. Auf dem Mobile Mapping System der 3D Mapping Solutions GmbH sind 2 der neuesten Laser Profiler 9012 von Z+F, eine IMU und Wegsensoren zur hochpräzisen Trajektorienbestimmung zum Einsatz gekommen. Zwecks Georeferenzierung wurden alle 50 m beidseitig Fixpunkte mit Zieltafeln signalisiert. Die daraus gewonnenen Resultate umfassen das Bestimmen von Koordinaten vordefinierter Profilpunkte der Innenschale im Profilabstand von 10 m, Differenzplänen des aktuellen Tunnels gegenüber der Soll-Geometrie und die Bestimmung der Schachtkoordinaten. Des Weiteren konnte gezeigt werden, dass mit der dichten Punktwolke auch Orthophotos erzeugt werden können. Diese können als Grundlage für die Tunnelinventarisierung oder das Kartieren von Wassereintritten bzw. größeren Rissen herangezogen werden.

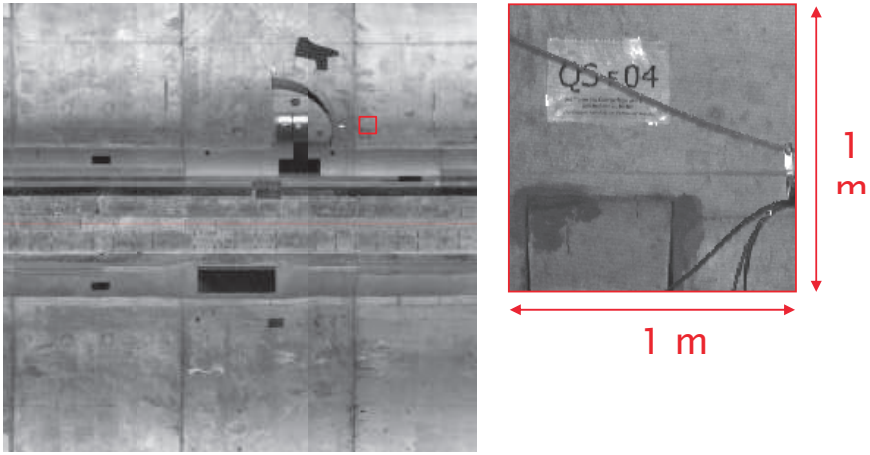


Abb. 5: Laserscanning-Orthophoto des Eisenbahntunnels nach dem Rohbau. Auflösung am Objekt beträgt 5 x 5 mm, eingefärbt mit den Intensitätswerten. Abwicklung ist mittig an Tunneldecke geschnitten.

4.2 Autobahn –St. Gallen, Bundesamt für Strassen (ASTRA)

Für das Erhaltungskonzept des Nationalstraßenabschnitts St. Gallen Ost-West von 11.5 km waren ein Fixpunktnetz zu realisieren, Datenerfassung mittels Mobile Mapping und deren Auswertung zu einem vektorisierten Objektkatalog (Straßenmöblierung), Erarbeiten des digitalen Terrainmodells (DTM) und Profilen, Ergänzungsaufnahmen an Kunstbauten und unterstützende Vermessungsarbeit in den laufenden Jahren durchzuführen.
Bereits vor der Beauftragung wurden ALS-Daten erfasst, diese dienten in erweiterten Gebieten den geologischen Untersuchungen. In einem weiteren Schritt wurde ein Grundlagenetz von über 500 neuen Fix- und Passpunkten realisiert und gemessen. Letztere wurden zur Einpassung der mobilen Laser- und Bilddatenerfassung genutzt. Schlüsseltechnologie zur Erfassung und Auswertung von 11.5 km Autobahn und 3 Tunnels war das Mobile Mapping System. Mit 55 km/h wurde bei laufendem Verkehr jede Fahrspur innerhalb von zwei halben Tagen erfasst. Aus den Laser- und Bilddaten und den manuellen Ergänzungsmessungen wurde ein digitaler Katalog mit über 40 verschiedenen Objektklassen (OK & UK Böschung, Fahrbahnrand, Verkehrssignale, Entwässerungsschacht, Fahrbahnachse, etc.) erstellt. Mobile Mapping vermag zwar Sperrzeiten und Arbeitsgefahren zu minimieren, doch das Arbeitsvolumen hat sich eindeutig hin zum Büro verlagert. Das Projekt hat allen Beteiligten aufgezeigt, dass aus Laserscanning-Aufnahmen jederzeit und fast beliebige geometrische Daten abgeleitet werden können.

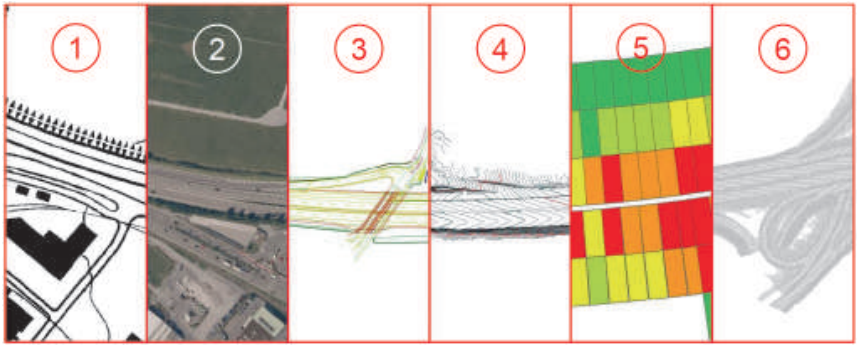


Abb. 5: Prozessierungs- und mögliche Produkteschritte: (1) Mittels Grundlagenplänen wird die Befliegung und Befahrung des Korridors geplant. (2) Die gesamte Breite wir mittels AP und ALS beflogen. (3) Sämtliche befahrbare Strassen werden mittels Mobile Mapping System erfasst und vektorisierte Pläne erstellt. (4) Daraus lassen sich Höhenmodelle in Form von Höhenlinien und Bruchkanten ableiten. (5) Die Laserdaten des Mobile Mappings werden für die Ebenheitsprüfung nach VSS-Norm herangezogen und je nach Zustandklasse eingefärbt. (5) Weitere Form der Geländedarstellung ist das triangulierte Netz mit Bruchkantenintegration

Die Erfahrung hat auch gezeigt, dass in Zukunft weiter an der Effizienzsteigerung auf Seiten Auswertung gearbeitet werden muss. Damit verbundenen sind Kostenersparnissen und Zeitgewinne für den Kunden. Durch diese Entwicklung werden periodische Infrastruktur-

Kontrollen erschwinglich und zu einem Bedürfnis von Straßenmeistereien, Tiefbauämter, Stromnetz- und Bahnbetreibern.

5 Ausblick

Durch die schnelleren Erfassungsmethoden und deren sinkenden Kosten wird die Nachfrage nach aktuelleren und hochwertigeren Geodaten weiter wachsen. Dies wiederum stellt neue Herausforderungen an Vermessungsingenieure beim Verständnis der Technologien und deren Einsatz für verschiedenste Ansprüche. Die größten Herausforderungen werden im Bereich von effizienten Auswertungsprozessen gesehen. Bereits heute findet eine Verlagerung der Arbeiten vom Feld ins Büro statt. Die Auswertungsprozesse konnten bislang nicht mit den Entwicklungen der Erfassungsmethoden mithalten. Die immer umfangreicheren Datenmengen stellen eine große Herausforderung an die Auswertungsprozesse betreffend Geschwindigkeit und Automatisierung. Auf der anderen Seite liegt die Produktvielfalt, die durch die Geodaten-Industrie aktiv gestaltet werden kann. Geländemodelle und Orthophotos sind nach wie vor ein großer Markt aber es muss an weiteren Anwendungsbereichen mit abgeleiteten Produkten gesucht werden wie beispielsweise:

- Komplette 3D Korridor-/Stadtmodelle
- Wärmebildkartierung (luftgestützt, terrestrisch)
- Bewertung Straßenbelag mittels Laserscanning und Bildkameras
- Risskartierung von Tunnelwänden
- Lärmschutzmodelle
- Vegetationsanalysen
- Sichtbarkeitsanalysen Fußgänger

Die Technologien verbinden folgende Vorteile: geringere Sperrzeiten, Gefahrenminimierung durch berührungslose Methoden, kürzere Produktionszeiten bei weiteren Produkten, kürzere Produktionszeiten bei Nachforderungen, Anwendungsvielfalt. Probleme ergeben sich durch die erforderliche Prozessierungszeit, die enormen Datenmengen, die Notwendigkeit verschiedene Daten zu integrieren und zu kontrollieren.

6 Literaturverzeichnis

STENGELE, R. 2009: Geodaten aus der Luft: aktueller Status, Trends undAusblick, September 2009: Geomatik Schweiz

WIEDEMANN, A. 2011: Geometrisches Potential von Schrägbildern aus dem System AOS. DGPF-Jahrestagung 2011

Referenzorientierung für Bilddaten aus Mehrkopfkamerasystemen

ALEXANDER WIEDEN[1] & KARSTEN STEBNER[2]

Zusammenfassung: Die Referenzorientierung kennzeichnet die Übertragung der äußeren Orientierung eines photogrammetrisch auswertbaren Sensors auf einen anderen simultan agierenden optischen Sensor. Auf diese Weise können Bilddaten aus Kamerasystemen orientiert werden, für die nicht jeder Einzelsensor zwangsläufig den photogrammetrischen Minimalanforderungen entsprechen muss. In diesem Beitrag wird die Referenzorientierung auf den Datensatz des Mehrkopfkamerasystems MACS 50/70 angewendet. Dabei wird das Potential des Verfahrens zur genauen Bestimmung der äußeren Orientierung für die Bilddaten der Einzelsensoren untersucht und gezeigt, dass sich die Referenzorientierung für den verwendeten Bilddatensatz bewährt.

1 Einleitung

Aufgrund verschiedener Aufgabenstellung im Bereich der Fernerkundung können Luftbildkameras nur selten vollständig den Ansprüchen der Auftraggeber gerecht werden. Gerade in der heutigen Zeit verlangt der Markt nach immer spezielleren Sensoren, welche Daten in hoher Qualität generieren sollen. Insbesondere die genaue Verortung und die Extraktion von Geoinformationen spielen eine große Rolle. Die Einrichtung *Optische Informationssysteme* des *Deutschen Zentrums für Luft- und Raumfahrt e.V. (DLR)* hat es sich daher unter anderem zur Aufgabe gemacht, optische Sensorsysteme zu speziellen Anwendungen und Fragestellungen zu konzipieren. Eine große Herausforderung stellt dabei z.B. die photogrammetrische Auswertung für Bilddaten aus Oblique-Sensoren, sehr schmalwinkliger Kameras, exzentrischer Sensorflächen aber auch thermaler Sensoren dar. Als Repräsentant für Luftbildkamerasystem mit niedrigen Öffnungswinkeln steht das MACS 50/70 (LEHMANN et al., 2011) des DLR. Neben einer kurzen Beschreibung dieses Sensorsystems in Abschnitt 2 und sich daraus ergebender Problemstellungen für die photogrammetrische Auswertung, werden im Kapitel 3 Luftbilddaten aus einer Testfeldbefliegung mit dem MACS 50/70 zum einen per Aerotriangulation und zum anderen durch Referenzorientierung prozessiert. Eine Genauigkeitsanalyse beider Orientierungsverfahren u.a. auf Basis abgeleiteter digitaler Oberflächenmodelle (DOM) ist Bestandteil des 4. Abschnittes.

1) Alexander Wieden, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Einrichtung Optische Informationssysteme, Sensorkonzepte und Anwendungen, Rutherfordstraße 2, 12489 Berlin, Email: alexander.wieden@dlr.de
2) Karsten Stebner, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Einrichtung Optische Informationssysteme Sensorkonzepte und Anwendungen, Rutherfordstraße 2, 12489 Berlin, Email: karsten.stebner@dlr.de

2 Sensorsystem

Das Multisensorsystem MACS 50/70 besteht aus drei Kameramodulen – der GeoCam mit 50 mm Brennweite und den hochauflösenden Kameras DOMCam 1 und DOMCam 2 mit jeweils 70 mm Brennweite. Die objektseitigen Bildfelder der DOMCam 1/2, die sich minimal überlappen, bilden dabei den Abbildungsbereich der GeoCam ab. Um das gesamte Bildfeld der GeoCam in den hochauflösenden DOMCam-Modulen zu erfassen und eine Maßstabsänderung durch Verkippung der Sensoren zu vermeiden, sind beide Module mit vershifteten Objektiven ausgestattet. Die Konsequenz bilden exzentrische Bildhauptpunktlagen für diese Kameras. Die nachfolgende Abbildung 1 kennzeichnet den grundsätzlichen Aufbau des MACS 50/70 und die sich daraus ergebenden Footprints der Einzelsensoren. Durch die vorliegende Bildhauptpunktexzentrität entstehen mehrere Probleme bei der photogrammetrischen Auswertung bzw. für die separate Triangulation der DOMCam-Module. Für einen gängigen Bildflug (z.B. 60/30) ist für die GeoCam, eine stabile geometrische Verknüpfung innerhalb und zwischen einzelnen Flugstreifen gewährleistet, während dies für die jeweiligen Bildstreifen der zwei hochauflösenden Kameras nicht der Fall ist. Wie die Abbildung 2 zeigt, entstehen bei einer klassischen photogrammetrischen Aufnahmekonstellation einseitige Bildanschlüsse. Die Aufnahmen einer jeweiligen DOMCam können aufgrund der Aufnahmefrequenz in Flugrichtung zwar ausreichend verknüpft werden, jedoch bewirkt der einseitige Bildanschluss keine geometrische Stabilisierung der Bildverbände quer zur Flugrichtung. Die in einer Bündelblockausgleichung auftretenden Restfehler werden daher in eine Richtung oder in den somit unterbestimmten Rollwinkel minimiert.

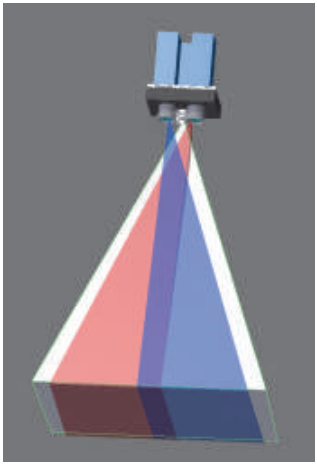


Abb. 2: Footprint MACS 50/70

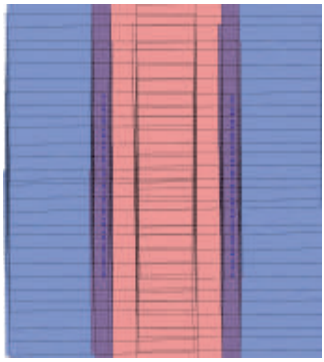


Abb. 1: Einseitiger Bildanschluss für einzelne DOMCam-Module des MACS 50/70

Sehr niedrige Öffnungswinkel (vgl. Tab. 1) sowie die Anordnung der DOMCam-Module im Hochformat bezogen auf die Flugrichtung, ziehen eine enorm schleifende Schnittgeometrie insbesondere in Querüberlappung nach sich. Erhöhte Fehlerellipsen in Vertikalrichtung resultieren und führen dazu, dass die Güte der ermittelten Projektionszentren in Vertikalrichtung innerhalb einer Aerotriangulation maßgeblich von der Genauigkeit der Näherungswerte der

äußeren Orientierung sowie gemessenen Referenzpunkten abhängig wird. Einen Überblick zu den einzelnen Kameramodulen des MACS 50/70 gibt Tabelle 1.

Tab. 1: Eigenschaften der Kameramodule des MACS 50/70

Sensor	Ck	Alpha/2	Beta/2	Bildhauptpunkt	Auflösungsfaktoren
GeoCam	50mm	19.8°	13.5°	zentrisch	1.4
DOMCam 1	70mm	14.4°	9.7°	exzentrisch	1
DOMCam 2	70mm	14.4°	9.7°	exzentrisch	1

3 Umsetzung

Folgende Untersuchungen basieren auf einen Bildflug über Berlin Adlershof, der im Frühjahr 2011 mit dem MACS 50/70 des DLR stattfand. Der Bildflug umfasste sechs parallele sowie zwei kreuzende Flugstreifen mit alternierender Flugrichtung und erfolgte mit einer Überlappung 80/60 in etwa 700 m über Grund. Damit stellen sich nominale Bodenauflösungen der GeoCam von 12.5 cm und der übrigen Kameras von etwa 9 cm ein. Insgesamt standen damit pro Kamera 365 Bilder zur Auswertung zur Verfügung. Die Abbildung 3 zeigt die Footprints aller GeoCam-Aufnahmen sowie tachymetrisch erfasste Passpunkte, von denen rund 30 innerhalb der Bilddatenorientierung Anwendung fanden. Insgesamt liegen für das aufgenommene Gebiet rund 400 natürliche Passpunkte mit einer Lagegenauigkeit von 2 cm und einer Höhengenauigkeit von 3 cm vor. Neben Fahrbahnmarkierungen, Pflasterwechsel und runden Schachtdeckeln sind ebenso höherliegende Passpunkte auf Gebäuden vorhanden.

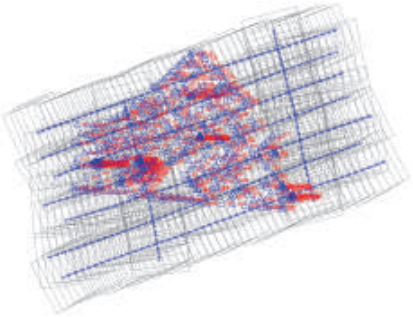


Abb. 3: Flugstreifen und Referenzfeld des Testfluges mit MACS 50/70

Zunächst wurden die Aufnahmen der GeoCam orientiert. Die erreichten Genauigkeiten der Aerotriangulation sind in Tabelle 2 dargestellt.

Tab. 2: Genauigkeiten der GeoCam-Orientierung

Sensor	East [m]	North [m]	Height [m]	Omega [gon]	Phi [gon]	Kappa [gon]
GeoCam	0.042	0.040	0.051	0.0048	0.0053	0.0007

Anschließend wurde die äußere Orientierung der GeoCam als Zwangsparameter mit entsprechender Gewichtung der Tabelle 2 für eine Aerotriangulation der Bilddaten der

DOMCam 1 und 2 eingeführt. Aufgrund der beschriebenen Einflussgrößen, in Bezug auf die geometrische Stabilität der Bildverbände, muss die Triangulation der exzentrischen Kameramodule durch die bessere geometrische Aufnahmekonstellation der GeoCam kompensiert werden, die daher auch als Übersichtskamera oder Orientierungsmodul verstanden wird. Die vorhandenen Leverarms zwischen den Projektionszentren aller drei Kameras konnten der CAD-Zeichnung zur Projektierung des MACS 50/70 näherungsweise entnommen werden und wurden innerhalb der Bilddatentriangulation der DOMCam-Module berücksichtigt. Weiter wurden konstante Boresight Misalignments zwischen der GeoCam und den DOMCam-Modulen bestimmt. Die Grundlage hierbei bildet die Minimierung der Winkelfehler über alle Bilder einer jeweiligen Kamera. Unter Berücksichtigung aller gegenseitigen Lage- und Winkeldifferenzen der Projektionszentren ergeben sich nach der Bündelblockausgleichung für DOMCam-Daten konstante Rest-Boresight-Misalignments unter einem Milligon (vgl. Tab. 3). Die Tabelle 4 kennzeichnet die Standardabweichungen der beschriebenen Aerotriangulation.

Tab. 3: Restkonstanter Anteil von Winkeldifferenzen zw. GeoCam und DOMCam-Module nach Aerotriangulation mit Zwangsbedingungen

Sensor	Omega [gon]	Phi [gon]	Kappa [gon]
DOMCam 1	0.0002	-0.0008	0.0000
DOMCam 2	0.0004	-0.0002	-0.0003

Tab. 4: Genauigkeiten der äußeren Orientierung für DOMCam-Aufnahmen durch Aerotriangulation mit Zwangsbedingungen

Sensor	East [m]	North [m]	Height [m]	Omega [gon]	Phi [gon]	Kappa [gon]
DOMCam 1	0.047	0.042	0.036	0.0040	0.0040	0.0030
DOMCam 2	0.025	0.028	0.038	0.0030	0.0030	0.0070

Sodann wurden die Bilddaten der hochauflösenden Sensoren erneut mit Hilfe der Referenzverfahren orientiert. Dazu wurde abermals die äußere Orientierung der Übersichtskamera verwendet und unter Beachtung der Leverarms und Boresight-Misalignments auf die DOMCam-Module übertragen. Eine nochmalige Bündelblockausgleichung sieht die Referenzorientierung nicht vor. Zur ersten Genauigkeitsabschätzung wurde dennoch zusätzlich eine Aerotriangulation durchgeführt, für die sämtliche Näherungswerte der äußeren Orientierung als fehlerfrei angenommen wurden. Dabei sank im Ergebnis weder die Bildpunktmessgenauigkeit a posteriori noch ergaben sich Widersprüche für die Genauigkeit (vgl. Tab. 5 und 6) der äußere Orientierung der Bilddaten.

Tab. 5: Genauigkeiten der äußeren Orientierung für DOMCam-Aufnahmen durch Referenzorientierung

Sensor	East [m]	North [m]	Height [m]	Omega [gon]	Phi [gon]	Kappa [gon]
DOMCam 1	0.000	0.000	0.000	0.0020	0.0010	0.0060
DOMCam 2	0.000	0.000	0.000	0.0020	0.0010	0.0040

Tab. 6: Restkonstante Anteile von Winkeldifferenzen zw. GeoCam und DOMCam-Module nach Referenzorientierung

Sensor	Omega [gon]	Phi [gon]	Kappa [gon]
DOMCam 1	0.0009	-0.0008	0.0001
DOMCam 2	0.0009	0.0001	-0.0005

4 Auswertung und Analyse

Damit lagen für beide hochauflösenden Kameras je zwei Orientierungen vor – eine aus direkter Referenzorientierung und eine aus gewichteter Aerotriangulation. Um die Güte der Referenzorientierung zu beurteilen behalf man sich zunächst der Epipolargeometrie. Der Verlauf der Epipolarlinie kennzeichnet dabei visuell die Genauigkeit der Orientierung. Durch das Messen eines Bildpunktes im Basisbild (vgl. Abb. 4, oben links) und einem weiteren Bild, sollten bei einer gelungenen Orientierung die ableitbaren Epipolarlinien der übrigen Bilder durch den identischen Punkt verlaufen. Wie die Abbildung 4 zeigt, trifft das für den referenzorientierten Datensatz zu. Die sich einstellenden Epipolarlinien verlaufen subpixelgenau durch den in allen Bildern abgebildeten Firstpunkt. Mit Hilfe der Epipolargeometrie konnte eine subpixelgenaue Referenzorientierung stichprobenartig über den gesamten Datensatz nachgewiesen werden.

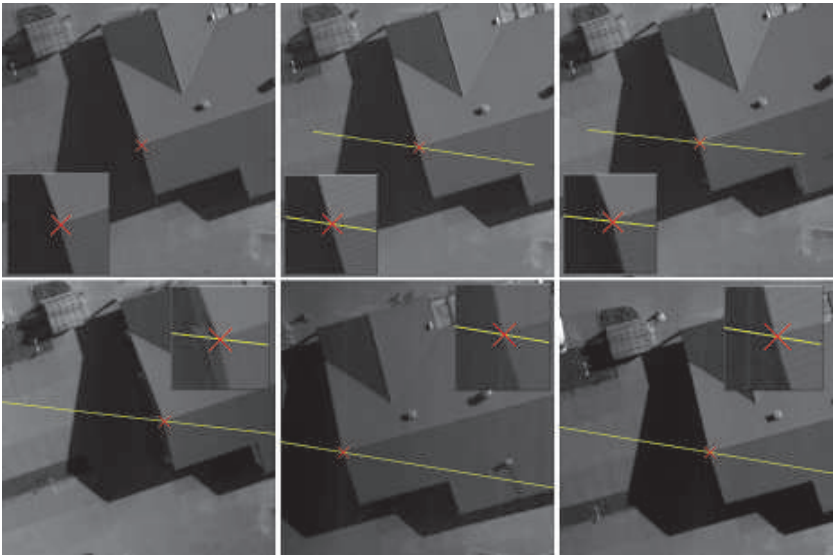


Abb. 4: Epipolargeometrie für Bilddaten der DOMCam 1 nach Referenzorientierung

Um die Fragen zu klären, ob die Referenzorientierung nun zur Auswertung des Sensorsystem MACS 50/70 geeignet ist bzw. ein Äquivalent zur gängigen Methode der Bündelblockausgleichung darstellt, wurden für beide vorliegenden Orientierungen DOMs mit jeweils 10 cm Rasterweite mittels Semiglobal-Matching (HIRSCHMÜLLER, 2008) bestimmt. Unzulänglichkeiten in der Bilddatenorientierung werden durch die Berechnung eines DOMs unmittelbar sichtbar. So erfolgte zunächst die visuelle Interpretation beider generierten DOMs (vgl. Abb. 5). Es wird deutlich, dass die hochauflösenden DOMs vorhandene Geometrien vollständig abbilden. So werden Gebäude und insbesondere Dächer sowie Dachaufbauten (vgl. Abb. 5 und 6, Detail) deutlich aufgelöst. Weiter bilden sich homogene Flächen wie Straßen oder Flachdächer sehr homogen und ohne Höhenversätze ab. Einzig im Bereich der Vegetation können Artefakte bzw. Fehlmatchings detektiert werden, was jedoch nicht im Einfluss der angewendeten Methode liegt. Aufgrund optisch konformer DOMs, wird an dieser Stelle darauf verzichtet, zusätzlich die abgeleiteten DOMs aus der Triangulation der Bilddaten abzubilden.

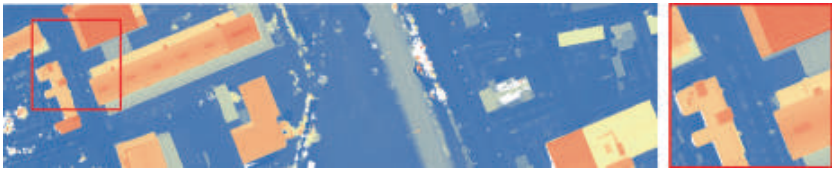


Abb. 5: DOM für DOMCam 1 (links) mit Detail (rechts)

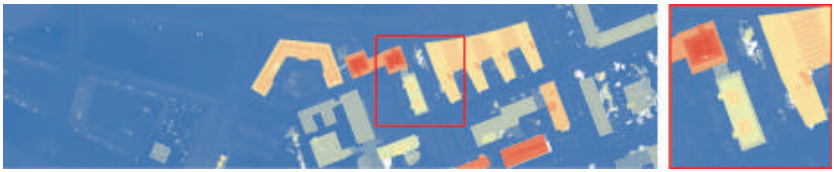


Abb. 6: DOM für DOMCam 2 (links) mit Detail (rechts)

Zur Abschätzung der Qualität der generierten DOMs wurden anschließend Höhendifferenzen zu vorliegenden signalisierten Referenzpunkten im Befliegungsgebiet berechnet, ausgenommen jene Passpunkte, die bereits innerhalb der Triangulation Anwendung fanden. Die Analyse sämtlicher vertikaler Differenzen zeigte unter Zuhilfenahme des Orthophotos aus GeoCam des identischen Fluges, dass sich maximale Werte durch Verdeckungen wie Bäume, Fahrzeuge und Container ergaben. Um die Resultate der DOMCam-Module vergleichbar zu machen und die Analyseergebnisse insgesamt nicht durch grobe Fehler zu verfälschen, wurden diese Höhendifferenzen zur Berechnung der RMS-Werte nicht berücksichtigt. In Tabelle 7 sind die statistischen Werte zu sämtlichen Auswertemöglichkeiten der Höhendifferenzen des Kamerasystems MACS 50/70 enthalten. Dazu zählen vertikale Differenzen zwischen Passpunkten zu DOMs des triangulierten Gesamtsystems sowie zu DOMs der referenzorientierten DOMCams. Es wird zunächst ersichtlich, dass man mit der Methode der

Referenzorientierung für beide DOMCams gleiche Genauigkeiten mit 5.7 und 5.3 cm erzielt. Der leichte Genauigkeitsverlust für die Triangulation der DOMCam 1 liegt in der Konstellation des Testfluges begründet. Dabei bildeten die äußeren Bildstreifen jeweils die Aufnahmen der DOMCam 1 ab (vgl. Abb. 2). Daher waren für diese Streifen keine Bildanschlüsse quer zur Flugrichtung vorhanden. Eine Stabilisierung war nur durch die hochgenaue Näherungswertlösung der GeoCam, Referenzpunkte und hohen Längsüberdeckung gegeben. Damit kann gesagt werden, dass beide Orientierungsverfahren gleich genaue Ergebnisse erzielen. Weiterhin entsprechen die Genauigkeiten der abgeleiteten DOMs aus MACS 50/70-Bilddaten nach direkter Referenzorientierung damit der Oberflächengenauigkeit anderer Luftbildkamerasysteme wie UltraCamX (HAALA et al., 2010).

Tab. 7: Höhendifferenzen zwischen DOM und Passpunkten

Sensor	Orientierung	DOM [cm] Raster	RMS [cm] ohne grobe Fehler	Mittel [cm]	Max/Min [cm]		Anz. Punkte
	AT mit Bündelblockausgleichung						
GeoCam	GSD 13cm	L=15	8.8	5.0	13.9	-17.5	87
DOMCam 1	GSD 9 cm	L=10	7.2	0.3	11.4	-9.1	54
DOMCam 2			5.7	-0.7	9.1	-10.7	53
	Referenzorientierung						
DOMCam 1	GSD 9cm	L=10	5.7	1.6	9.4	-6.2	54
DOMCam 2			5.3	-0.7	8.4	-10.9	53

5 Fazit/Ausblick

Im vorliegenden Artikel wurde die Referenzorientierung als Verfahren zur Orientierung von Bilddaten eines Multisensorsystems vorgestellt und analysiert. Ein Bestreben innerhalb der Luftbildkameraentwicklung ist es immer, die Auswertung so nutzerfreundlich wie möglich zu gestalten, um so einen praktikablen Einsatz zu gewährleisten. Die getrennte Auswertung von Bilddatensätzen jedes Einzelsensors ist daher nicht immer nötig. Verschiedene Kamerahersteller gehen an dieser Stelle dazu über, aus den Bilddaten der einzelnen Kameras des Gesamtsystems virtuelle Negative (DOERSTEL et al., 2002, GRUBER & LADSTÄDTER, 2008) zu identischen Aufnahmezeitpunkten zu berechnen, um den Aufwand bei der Orientierung zu minimieren. Dabei entstehende nicht vermeidbare Restfehler im virtuellen Kameramodell sind die Folge (JACOBSEN et al., 2010). Die Analyse zeigt für das MACS 50/70, dass die Referenzorientierung ein valides Verfahren ist, um die Bilddaten des Sensorsystems zu orientieren. Damit verbunden beschränkt sich die photogrammetrische Auswertung des Gesamtsystems auf nur noch einen Sensor. Herauszustellen sind dabei die hohen erreichten Genauigkeiten für einen aus photogrammetrischer Sicht zunächst unvorteilhaft konfiguriertes Sensorsystem. Neben schwer zu orientierenden Sensoren wie Obliquekameras oder Thermalsensoren, kann die Referenzorientierung z.B. die Anzahl notwendiger Flugstreifen zur hochauflösenden Erfassung eines Zielgebietes minimieren und damit einen wirtschaftlichen Mehrwert entwickeln. Weiter

verringern sich zusätzlich der Speicherbedarf und die Rechendauer für notwendige Bildpyramiden. Das Verfahren der Referenzorientierung setzt die Kenntnis der Transformationsparameter zwischen den Einzelsensoren voraus. Weiter ist sie abhängig von einer hohen räumlichen Lagestabilität der Sensoren sowie insbesondere synchron auslösenden Kameras.

6 Literaturverzeichnis

DOERSTEL, C., ZEITLER, W. & JACOBSEN, K., 2002: Geometric Calibration of the DMC: Method and Results. ISPRS Com I and PECORA. Denver 2002.

GRUBER, M. & LADTSTÄDTER, R., 2008: Calibrating the digital large format aerial camera UltraCamX. International Calibration and Orientation Workshop EuroCOW 2008 Proceedings. Castelldefels, Spain 2008.

JACOBSEN, K., CRAMER, M., LADTSTÄDTER, R., RESSL, C. & SPRECKELS, V., 2010: DGPF Project: Evaluation of Digital Photogrammetric Camera Systems – Geometric Performance. PFG 2/2010: S. 83-97.

HAALA, N., HASTEDT, H., RESSL, C. & WOLFF, K., 2010: DGPF Projekt: Evaluierung digitaler photogrammetrischer Luftbildkamerasysteme – Themenschwerpunkt Höhenmodelle. Proceedings 3-Ländertagung Wien 2010.

HIRSCHMÜLLER, H., 2008: Stereo Processing by Semi-Global Matching and Mutual Information. – IEEE Transactions on Pattern Analysis and Machine Intelligence 30 (2): 328–341. Los Alamitos, CA, USA 2008.

LEHMANN, F., BERGER, R., BRAUCHLE, J., HEIN, D., MEISSNER, H. & PLESS, S., 2011: MACS – Modular airborne camera system for generating photogrammetric high-resolution products. PFG 6/2011: S. 435–446.

DIN 18714-8 − Anforderungen an die Bildqualität

RALF REULKE[1]

Zusammenfassung: In den letzten 20 Jahren gab es eine Reihe von Aktivitäten, um adäquate Ansätze zur Beschreibung der Qualität von optischen Fernerkundungsdaten abzuleiten. Grundsätzlich kann man zwischen generischen und aufgabenbezogenen Ansätzen unterscheiden. Ein aufgabenbezogener Ansatz bedeutet, dass die Bildqualität (IQ – image quality) letztlich nur durch die Beurteilung der Qualität des Endproduktes (z.B. Objektdetektion, Klassifizierung) eingeschätzt werden kann. Fokussiert man sich auf die Objektdetektion, dann kann der National Image Interpretability Rating Scales (NIIRS) verwendet werden. Der vorliegende Beitrag zeigt einige Ansätze zur Beschreibung der Bildqualität auf und beschreibt erste Ansätze zur Normung.

1 Einleitung

Photogrammetrie und Fernerkundung stellen Verfahren bereit, die aus Bilddaten geometrische und thematische Informationen gewinnen und verarbeiten. Zur Erfassung von Bilddaten existiert eine Vielzahl von Flugzeug- und Weltraumsensoren. Entsprechend der Zielsetzung der Normenreihe DIN 18740 soll diese Norm die Qualitätsanforderungen an optische Fernerkundungsdaten festlegen. Aus der absoluten radiometrischen Kalibrierbarkeit digitaler Sensoren ergeben sich neue Möglichkeiten für die Erstellung abgeleiteter Produkte, da diese Kameras neben hoher geometrischer Qualität auch die radiometrischen Standards von Erdbeobachtungssystemen erfüllen. Um eine Vergleichbarkeit der Sensoren und der gewonnenen Bildprodukte zu gewährleisten, verwendet man bestimmte Beschreibungsgrößen oder Metriken für Bildqualität. Da diese bisher nur teilweise miteinander vergleichbar sind, ist eine Normierung der Bildqualität notwendig.
Die Bestimmung der Güte von Fernerkundungsdaten lässt sich prinzipiell in (spektral) radiometrische und geometrische Aspekte unterscheiden. Für die Beschreibung lassen sich unterschiedliche Metriken finden, die zum einen (spektral radiometrische und geometrische) Genauigkeitsaspekte (Wandlung von DN in Strahlungsgrößen, geometrische Punktgenauigkeit, z.B. CE90) aber auch Performancegrößen (SNR, MTF) betreffen. Der Beitrag ist folgendermaßen aufgebaut. Zuerst werden einige Grundlagen zur Bildqualität und eine Literaturübersicht gegeben. Anschließend werden einige Normungsansätze aufgezeigt. Der Beitrag schließt mit Zusammenfassung und Ausblick.

2 Bildqualität

2.1 Anwendungsbereich

Diese Norm soll für die Bestimmung der Güte optischer Fernerkundungsdaten auf der Grundlage multispektraler Luft- und Satellitenbilder gelten, die mittels digitaler optischer Sensoren

[1] Ralf Reulke, Humboldt-Universität zu Berlin, Institut für Informatik, Computer Vision, Unter den Linden 6, 10099 Berlin; E-Mail: ralf.reulke@informatik.hu-berlin.de

(Flächen- oder Zeilenkameras) erfasst werden. Diese Sensoren müssen messtechnische Anforderungen (radiometrisch und geometrisch) erfüllen. In dieser Norm werden aktive Sensoren (Radarsysteme und Laserscanner) nicht betrachtet.

Betrachtet man die Bildgüte, gibt es grundsätzlich zwei Herangehensweisen. Zum einen kann generisch festgestellt werden, wie sich ein ideales zu einem realen System verhält (JAHN, et al. 2012). Zum anderen kann man Bildqualität in Bezug auf spezielle Anwendungen definieren.

2.2 Grundlegende Parameter

Zur Charakterisierung von Bilddaten oder abgeleiteten Produkten wird oft die räumliche Auflösung (GSD – ground sampling distance) verwendet (siehe DIN 18716). Diese steht im Zusammenhang zur Pixelgröße, der Brennweite und der Flughöhe. Es gibt allerdings weitere Einflüsse des Bildgebungssystems, welche die Bildqualität beeinträchtigen. Die wesentlichen Parameter sind die Punktbildverschmierungsfunktion (PSF – point spread function) und das Signal-zu-Rausch-Verhältnis (SNR) des Bildes. Sie hängen eng mit subjektiven Parametern wie Bildschärfe zusammen und werden im Folgenden als Performancegrößen bezeichnet.

Metriken zur Beschreibung der Bildqualität können für jedes abbildende System abgeleitet werden. Sie erlauben eine quantitative Charakterisierung des Systems in Bezug auf Abbildungsschärfe und zusätzlichen Bildeigenschaften, wie zum Beispiel Rauschen, Nichtlinearitäten und Artefakten nach der Kompression. Im Allgemeinen hängt die Bildqualität letztlich von der zu lösenden Aufgabenstellung ab und kann nicht für jedes Bild allgemein definiert werden. Will man zum Beispiel Sternpositionen messen, ist die Genauigkeit der abgeleiteten Position besser, wenn das Bild unscharf ist. Besteht die Aufgabe in der Objekterkennung, sollte das Bild möglichst scharf sein.

Um eine bestimmte Aufgabe wie Objekterkennung oder Erkennung in einer bestimmten Umgebung (z.B. die Erfassung eines in einer landwirtschaftlichen Fläche) mit einem satellitengestützten Sensor zu lösen, wird dieser Sensor auf die Lösung dieser Aufgaben optimiert. Daraus resultieren Sensorparameter wie Brennweite, Blende und so weiter. Unter diesen Bedingungen lässt sich ein idealer Sensor definieren und simulieren, der optimale Ergebnisse, bzw. die beste Bildqualität liefert. Dazu wurden Simulationssysteme entwickelt (siehe BÖRNER, et al. 2001).

2.3 Literaturübersicht

Das wohl am häufigsten genutzte Beschreibung der Bildqualität ist der NIIRS. NIIRS wurde vom Imagery Resolution Assessments and Reporting Standards (IRARS) Ausschuss entwickelt. Er bewertet die Bildqualität von 0-9. Höhere Werte unterstützen eine detailliertere Objektanalyse.

(LEACHTENAUER, et al. 1997) führten einen heuristisch abgeleitete mathematischen Zusammenhang ein, der die Bildqualität auf Basis des NIIRS berechnet (GIQE - General Image Quality Equation). GIQE ist eine Metrik welche die Qualität eines Bildes auf der Basis der GSD, relativer Sprungantwort und Signal-zu-Rausch-Verhältnis bestimmt. Zusätzliche Beiträge kommen von der Bildrestaurierung:

$$\mathrm{NIIRS} = 10.251 - a \cdot \log_{10}\mathrm{GSD} + b \cdot \log_{10}\mathrm{RER} - 0.656 \cdot H - 0.344 \cdot G / \mathrm{SNR}$$

GSD	Ground sampled distance (legt die Nyquist frequenz fest)
RER	Relative Edge Response (bezieht sich auf eine Kante im Bild, wobei RER aus dem Anstieg an der Kante bestimmt wird)
G	Gain (Verstärkung des Rauschens durch Bilddatenrestauration)
SNR	Signal-to-noise ratio
H	Overshoot Term (Bestimmt die Größe des Ringing an einer Kante, das durch Bilddatenrestauration hervorgerufen wird)

Probleme bei der Anwendung einer Qualitätsmetrik im Fernerkundungsbereich werden von (MIETTINEN, 2004) beschrieben. Er führt deshalb eine NIIRS basierende Metrik ein, die das Powerspektrum des additiven inkohärenten Rauschens analysiert. In dem Beitrag von (KIM, 2010) wird eine Metrik eingeführt, die den Phasenquantisierungscode (PQC) nutzt. (DUMIC, et al., 2010) präsentieren einen Ansatz, zur Bewertung der Bildqualität, unter Verwendung der mittlere Differenzen zwischen der ursprünglichen und untersuchten Bilder in verschiedenen Wavelet-Subbänder. Das Qualitätsmaß wird dann aus der gewichteten Differenz zwischen den Koeffizienten von Original- und degradierten Bildern berechnet. Die Bildqualität beeinflusst die Ergebnisse der nachfolgenden Bildverarbeitung (z.B. automatisierte Klassifizierung von Bildern). In (YAN, et al. 2009) wurde eine Datenbank von Referenzbildern erstellt, die eine automatisierte, kundenspezifischen Änderung der Bildqualität ermöglicht, um die Klassifizierung neuer Bilder zu verbessern. Sie führen auch eine anwendungsbasierte Definition von Bildqualität ein. (GERWE, et al. 2009) präsentieren eine neue informationstheoretische Bildqualitäts-evaluierung (ITIQUE) für die Modellierung und Vorhersage des NIIRS, basierend auf der Visual Information Fidelity (VIF) IQ Metrik. Die Auswertung zeigt eine gute Übereinstimmung mit der GIQE. (CHEN XUE et al. 2008) konzentrieren sich auf die mutual Information basiertes Gütemaß. Basierend auf einem Modell für die Bildergeneration, erhalten sie einen geschlossenen Ausdruck für das Qualitätsmaß.

Zuletzt soll noch auf die DIN 62464-1 hingewiesen werden, die sich mit der Bestimmung der wesentlichen Bildqualitätsparameter für die medizinische Bildgebung befasst.

3 Normungsansätze

3.1 Einleitung

Die Erfassung der Güte von Fernerkundungsdaten ist Teil der Validierung von Fernerkundungssystemen. Ausgangsdaten für die Bestimmung der Güte sind (georeferenzierte) Bilddaten eines oder mehrerer Sensoren zu einem oder mehreren Zeitpunkten.

Die Bestimmung der Bildgüte kann auf unterschiedlichen Verarbeitungsniveaus erfolgen. Dies erfordert eine Vorverarbeitung der Bilddaten zur Herstellung einer vergleichbaren Datenqualität, die Gewinnung, Bereitstellung und Bearbeitung von Referenzdaten und Zusatzinformationen.

Externe Zusatzinformationen wie Referenzdaten, Ground Truth Daten oder Daten aus bereits existierenden Datenbeständen, sowie die Kenntnis der Steuerparameter, z.B. bei der Nachverarbeitung (Pansharpening, resampling kernal) sind ebenfalls notwendig.

3.2 Güte- und Performancebestimmung

Der Evaluierungsprozess ist nach verschiedenen Gesichtspunkten zu bewerten und zu untergliedern:

- Visuelle Inspektion
  - Artefakte wie z.B. Farbsäume;
  - Erfasster Dynamikbereich und Grauwertverteilung in den Bilddaten;
  - Subjektiver Bildeindruck (Bildschärfe, Farbbalance, etc.);
- Genauigkeitsanalyse
  - Genauigkeit der Koregistrierung bei der Verwendung von mehreren Spektralkanälen oder Sensoren;
  - Genauigkeit der Punktbestimmung;
  - Genauigkeit der Umrechnung DN in Strahldichteeinheiten bzw. der Berechnung von Reflektanzwerten am Boden;
  - Farbfehler bei True Color;
- Performanceanalyse
  - Signalabhängiges SNR
  - MTF in unterschiedlichen Richtungen und an unterschiedlichen Orten im Bild
  - Sonstiges (z.B. Overshoot Term in der GIQE)

Es gibt eine Reihe von Möglichkeiten zur Analyse und Ableitung dieser Parameter. Ein wichtiger Punkt ist dabei die Erfassung von Bilddaten, die eine solche Analyse erlauben. Dabei kommt der Planung und Durchführung von Validierungskampagnen eine besondere Bedeutung zu. Prinzipiell gibt es folgende Möglichkeiten:

- Analyse mit Testmustern wie zum Beispiel der Siemensstern. Beispiele für natürliche und künstliche Testmuster findet man in dem Beitrag von (PAGNUTTI, 2010)
- Automatische Verfahren aus realen Daten zur Bestimmung von SNR, MTF
- Vergleiche mit Modellen und Simulationen

Im Folgenden soll kurz auf einige Aspekte, wie Artefakte, die SNR / PSF Bestimmung und der Einfluss der Bildvorverarbeitung (Bilddatenrestauration) eingegangen werden.

3.3 Artefakte

Ursachen für Artefakte, also Fehler in den Bilddaten, ist der Detektor (Pixelfehler, Datenausfälle, Blooming, usw.) selbst, aber auch die nachfolgende Verarbeitung (z.B. PAN-Sharpening). Typische Artefakte sind zum Beispiel Überstrahlungen mit Farbsäumen, Verfärbungen und Verschmierung von Objektstrukturen. Hinzu kommt der Einfluss von Bewegungen des Trägers und bewegte Objektstrukturen (z.B. Windräder).

Solche Bildfehler sind abhängig bzw. werden erzeugt von

- der Vorverarbeitung (Rohdaten, PRNU, etc., Pansharpening),

- Fehler durch Algorithmen und mangelhafte Zusatzdaten (DGM, Atmosphärenparameter)
- Übertragungsfehler, etc

3.4 SNR-Bestimmung

Es werden drei Komponenten des Rauschens betrachtet:

- Photonenrauschen (shot noise)
- Dunkelstromrauschen (dark noise)
- Rauschen des elektronischen Kanals (read-out noise)

Während das elektronische Rauschen normalverteilt ist, verhalten sich Photonen- und Dunkelstromrauschen entsprechend einer Poissonverteilung und sind damit signalabhängig.

Um das signalabhängige Rauschen mittels realer Bilddaten bestimmen zu können, werden häufig homogene Bereiche im Bild verwendet. Das können künstliche aber auch natürliche Flächen sein. Einen alternativen Ansatz findet man in (REULKE, 2012).

3.5 PSF/RER-Bestimmung

In realen Bilddaten wird häufig statt der PSF die Kantenverschmierungsfunktion (ESF) bestimmt. Ein Ansatz wird in dem Buch von (JAHN, 1995) beschrieben. Einen Überblick findet man auch in der Veröffentlichung von (HELDER, 2004). Aus der ESF lässt sich leicht der RER, der für die Berechnung des NIIRS von Bedeutung ist, ableiten.

Häufig geht man von einer normalverteilten PSF aus:

$$H(x) = \frac{1}{\sigma_H \sqrt{2\pi}} \cdot e^{-\frac{x^2}{2 \cdot \sigma_H^2}}$$

Daraus lässt sich die ESF als Fehlerfunktion (plus einige zusätzliche Terme) ableiten:

$$y(x) = a_0 \cdot \Phi\left(\frac{x - a_1}{a_2 \sqrt{2}}\right) + a_3 + a_4 \cdot x$$

mit

x	Ortskoordinate
y	Messwert
◘	Fehlerfunktion
a_i	Parameter

Der Parameter a_2 entspricht der Größe σ_H. Der RER entspricht dem Anstieg der ESF:

$$RER = \frac{\partial}{\partial x}\left\{\frac{1}{2}\left(I_{max} - I_{min}\right) erf\left[\frac{x - x_0}{\sqrt{2}\sigma}\right]\right\} = \frac{I_{max} - I_{min}}{\sqrt{2}\sigma} e^{\frac{(x - x_0)^2}{2\sigma^2}} \xrightarrow[x = x_0]{} \frac{I_{max} - I_{min}}{\sqrt{2}\sigma}$$

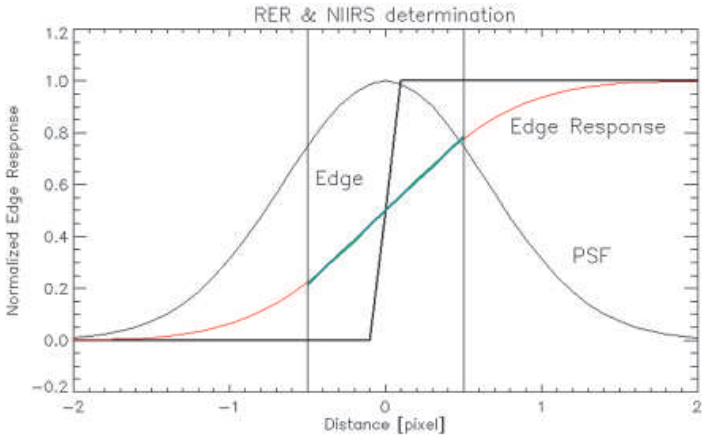


Abb. 1: - Normalized Relative Edge Response (RER)

Abbildung 1 zeigt ein Beispiel. Die Kante (schwarz) wird mit der PSF (grau) gefaltet. Daraus ergibt sich die verschmierte Kante (edge response), die der ESF entspricht. Der Anstieg der grünen Lienie um 0 herum ist der RER.

3.6 Einfluss der Bilddatenrestauration

Zur Verdeutlichung dieses Effekts soll die Wirkung von Bilddatenrestaurationsalgorithmen dargestellt werden (siehe Abbildung 2).

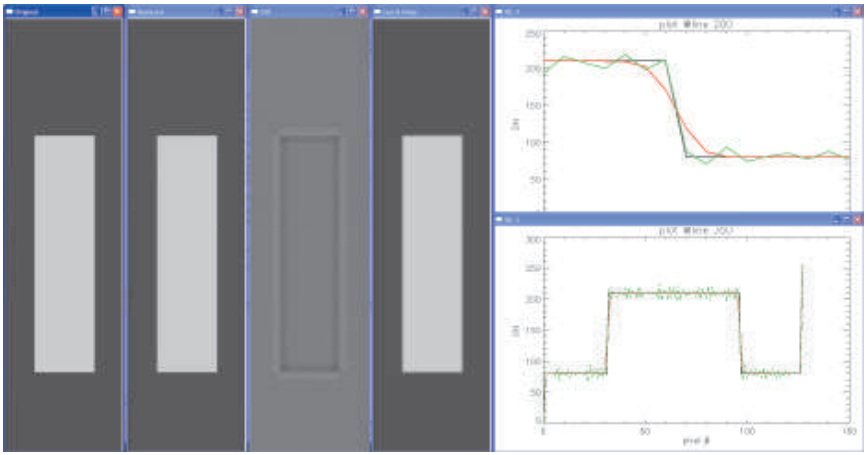


Abb. 2: Bilddatenrestauration, von links nach rechts: Original, restauriertes Bild Differenzbild, verschmiertes Bild, Profileausschnitt (oben), gesamtes Profil (unten)

Bilddatenrestauration gehört (mathematisch gesehen) zur Klasse der inkorrekt gestellten Aufgaben. In Abb. 2 ist das Ergebnis eines solchen Algorithmus dargestellt. Der Algorithmus wurde in (REULKE, 2004) vorgestellt. Man sieht, dass geringfügiges Rauschen im Ausgangsbild verstärkt wird. Das Überschwingen an den Kanten ist deutlich sichtbar. Dieser Effekt ist deutlich in der Teilabbildung oben rechts sichtbar: Die originale Kante (schwarz) lässt sich aus den realen Daten (rot) restaurieren (grün). Im Ergebnis wird das Rauschen drastisch verstärkt. Überschwingeffekte an den Kanten (ringing) sind deutlich sichtbar.

4 Zusammenfassung und Ausblick

In dem Beitrag wurden Algorithmen zur Bestimmung der Bildqualität von optischen Fernerkundungssensoren zusammengefasst, einige Probleme bei der Bestimmung der Parameter aufgezeigt und erste Ansätze zur Normung dieser Größen dargestellt.

Während die quantitativen Ansätze zur Bewertung der geometrischen und radiometrischen Güte der Bilddaten bereits gut etabliert sind, findet man für die Bestimmung von Performancegrößen keine einheitlichen Bewertungskriterien. Es kommt hinzu, dass häufig der Einfluss der Bildvorverarbeitung vom Bilddatenprovider nicht von der eigentlichen Bildqualität der originalen Bilddaten getrennt werden kann.

In Zukunft sind zwei Aspekte von Bedeutung: Zum einen müssen die vorhandenen Performancegrößen zusammengefasst und verglichen werden. Insbesondere muss die Verbindung zur Aufgabenstellung (z.B. Objekterkennung, Klassifizierung) hergestellt werden. Zum Anderen müssen einheitliche Ansätze zur Bestimmung der Komponenten (z.B. SNR, PSF) dieser, meist sehr komplexen Größen, vorgegeben werden. Dazu sind etablierte und vergleichbare Ansätze zur (halbautomatischen) Bestimmung dieser Größen notwendig.

Es zeigt sich, dass für das Verständnis und die Nutzung noch grundlegende Untersuchungen notwendig sind. Hier müssen die Anwender einbezogen werden, um sinnvolle und praktikable Ansätze zur Bestimmung der Bildqualität zu etablieren.

5 Literaturverzeichnis

BÖRNER, A., WIEST, L., KELLER, P., REULKE, R., RICHTER, R., SCHAEPMAN, M. & SCHLÄPFER, D. Sensor: a tool for the simulation of hyperspectral remote sensing systems 2001: IN ISPRS Journal of Photogrammetry and Remote Sensing, volume 55, pages 299-312, 2001.

CHEN, Y. & XUE, Z. Y., et al. 2008. Theoretical analysis of an information-based quality measure for image fusion." Information Fusion 9(2): 161-175.

DUMIC, E., GRGIC, S., ET AL., 2010: New image-quality measure based on wavelets, JOURNAL OF ELECTRONIC IMAGING **19**(1).

GERWE, D., R. , CARLOS, E. L., et al., 2009: Information Theoretic Based Image Quality Evaluation, Optical Society of America.

HELDER, D., CHOI, T. & RANGASWAMY, M., 1995: In-flight characterization of spatial quality using point spread function. In: Morain, S. A. and A.M. Budge (2004). "Post-Launch Calibration of Satellite Sensors", Taylor & Francis Group, London UK

Jahn, H. & Reulke, R. 1995: Systemtheoretische Grundlagen optoelektronischer Sensoren, WILEY-VCH Verlag GmbH & Co.

JAHN, H., & REULKE, R., 2012: A Sensor-Based Approach to Image Quality. Photogrammetrie, Fernerkundung, Geoinformation (PFG), **1**, S. 19-27.

LEACHTENAUER, J. C. & W. MALILA, et al., 1997: General Image-Quality Equation: GIQE. Applied Optics **36**(32), S. 8322-8328.

PAGNUTTI, M. & BLONSKI, S., et al. 2010: Targets, methods, and sites for assessing the in-flight spatial resolution of electro-optical data products, Canadian Journal of Remote Sensing, 36(5): 583-601.

REULKE, R., & WEICHELT, H., 2012: SNR Evaluation of the RapidEye Space-borne Cameras. Photogrammetrie, Fernerkundung, Geoinformation (PFG), **1**, S. 29-38.

REULKE, R., TEMPELMANN, U., STALLMANN, D., CRAMER, M. & HAALER, N., 2004: Improvement of Spatial Resolution with Staggered Arrays As Used in The Airborne Optical Sensor Ads40, volume IAPRS - XXXV, 2004.

REULKE, R. & WEICHELT, H., 2012: SNR Evaluation of the RapidEye Space-borne Cameras, Photogrammetrie - Fernerkundung - Geoinformation, Volume 2012, Number 1, February 2012 , pp. 29-38(10)

YAN, S., SAYAD, S., et al., 2009: Image quality in image classification: Design and construction of an image quality database. Computers & Chemical Engineering **33**(2): 421-428.

DIN 18716, Photogrammetrie und Fernerkundung - Begriffe, August 2012.

DIN 62464-1, Magnetresonanzgeräte für die medizinische Bildgebung - Teil 1: Bestimmung der wesentlichen Bildqualitätsparameter

Sind das landesweite Netzwerk von Grenzen und die Grenzsteine ein Weltkulturerbe?

PETER WALDHÄUSL[1], GÜNTHER ABART, KLAUS HANKE, MICHAEL HIERMANSEDER, HEINZ KÖNIG, REINFRIED MANSBERGER, GERHARD NAVRATIL, GERDA SCHENNACH & CHRISTOPH TWAROCH

Zusammenfassung: Die UNESCO hat dazu aufgerufen, auch technisch fundamental Außerordentliches als Welterbe einzureichen. Die Triangulierungskette des Struve-Meridianbogens ist ein klarer Präzedenzfall. Das landesweite Netz einvernehmlich definierter, vertraglich gesicherter und neutral verwalteter Grenzen ist überaus bedeutend. Es hat befriedende Wirkung. Es existiert auch in Ländern mit derzeit wenigen Weltkulturerbe-Stätten. Gesichertes Eigentum an Grund und Boden ermöglicht Lebensunterhalt, Kredite und damit wirtschaftlichen Aufschwung.
Österreich wird Geschichte, Bedeutung und die kulturellen Werte in einem allgemeinen Antragsteil behandeln, das spezifisch Österreichische über Kataster und Grundbuch in seinem nationalen Teil. Es ist vorgesehen, andere Staaten einzuladen, den ersten Teil zu ratifizieren und mit eigenem National-Teil dem Antrag gleich oder später beizutreten. Der Beitrag informiert über Welterbe und berichtet über den Stand des Vorhabens und die beabsichtigte internationale Zusammenarbeit.

1 Einleitung

Grund und Boden zählt neben Arbeit und Kapital zu den klassischen Produktionsfaktoren der Volkswirtschaft, allerdings mit dem Unterschied, dass *Grund und Boden* kein homogenes Gut und auch nicht vermehrbar ist. *Land* ist eine wesentliche Grundlage des Lebens: Es ist die Produktionsstätte von Nahrung, Voraussetzung für Fauna und Flora, Speicher natürlicher Ressourcen (z.B. Wasser, Mineralien), Träger von Infrastruktur – aber *Grund und Boden* ist auch Heimat, Heimatgemeinde, Heimatland.
Diese Multifunktionalität von Land war und ist immer eine Ursache für Nutzungskonflikte. Das Verlangen nach Land, nach mehr Land oder nach einer Nutzung von Land führte und führt oft zu schwerwiegenden Differenzen oder sogar zu bewaffneten Auseinandersetzungen zwischen einzelnen Personen, Personengruppen oder Staaten. Seit jeher werden aber Streitigkeiten um Land auch durch unklar definierte Besitzverhältnisse und durch mangelnde bzw. unvollständige Dokumentation der Nutzungsrechte ausgelöst. [ENGLERT & MANSBERGER 2008] Eine wohlorganisierte Bodenordnung durch sauber definierte Grenzen, richtig gesetzte Grenzsteine und einvernehmlich beschlossene Landrechte ist Voraussetzung für eine möglichst konfliktfreie Verteilung von Land, für eine optimale Bewirtschaftung von Grund und Boden und damit für eine nachhaltige organisatorische und wirtschaftliche Entwicklung von Personen und Staaten.
Seit Menschen sesshaft wurden, Häuser bauten und Ackerbau betrieben, stellten sich Grundbesitzfragen: Es wurde abgegrenzt, sichtbar vermarkt und vermessen [ABART et al. 2011]. In Österreich wurden im 18. Jahrhundert mit der Einführung von Kataster und Grundbuch auch die

1) Em.Univ.Prof. Dr. Peter Waldhäusl ist Leiter der Arbeitsgruppe „Grenzstein“ der Österreichischen Gesellschaft für Vermessung und Geoinformation. Alle genannten Mitglieder sind Mitautoren. Adresse: Weimarerstr.114/2, 1190 Wien, E-Mail: pw@ipf.tuwien.ac.at

Grenzen von Parzellen und die Eigentumsverhältnisse der Liegenschaften dokumentiert – vorerst mit dem Hauptziel der Grundsteuer-Einhebung. Seitdem haben sich Kataster und Grundbuch kontinuierlich zum Landadministrationssystem weiter entwickelt, bedingt durch geänderte gesellschaftspolitische Anforderungen, durch neue rechtliche Rahmenbedingungen und durch technische Innovationen [TWAROCH ET AL. 2011].
Die lange Tradition, die Dynamik der Entwicklung sowie die große friedensstiftende und friedenserhaltende Bedeutung zeichnen Grenzstein und Grenzen, Grundbuch und Kataster als potentielles Weltkulturerbe aus.

2 Die Liste des Welterbes

Am 16. November 1972 hat die Generalversammlung der UNESCO das Übereinkommen zum Schutz des Kultur- und Naturerbes der Welt, die Welterbekonvention, beschlossen. Leitidee dieser Konvention ist die „Erwägung, dass Teile des Kultur- oder Naturerbes von außergewöhnlicher Bedeutung sind und daher als Bestandteil des Welterbes der ganzen Menschheit erhalten werden müssen“ [SCHLÜNKERS, 2009]. Kultur- und Naturgüter, die in der Liste des Welterbes eingetragen sind, genießen besonderen Schutz nach internationalem Völkerrecht. Die per 19. September 2012 190 Signatarstaaten haben sich verpflichtet, ihr eigenes Welterbe zu schützen und zu erhalten sowie zu Schutz und Erhaltung des Welterbes aller anderen Staaten bestmöglich beizutragen[2] [UNESCO 2012 A].
Insgesamt sind derzeit 962 Welterbe-Stätten gelistet, 745 davon sind Kulturerbe, 188 Naturerbe

und 29 gehören beiden Kategorien an. In Österreich gibt es derzeit neun Welterbe-Stätten,

Deutschland hat 37, die Schweiz 11. Auf den 173 Vorschlagslisten warten weitere 1567 Anträge, aus Deutschland 10, Österreich 10 und aus der Schweiz 1. Diese sind teilweise alte Vorschläge. Viele davon haben nur geringe oder gar keine Aussicht auf Bewilligung durch die alles entscheidende Welterbe-Kommission, solange nichts wesentlich Neues oder besonders Gefährdetes dabei ist und solange nicht alle anderen UNESCO-Mitgliedsstaaten ebenfalls Welterbe-Stätten in vergleichbarem Umfang haben [UNESCO 2012A]. Die Welterbe-Kommission ist aus 21 gewählten Vertretern der 190 Mitgliedsstaaten zusammengesetzt und wird alle 4 Jahre durch in der Regel 7 Staaten nachbesetzt. Deutschland und die Schweiz waren schon wiederholt in der Kommission und sind es auch derzeit bis 2015. Österreich war noch nie in der Kommission vertreten.
Was kann denn zu Weltkulturerbe erklärt werden? Nach der Konvention sind als Kulturerbe etwa Werke der Architektur, Großplastiken und Monumentalmalereien, Objekte oder Überreste archäologischer Art, Gruppen einzelner oder verbundener Gebäude sowie *Werke von Menschenhand* oder gemeinsame Werke von Natur und Mensch zu verstehen, und zwar dann, *wenn sie aus geschichtlichen*, künstlerischen, *wissenschaftlichen,* ästhetischen, anthro-pologischen *oder* ethnologischen *Gründen von außergewöhnlichem Interesse oder Wert für die gesamte*

[2] Die Schweiz ist am 17. September 1975, Deutschland am 23. August 1976 (DDR am 12.12.1988, Vereinigung der Abkommen per 3. Oktober 1990) und Österreich am 18. Dezember 1992 beigetreten.

Menschheit sind. Inzwischen wurde der Begriff des Welterbes auch auf das „intangible heritage" ausgedehnt, also auch auf Stätten hoch bewerteter Bräuche, Kulte oder religiöser Traditionen.

3 Lücken in der Welterbe-Liste und Neuanträge

Die „Globale Strategie" der UNESCO strebt eine repräsentative, ausgewogene und glaubwürdige Liste des Erbes der Welt an. 2002 hat das Welterbe-Komitee mit dem Dokument 26COM13 ICOMOS[3] beauftragt, in der Welterbe-Liste und den Vorschlagslisten unterrepräsentierte Kategorien und Länder zu identifizieren. Beim ICOMOS Meeting 2004 in Bergen wurde die Studie "..Filling the Gap.." [PETZET 2004] präsentiert und angenommen. Darin wird die Aufnahme von *außerordentlichen technischen Errungenschaften* in die Liste gefordert, weil diese bisher mit nur 2% stark unterrepräsentiert seien. Es fehlten auch Beispiele zur *Entwicklung der technischen Infrastrukturen,* wozu eindeutig die geodätischen Netze, die Fundamente der Geoinformationssysteme, der Astronomie und der Satellitengeodäsie gehören. In den operationellen Richtlinien für Einreichungen weist das Welterbe-Komitee ausdrücklich auf diese und andere fundamentale Studien hin [UNESCO 2012B].
Als Folge davon wurde 2006 der Struve-Bogen als herausragende technische und intellektuelle menschliche Leistung in die Liste aufgenommen [UNESCO 2012C]. Ein völlig neuer Welterbe-Typ: Eine exakt vermessene und dokumentierte, jedoch unsichtbare Dreieckskette mit Resten von Bodenpunkten und Dokumenten. Er war ein wichtiger Beitrag zur Bestimmung von Größe und Form der Erde. Im UNESCO-Akt wurde auf eine mögliche Verlängerung des Struve Meridianbogens zum Ostafrika-Meridianbogen hingewiesen, was ein einzigartiges, viele Staaten verbindendes Welterbe ergeben könnte [WALDHÄUSL 2007]. Südafrika hat 2009 jedoch den historisch interessanteren, kurzen Cape Arc auf seine Vorschlagsliste gesetzt [UNESCO 2012D].
Die Mitgliedsstaaten wurden aufgefordert, nach freiem Ermessen qualifizierte und von der Regierung genehmigte neue Vorschläge bekanntzugeben. Das Welterbe Zentrum prüft, ob diese Vorschläge vor allem formal den „Operational Guidelines" entsprechen. Die Vorschlagslisten, gedacht für Einreichungen in den nächsten 5-10 Jahren, werden im Internet veröffentlicht.
Auch hinsichtlich der geographischen Verteilung des bestehenden und beantragten Welterbes gibt es Lücken. Nach den Richtlinien der UNESCO soll sich nun jeder Staat überlegen, ob er sich in der Liste über- oder unterrepräsentiert sieht (§54 ff). Den reichlicher repräsentierten Staaten wird geraten, ihre Anträge freiwillig zeitlich zu staffeln, nur *Güter zu unterdurchschnittlich vertretenen Kategorien* vorzuschlagen, allenfalls *mit ähnlichen Anträgen anderer Vertragsstaaten*, *insbesondere mit Staaten, die unterdurchschnittlich vertreten sind, zu verknüpfen,* oder überhaupt zu verzichten. Die unterdurchschnittlich vertretenen Vertragsstaaten sind aufgefordert, mehr einzureichen und *mit anderen Vertragsstaaten zusammen zu arbeiten und so ihr Fachwissen zu mehren.*- Es geht also auch um *Wissens-Transfer und Zusammenarbeit*.
Das Welterbe-Komitee hat sich auf die Bearbeitung von 45 Titeln pro Jahr beschränkt, und das gilt auch für Sammelgüter und grenzüberschreitende Güter. Geringfügige Ausnahmen über die 45 sind möglich nach folgender Priorität: i) Erstvorschläge von Vertragsstaaten; ii) *nur mit maximal drei Gütern vertretene Kategorien*; iii) langjährig wartende Vorschläge; iv) Naturerbe;

[3] ICOMOS (International Council on Monuments and Sites) ist UNESCO Fachbeirat für das immobile Kulturerbe; www.icomos.org. ICOM (International Council of Museums) ist der Fachbeirat für das mobile Kulturgut; www.icom.org. IUCN (International Union for the Conservation of Nature) ist für das Naturerbe zuständig; www.iucn.org. ICCROM (International Center for the Study of Preservation and Restoration of Cultural Property) ist ebenfalls Fachbeirat in Welterbe Angelegenheiten; www.iccrom.org.

v) gemischtes Erbe; vi) *grenzüberschreitendes Erbe*; vii) Afrika, Pazifik und Karibik; viii) 10 Jahre junge Vertragsstaaten; ix) Vorschläge von Staaten, die 10 Jahre nichts beantragt haben.

4 Grenzen und Grenzsteine

Österreich und auch andere Länder sind von einem *Netzwerk aus Eigentumsgrenzen* und/oder administrativen Grenzlinien überzogen. Dieses unterliegt, entsprechend den gesetzlichen Rahmenbedingungen, den wirtschaftlichen Notwendigkeiten und den aktuellen Eigentumsverhältnissen, einem kontinuierlichen Anpassungsprozess. Grenzen und Grenzsteine definieren die räumliche Ausdehnung von Landrechten und Verpflichtungen einzelner und/oder juridischer Personen. Das Grenzlinien-Netzwerk ist von außerordentlicher Bedeutung für ein respekt- und friedvolles Zusammenleben von Nachbarn, für den Schutz von Landeigentum, für die wirtschaftliche Entwicklung einzelner Bürger sowie der gesamten Gesellschaft, aber auch etwa für die Festlegung und Administration von grundstücksbezogenen Steuern. Grenzen sind mehr als nur räumliche Trennlinien: Sie stellen auch den Bezug zu den rechtlichen, kulturellen, geschichtlichen, sozialen, wirtschaftlichen und moralischen Entitäten her [SIMMERDING, 1996].



Abb. 1: Beispiel für einen Grenzstein mit besonderer Bedeutung (links: Foto H. König, 2012; rechts: Darstellung auf Basis der Digitalen Katastralmappe, Quelle: Bundesamt für Eich- und Vermessungswesen). Der Stein stammt aus dem 17. Jahrhundert. Bei ihm stoßen 6 Grundstücke, 5 Katastralgemeinden (daher ist er fünfeckig), 3 Ortsgemeinden, 2 Bezirke und 2 Viertelgrenzen Niederösterreichs zusammen. Eine besondere Kombination.

Grenzen und Grenzsteine haben eine lange Geschichte. Im Römischen Reich war das Wort *Terminus* sowohl Grenzstein als auch Bezeichnung eines Gottes. Aberglaube, Rituale und

Traditionen waren und sind mit Grenzsteinen und Grenzen verbunden. Eine Grenzverletzung wurde in vielen Ländern mit dem Tod bestraft. Die Wichtigkeit der Grenzlinie ist auch im Alten Testament dokumentiert: *Verflucht, wer den Grenzstein seines Nachbarn verrückt* (BIBEL, u.a. 5. Mose 27/17). Allein mit diesen zwei Hinweisen sind wir in zwei Jahrtausenden vor Christus. Jede gemeinsam anerkannte, eindeutig definierte und in Verträgen bei Gericht hinterlegte Grenze trägt zu einem dauerhaft möglichen Nachbarfrieden bei. Das ist eine ganz wesentliche Voraussetzung für ein gutes Zusammenleben und Zusammenarbeiten der Bürger.
Grenzen und Grenzsteine sind die Basiselemente eines Landadministrations-Systems. In Österreich wurde am 23.12.1817 von Kaiser Franz I. flächendeckend ein koordiniertes Kataster-System, der „Franziszeische Kataster" eingerichtet, der in administrativer, technischer und inhaltlicher Hinsicht ständig für die geänderten Anforderungen weiterentwickelt wurde [MUGGENHUBER & TWAROCH 2008].
Grenzen und Grenzsteine als Symbole für das Grundeigentum haben aus vier Gründen gute Chancen für eine Aufnahme in die Welterbe-Liste: Erstens sucht die UNESCO besonders vorbildliche technische Errungenschaften, wie das Beispiel Struve-Bogen gezeigt hat. Zweitens erwartet ICOMOS [ICOMOS 2012] von zukünftigem Kulturerbe Beiträge zur Schaffung und Erhaltung des Weltfriedens. Drittens bevorzugt die Welterbe-Kommission Projekte, die die internationale Zusammenarbeit fördern und viertens auch Länder beteiligt, die bislang noch kein oder wenig Welterbe haben. Grundbuch und Kataster können alle Wünsche erfüllen.

5 Kriterien für die Aufnahme in die Welterbe-Liste

In den operationellen Richtlinien der UNESCO sind sechs Kriterien für die Aufnahme von Kulturstätten in die Liste des Welterbes angeführt. Mindestens eines der ersten fünf muss für eine Annahme des Antrags erfüllt sein, das sechste ist optional. Grenzsteine und Grenzen erfüllen völlig die Kriterien (ii) und (iv), andere aber auch oder zumindest teilweise:

(i) *Meisterwerk menschlicher Schöpferkraft:* Das gilt für den Kataster an sich ebenso wie für einzelne, künstlerisch einzigartig gestaltete Grenzsteine.

(ii) *Bedeutender Austausch menschlicher Werte:* Gut definierte Grenzen in einem gut funktionierenden Landadministrations-System unterstützen den Austausch menschlicher Werte in Bezug auf sozi*ale, wirtschaftliche und Landschafts-Entwicklung*

(iii) *Einzigartiges Zeugnis einer kulturellen Tradition:* Grenzsteine und Grenzen sind bemerkenswerte Zeugnisse einer jahrtausendelangen Tradition.

(iv) *Hervorragendes Beispiel eines technologischen Ensembles, das bedeutsame Abschnitte der Menschheitsgeschichte versinnbildlicht:* Wegen des Sesshaftwerdens des Menschen hat sich die Einrichtung und die Nachführung eines landesweiten Netzes aus vereinbarten Grenzen entwickelt, eine hervorragende technische und Errungenschaft.

(v) *Herausragendes Beispiel für Landnutzung, die repräsentativ für die menschliche Interaktion mit der Umwelt ist:* Landadministrations-Systeme bilden den Rahmen, die Planungsgrundlage und die Kontrolle für eine nachhaltige Nutzung und Entwicklung von Grund und Boden. Sie sind unverzichtbare Instrumente zur Milderung der Verletzbarkeit der Umwelt durch menschliches Handeln.

(vi) *Verknüpfung mit überlieferten Lebensformen und künstlerischen Werken von außergewöhnlicher universeller Bedeutung:* Grenzsteine und Grenzen sind lebendige Zeugnisse für soziale, wirtschaftliche, rechtliche, administrative und technische Innovation. Sie sind außerdem eng mit Traditionen und Ritualen verknüpft.

Der Welterbe-Antrag "Grenzen und Grenzsteine" soll zunächst von Österreich vorbereitet werden, aber offen für andere bleiben, die sich zu einem späteren Zeitpunkt anschließen wollen. Die Österreichische Gesellschaft für Vermessung und Geoinformation (OVG) wird die Aufnahme in die österreichische Vorschlagsliste beantragen. Der Vorschlag wird dafür zur Prüfung der formalen Richtigkeit an das UNESCO-Welterbe-Zentrum in Paris weitergeleitet [BRINCKS-MURMANN, 2009].

Nach Aufnahme des Antrages in die Vorschlagsliste wird der Hauptantrag von der österreichischen Arbeitsgruppe nach den Richtlinien der UNESCO erarbeitet. Dieser Hauptantrag besteht aus zwei Teilen. Der erste beinhaltet die allgemeinen Grundsätze, die Geschichte der Grenzbildung und die allgemeine Beschreibung der Landadministrations-Systeme, das Glossar mit den wichtigsten Begriffen, und eine Bibliographie. Im Antrag wird auch begründet werden, dass das landesweite Netz „Grenzen und Grenzsteine" als bedeutende menschlich-technische Entwicklung und als Gut von außergewöhnlichem und universellem Wert zu beurteilen ist und damit den UNESCO-Kriterien für die Aufnahme in die Welterbe-Liste entspricht. Bei der Abfassung dieses ersten Teils soll es schon eine Zusammenarbeit mit Fachkollegen aus Partner-Staaten geben. Im zweiten Teil wird Österreich seinen spezifisch nationalen Antrag gestalten, der danach den Partnerstaaten als Formatvorlage dienen soll, so dass später eine zusammenfassende Bearbeitung möglich ist. Zu diesem zweiten Teil gehört auch die Dokumentation einiger außergewöhnlicher, repräsentativer und tatsächlich „aktiver" Grenzsteine als sichtbare Zeichen für dieses Welterbe. Diese sollen später auch in der Natur besonders gekennzeichnet werden.

Um zu prüfen, ob es international Interesse an einem UNESCO-Welterbe „Grenzen und Grenzsteine" gibt, wurde im Rahmen der FIG[4] Working Week 2012 in Rom eine Informationsveranstaltung organisiert. Dabei ergaben sich zahlreiche Kontakte mit Ländern aus allen Teilen der Welt. Einige haben ihr Interesse auch in Form eines *Letter of Intent* schriftlich bestätigt.

Für die UNESCO muss im Hauptantrag das Kulturgut ausführlich beschrieben werden. Gegenüber den bisher eingereichten Welterbe-Stätten mit definierten Kern- und Pufferzonen erstreckt sich das Netzwerk aus Grenzen über ein ganzes Land. Außerdem ist es seiner Art nach ein zeitlich variables Gebilde. Grundstücke und damit ihre Grenzen sind durch neue Verträge, wie Kauf, Verkauf, Teilung, Zusammenlegung, veränderlich. Änderungen sind also eine immanente Qualität. Veränderungen werden aber nur im Rahmen der gesetzlichen Bestimmungen vorgenommen und bleiben daher nachvollziehbar.

Vom Einreicher wird auch ein Managementplan für nachhaltigen Schutz, Überwachung und Verwaltung des Welterbes verlangt. Das ist für Grenzsteine einfach, da die Grundeigentümer zum Erhalt ihrer Grenzzeichen gesetzlich verpflichtet sind und das Bundesamt für Eich- und Vermessungswesen mit der Dokumentation der Grenzen beauftragt ist. In Österreich kommt noch die Aufsicht durch das Bundesdenkmalamt für historisch oder künstlerisch wertvolle Grenzzeichen hinzu, weil getrachtet wird, die ausgewählten Grenzsteine unter Denkmalschutz zu stellen. Das ist nichts Neues, schon jetzt stehen 134 Grenzsteine unter Denkmalschutz. Das diesbezügliche Management hat die OVG übernommen. Es muss bei der Auswahl auch darauf geachtet werden, dass das Welterbe kein Faustpfand bei regionalen Grenzstreitigkeiten wird. Daher ist geplant, nur innerstaatliche Grenzsteine als Symbol für das Welterbe zuzulassen.

Partnerstaaten sollten eine repräsentative geographische Verteilung über alle Kontinente aufweisen, ihre Landinformationssysteme sollten die unterschiedlichen Katastersysteme umfassen. Es ist noch mit ICOMOS und dem UNESCO WHC abzuklären, ob man besser

[4] Fédération Internationale des Géomètres (FIG), http://www.fig.net

zunächst nur einen Einzelantrag stellt oder ob gleich ein Sammelantrag gestellt werden soll. Es muss dabei auch beraten werden, welchen der interessierten Länder, viele davon sind europäische, aus der Sicht der UNESCO der Vorzug gegeben werden soll, weil Europa ohnehin stark überrepräsentiert ist.

6 Vorteile eines Welterbetitels

Die Aufnahme von Grenzen und Grenzsteinen in die Liste des Welterbes wird mediales Interesse hervorrufen. Hinter den Begriffen Grenze und Markstein verbergen sich viele Gedanken, die von den Medien angesprochen werden können. Die Bedeutung der in der Natur und in Registern dokumentierten Eigentumsverhältnisse, Rechte und Pflichten, können einer breiten Öffentlichkeit bewusst gemacht werden. Der Grenzstein soll auch als Symbol für geordnete Verhältnisse sowie für den Frieden zwischen Nachbarn gelten und die soziale Bedeutung der Grenze als ordnendes Element ins Bewusstsein rücken. Der Bevölkerung wird die Bedeutung von Grenzen und Grenzmarken als einzigartiges Zeugnis einer kulturellen Tradition vor Augen geführt.
Durch die Teilnahme am Weltkulturerbe stehen auch die beteiligten Berufsstände und öffentlichen Institutionen verstärkt im Rampenlicht. Das wiederum wird ihr Ansehen in der Bevölkerung heben und kann genutzt werden, um das hohe Berufsethos des Vermessungsingenieurs zu betonen. Die Marke „Weltkulturerbe“ unterstreicht auch, dass geordnete und nachvollziehbare Verhältnisse über Grundbesitz und dessen Dokumentation in öffentlichen Registern Voraussetzung für nachhaltige Eigentumssicherung, für einen gesicherten Grundstückshandel und für wirtschaftliches Wohlergehen sind. Ländern ohne funktionierende Eigentumssicherung an Grund und Boden wird die Bedeutung des Traditionskonzeptes Grenzen und Grenzsteine verdeutlicht. Es steht zu hoffen, dass diese Länder es als Anreiz sehen werden, ein System zur Dokumentation von Grenzen und Grenzsteinen zu implementieren, beispielsweise mittels Kataster und Grundbuch.

7 Zusammenfassung und Ausblick

Es wurde eine Initiative beschrieben, die von zahlreichen österreichischen Fachinstitutionen, vor allem der Österreichischen Gesellschaft für Vermessung und Geoinformation, getragen wird. Die Idee, Grenzen zu setzen, des Katasters mit dokumentierten Grenzen und der sie repräsentierenden Grenzsteine sollen als jahrtausendealtes Kulturerbe in die Liste des Welterbes aufgenommen werden. Der Bevölkerung sollte dadurch die Bedeutung und der große Wert von einvernehmlich definierten und durch den Staat geschützten Grenzen ins Bewusstsein gerufen werden: Grenzen zu den Nachbarn und zum öffentlichen Gut, Grenzen zwischen Gemeinden, Ländern und Staaten. Es sollte weiters bewusst werden, dass eine gute Grundstücksverwaltung ein friedensschaffender und friedenserhaltender Prozess ist, dass die Achtung der Rechte des Nachbarn die Grundlage für sozialen Frieden darstellt und dass gesichertes Eigentum an Grund und Boden auch von großer wirtschaftlicher Bedeutung für den einzelnen wie für die ganze Gesellschaft ist.
Die Arbeiten zur Vorbereitung des Antrages laufen jetzt schon seit mehr als einem Jahr. Der Eintragungsprozess wird noch Jahre dauern, aber konsequent verfolgt werden. Eine internationale und interdisziplinäre Unterstützung dafür ist nicht nur wünschenswert, sondern wirklich notwendig. Die österreichische Arbeitsgruppe ist überzeugt, dass das Netz geordneter Grenzen ein wirklich bedeutendes Werk aus Menschenhand darstellt, das aus geschichtlichen

und interdisziplinär technischen wie wissenschaftlichen Gründen für die gesamte Menschheit von außerordentlichem Interesse und Wert ist.

8 Literaturverzeichnis

ABART, G., ERNST, J., TWAROCH, C. (2011): Der Grenzkataster – Grundlagen, Verfahren und Anwendungen. Neuer Wissenschaftlicher Verlag. Wien. Graz. ISBN 978-3-7083-0770-1.

BRINCKS-MURMANN, C. (2009): Das Aufnahmeverfahren. Welterbe-Manual, pp. 74-79. Deutsche UNESCO-Kommission e.V. Bonn. ISBN 978-3-940785-05-3.

ENGLERT, B., MANSBERGER, R. (2008): Gender & Landrechte. Vergleichsstudie zu Gender und Landrechte in den Schwerpunkt- und Kooperationsländern der Österreichischen Entwicklungs- und Ostzusammenarbeit. Wiener Institut für Internationalen Dialog und Zusammenarbeit (VIDC). Wien.

ICOMOS (2012): Proceedings of the General Assembly 2011, Document 29, Paris 2012

MUGGENHUBER, G. & TWAROCH, CH. (2008). Dynamisches Vermessungsrecht. VGI, 96(2): 135-145.

PETZET, M. ET AL (2004): The World Heritage List: Filling the Gaps - an Action Plan for the Future. An Analysis by ICOMOS. Annex 3, VI Developing Technologies. http:// whc.unesco.org/uploads/activities/documents/activity-590-1.pdf (Last access: 15.12.2012)

SCHLÜNKERS, K. (2009): Das UNESCO-Welterbe. In: Welterbe-Manual. Handbuch zur Umsetzung der Welterbe-Konvention in Deutschland, Luxemburg, Österreich und der Schweiz. UNESCO-Kommissionen der jeweiligen Länder (Hrsg.). Grafische Werkstatt Druckerei und Verlag Gebrüder Kopp GmbH & Co. Köln. 2009.

SIMMERDING, F. (1996): Grenzzeichen, Grenzsteinsetzer und Grenzfrevler. DVW Bayern. ISBN 3-923825-08-0.

TWAROCH, CH.; NAVRATIL, G.; MUGGENHUBER, G.; MANSBERGER, R. (2011): Potenziale der Landadministration - Ist der Kataster noch zeitgemäß? In: Grimm-Pitzinger, A.; Weinold, Th. (Hrsg.), 16. Internationale Geodätische Woche Obergurgl 2011, pp. 176-186. Wichmann Verlag, Berlin-Offenbach. ISBN 978-3-87907-505-8.

TWAROCH, CH. (2003): Wie viel Erde braucht der Mensch. In: Chesi, G.; Weinold, Th. (Hrsg.), 12. Internationale Geodätische Woche Obergurgl 2003, pp. 153-162. Wichmann Verlag, Berlin-Offenbach. ISBN 3-87907-401-14.

UNESCO-WHC (2012A): United Nation Educational, Scientific and Cultural Organisation - World Heritage Centre. Homepage. http://whc.unesco.org/. (Last access: 15.12.2012).

UNESCO-WHC (2012B): The Operational Guidelines for the Implementation of the World Heritage Convention. http://whc.unesco.org/en/guidelines (Last access 15.12.2012)

UNESCO-WHC (2012C): Struve Geodetic Arc. http://whc.unesco.org/en/list/1187. (Last access: 15.12.2012)

UNESCO-WHC (2012D): The Cape Arc of Meridian. whc.unesco.org/en/tentativelists/5461 (Last access 19.12.2012)

WALDHÄUSL, P. (2007): Vermessung und Weltkulturerbe. In: Chesi, G.; Weinold, Th. (Hrsg.), 14. Internationale Geodätische Woche Obergurgl 2007, pp. 116-126. Wichmann Verlag, Berlin-Offenbach. ISBN 978-3-87907-446-4.

WALDHÄUSL, P. ET AL. (2013): Ist der Grenzstein etwas Besonderes? In: Hanke, K.; Weinold, Th. (Hrsg.) 17. Internationale Geodätische Woche Obergurgl 2013, Wichmann Verlag, Berlin-Offenbach (im Druck).

AUTORENVERZEICHNIS

ABART, G.: Siehe WALDHÄUSL, P., ABART, G., HANKE, K., HIERMANSEDER, M., KÖNIG, H., MANSBERGER, R., NAVRATIL, G., SCHENNACH, G. & TWAROCH, C.: Sind das landesweite Netzwerk von Grenzen und die Grenzsteine ein Weltkulturerbe? 534
ADLER, P. & PETERS, S.: Automatische Klassifizierung von Fichten in digitalen 4-Kanal Luftbildern. 9
BACH, E., FRÖHLICH, B., WALDE, I., HESE, S., SCHMULLIUS, C. & DENZLER, J.: Landbedeckungsklassifikation von Satellitenaufnahmen unter Verwendung von Kontextinformation mit minimaler menschlicher Interaktion. 17
BAYER, S., MIRI, M. & BUCHER, T.: Automatisierte Einzelbaumerfassung auf Basis von hochaufgelösten TrueOrtho-Mosaiken und photogrammetrisch abgeleiteten Oberflächenmodellen. 26
BAYER, S.: Siehe MIRI, M., BAYER, S. & BUCHER, T.: Evaluation of high resolution digital surface models for single tree extraction approaches in mixed forests. 137
BECHTEL, B., ZAKŠEK, K. & BÖHNER, J.: Hochauflösende Datensätze urbaner Oberflächentemperaturen zur Beobachtung der städtischen Wärmeinsel. 36
BECKER, T.: Siehe DARR, T., GÖTZ, C., TUTTAS, S., BECKER, T. & STILLA, U.: Genauigkeit der inversen Navigation bei hoher Passpunktsichtbarkeit zur Absolutpositionierung eines geführten Messsystems. 279
BECKER, T.: Siehe HANEL, A., GÖTZ, C., TUTTAS, S., BECKER, T. & STILLA, U.: Orientierungsgenauigkeit eines Multi-Kamerasystems bei Roboterbewegungen in einer Fahrzeugkarosserie. 298
BEHLING, R., ROESSNER, S., WETZEL, H.-U. & KAUFMANN, H.: Automated Multi-temporal Landslide Identification in Kyrgyzstan Using Satellite Remote Sensing Data. 244
BELLINGHOFF, M.: Siehe GRIWODZ, U. & BELLINGHOFF, M.: Traditionelle Luftbilder in 3D. 210
BOESCH, R. & GINZLER, C.: Automatische Berechnung einer nationalen Gehölzmaske. .. 46
BOGDANOV, S.: Siehe GRENZDÖRFFER, G. & BOGDANOV, S.: UAS-basierte automatisierte Vogelzählung am Beispiel einer Sturmmöwenkolonie der Vogelschutzinsel Langenwerder. 89
BÖHNER, J.: Siehe BECHTEL, B., ZAKŠEK, K. & BÖHNER, J.: Hochauflösende Datensätze urbaner Oberflächentemperaturen zur Beobachtung der städtischen Wärmeinsel. 36
BRANDENBURGER, W., DRAUSCHKE, M., NGUATEM, W. & MAYER, H.: Detektion von Fensterverdachungen und Gesims in Fassadenbildern und Punktwolken. 202
BRENNER, A.: Siehe MAKSYMIUK, O., BRENNER, A. & STILLA, U.: Detection of Stationary Vehicles in Airborne Decimeter Resolution SAR Intensity Images using Morphological Attribute Filters. 391
BRENNER, C.: Siehe SCHLICHTING, A., BRENNER, C. & SCHÖN, S.: Bewertung von inertialen Messsystemen mittels Laserscannern und bekannter Landmarken. ... 326
BREUER, M.: Siehe HERDA, H. & BREUER, M.: Different Ways to Process UAV Imagery – Results of Comparative Benchmark Tests. 99
BRÜCK, A.: Siehe SCHMITT, M., BRÜCK, A., SCHÖNBERGER, J. & STILLA, U.: Potential of Airborne Single-Pass Millimeterwave InSAR Data for Individual Tree Recognition. 426
BUCHER, T.: Siehe BAYER, S., MIRI, M. & BUCHER, T.: Automatisierte Einzelbaumerfassung auf Basis von hochaufgelösten TrueOrtho-Mosaiken und photogrammetrisch abgeleiteten Oberflächenmodellen. 26
BUCHER, T.: Siehe MIRI, M., BAYER, S. & BUCHER, T.: Evaluation of high resolution digital surface models for single tree extraction approaches in mixed forests. 137
BÜHLER, Y.: Siehe GINZLER, C., MARTY, M. & BÜHLER, Y.: Grossflächige hochaufgelöste Schneehöhenkarten aus digitalen Stereoluftbildern. 71

CRAMER, M., HAALA, N., ROTHERMEL, M., LEINSS, B. & FRITSCH, D.: UAV-gestützte Datenerfassung für Anwendungen der Landesvermessung – das Hessigheim-Projekt. 450
DARR, T., GÖTZ, C., TUTTAS, S., BECKER, T. & STILLA, U.: Genauigkeit der inversen Navigation bei hoher Passpunktsichtbarkeit zur Absolutpositionierung eines geführten Messsystems. 279
DENZLER, J.: Siehe BACH, E., FRÖHLICH, B., WALDE, I., HESE, S., SCHMULLIUS, C. & DENZLER, J.: Landbedeckungsklassifikation von Satellitenaufnahmen unter Verwendung von Kontextinformation mit minimaler menschlicher Interaktion. 17
DÖRING, I.: 3D GDI zur Unterstützung der Standortbewertung von Windenergieanlagen (WEA). 337
DRAUSCHKE, M. & REIDELSTÜRZ, P.: Klassifikation von Fernerkundungsdaten zur Früherkennung von Borkenkäferschäden im Bayerischen Wald. 54
DRAUSCHKE, M.: Siehe BRANDENBURGER, W., DRAUSCHKE, M., NGUATEM, W. & MAYER, H.: Detektion von Fensterverdachungen und Gesims in Fassadenbildern und Punktwolken. 202
ELTNER, A.: Konzeption einer Methode zur quantitativen Erfassung von Bodenerosion durch integrierte Aufnahme mit einer Mikrodrohne und einem terrestrischen Laserscanner. 289
EUGSTER, H., GOTTSMANN, F., KÄSER, C., NEBIKER, S. & KOCH, R.: 3DRoadTV-Service – Georeferenzierter 3D-Bilddatendienst für die Infrastrukturpflege und -verwaltung. 344
EUGSTER, H.: Siehe NEBIKER, S., KAMER, K. & EUGSTER, H.: Integrierte Georeferenzierung von luft- und bodengestützten Stereobilddaten im Straßenbereich – Strategien und Ergebnisse. 499
FANG, L., MAKSYMIUK, O., SCHMITT, M. & STILLA, U.: Improvement of Motion Estimation of the Taku Glacier using Spaceborne SAR Images. 62
FARKAS, E.: Siehe RUB, M., FARKAS, E. & WIEDEMANN, A.: Dynamische Erfassung linearer Infrastrukturen unter und über der Erde. 509
FRANZKE, H. J. & WETZEL, H.-U.: Schlussfolgerungen aus der Anwendung der Geo-Fernerkundung in geowissenschaftlichen Projekten in SW-Deutschland. 256
FRITSCH, D., GRIMM, A., KREMER, J., ROTHERMEL, M. & WENZEL, K.: Bilddatenerfassung mit einem Gyrocopter – Erste Erfahrungen zur „Photogrammetrie nach Bedarf". 470
FRITSCH, D.: Siehe CRAMER, M., HAALA, N., ROTHERMEL, M., LEINSS, B. & FRITSCH, D.: UAV-gestützte Datenerfassung für Anwendungen der Landesvermessung – das Hessigheim-Projekt. 450
FRÖHLICH, B.: Siehe BACH, E., FRÖHLICH, B., WALDE, I., HESE, S., SCHMULLIUS, C. & DENZLER, J.: Landbedeckungsklassifikation von Satellitenaufnahmen unter Verwendung von Kontextinformation mit minimaler menschlicher Interaktion. 17
GIAQUINTO, G.: Siehe GIOVENCO, M.M. & GIAQUINTO, G.: Comparison of Image Quality in Optical Radar Fused Image Sets. 79
GINZLER, C., MARTY, M. & BÜHLER, Y.: Grossflächige hochaufgelöste Schneehöhenkarten aus digitalen Stereoluftbildern. 71
GINZLER, C.: Siehe BOESCH, R. & GINZLER, C.: Automatische Berechnung einer nationalen Gehölzmaske. 46
GIOVENCO, M.M. & GIAQUINTO, G.: Comparison of Image Quality in Optical Radar Fused Image Sets. 79
GOTTSMANN, F.: Siehe EUGSTER, H., GOTTSMANN, F., KÄSER, C., NEBIKER, S. & KOCH, R.: 3DRoadTV-Service – Georeferenzierter 3D-Bilddatendienst für die Infrastrukturpflege und -verwaltung. 344
GÖTZ, C.: Siehe DARR, T., GÖTZ, C., TUTTAS, S., BECKER, T. & STILLA, U.: Genauigkeit der inversen Navigation bei hoher Passpunktsichtbarkeit zur Absolutpositionierung eines geführten Messsystems. 279
GÖTZ, C.: Siehe HANEL, A., GÖTZ, C., TUTTAS, S., BECKER, T. & STILLA, U.: Orientierungsgenauigkeit eines Multi-Kamerasystems bei Roboterbewegungen in einer Fahrzeugkarosserie. 298

GRENZDÖRFFER, G. & BOGDANOV, S.: UAS-basierte automatisierte Vogelzählung am Beispiel einer Sturmmöwenkolonie der Vogelschutzinsel Langenwerder. 89
GRIMM, A.: Siehe FRITSCH, D., GRIMM, A., KREMER, J., ROTHERMEL, M. & WENZEL, K.: Bilddatenerfassung mit einem Gyrocopter – Erste Erfahrungen zur „Photogrammetrie nach Bedarf". 470
GRIWODZ, U. & BELLINGHOFF, M.: Traditionelle Luftbilder in 3D. … 210
HAALA, N.: Siehe CRAMER, M., HAALA, N., ROTHERMEL, M., LEINSS, B. & FRITSCH, D.: UAV-gestützte Datenerfassung für Anwendungen der Landesvermessung – das Hessigheim-Projekt. 450
HANEL, A., GÖTZ, C., TUTTAS, S., BECKER, T. & STILLA, U.: Orientierungsgenauigkeit eines Multi-Kamerasystems bei Roboterbewegungen in einer Fahrzeugkarosserie. 298
HANKE, K.: Siehe WALDHÄUSL, P., ABART, G., HANKE, K., HIERMANSEDER, M., KÖNIG, H., MANSBERGER, R., NAVRATIL, G., SCHENNACH, G. & TWAROCH, C.: Sind das landesweite Netzwerk von Grenzen und die Grenzsteine ein Weltkulturerbe? … 534
HASTEDT, H. & LUHMANN, T.: Untersuchungen zum Genauigkeitspotential eines hybriden Kamerasystems aus RIM- und RGB-Kamera. 308
HEINZE, N.: Siehe SAUR, G., HEINZE, N. & KRÜGER, W.: Video-like exploitation of MiSAR sequences from small UAVs. 411
HERDA, H. & BREUER, M.: Different Ways to Process UAV Imagery – Results of Comparative Benchmark Tests. 99
HESE, S.: Siehe BACH, E., FRÖHLICH, B., WALDE, I., HESE, S., SCHMULLIUS, C. & DENZLER, J.: Landbedeckungsklassifikation von Satellitenaufnahmen unter Verwendung von Kontextinformation mit minimaler menschlicher Interaktion. 17
HESE, S.: Siehe MICHAELIS, R. & HESE, S.: Objektbasierte spektrale Analyse und Klassifikation von ölkontaminierter Landoberfläche in Westsibirien mit hochaufgelösten RapidEye Daten. 127
HEURICH, M.: Siehe YAO, W., KRZYSTEK, P. & HEURICH, M.: A study on the detection of understory regenerations in temperate forest areas using airborne full-waveform LiDAR. 437
HIERMANSEDER, M.: Siehe WALDHÄUSL, P., ABART, G., HANKE, K., HIERMANSEDER, M., KÖNIG, H., MANSBERGER, R., NAVRATIL, G., SCHENNACH, G. & TWAROCH, C.: Sind das landesweite Netzwerk von Grenzen und die Grenzsteine ein Weltkulturerbe? 534
HINZ, S.: Siehe HOEGNER, L., WEINMANN, M., JUTZI, B., HINZ, S. & STILLA, U.: Co-registration of Time-of-Flight (TOF) camera generated 3d point clouds and thermal infrared images (IR). 481
HOEGNER, L., WEINMANN, M., JUTZI, B., HINZ, S. & STILLA, U.: Co-registration of Time-of-Flight (TOF) camera generated 3d point clouds and thermal infrared images (IR). 481
HOEGNER, L.: Siehe IWASZCZUK, D., HOEGNER, L. & STILLA, U.: Zuordnung von 3D Gebäudemodellen und IR-Videosequenzen mit Linienverfolgung. 220
HOEKMAN, D. H.: Siehe SCHLUND, M., VON PONCET, F., KUNTZ, S., HOEKMAN, D. H. & SCHMULLIUS, C.: TanDEM-X als Datengrundlage für Referenzkartierungen im Kontext von REDD. 418
HOLLAUS, M.: Siehe MÜCKE, W., SCHROIFF, A., HOLLAUS, M. & PFEIFER, N.: Erkennen von liegendem Totholz in full-waveform Airborne Laserscanning Daten. 401
IWASZCZUK, D., HOEGNER, L. & STILLA, U.: Zuordnung von 3D Gebäudemodellen und IR-Videosequenzen mit Linienverfolgung. 220
JACOBSEN, K.: Erstellung großflächiger Höhenmodelle und Vergleich mit anderen Daten. 489
JUTZI, B.: Siehe HOEGNER, L., WEINMANN, M., JUTZI, B., HINZ, S. & STILLA, U.: Co-registration of Time-of-Flight (TOF) camera generated 3d point clouds and thermal infrared images (IR). 481
KAMER, K.: Siehe NEBIKER, S., KAMER, K. & EUGSTER, H.: Integrierte Georeferenzierung von luft- und bodengestützten Stereobilddaten im Straßenbereich – Strategien und Ergebnisse. 499

KARDOŠ, M.: Siehe SAČKOV, I. & KARDOŠ, M.: Point Clouds from Airborne Laser Scanning and Aerial Images for Forestry Application. 160
KÄSER, C.: Siehe EUGSTER, H., GOTTSMANN, F., KÄSER, C., NEBIKER, S. & KOCH, R.: 3DRoadTV-Service – Georeferenzierter 3D-Bilddatendienst für die Infrastrukturpflege und -verwaltung. 344
KATTENBORN, T.: Forest Change Assessment and Corresponding Driver Analysis in the Magdalena Department, Colombia (1985-2010). 108
KAUFMANN, H.: Siehe BEHLING, R., ROESSNER, S., WETZEL, H.-U. & KAUFMANN, H.: Automated Multi-temporal Landslide Identification in Kyrgyzstan Using Satellite Remote Sensing Data. 244
KENKMANN, T.: Siehe WULF, G., POELCHAU, M. & KENKMANN, T.: Structural analysis of Martian impact craters using several remote sensing data. 263
KNAUER, U. & SEIFFERT, U.: An Approach for Automated Registration of Hyperspectral Images for Boresight Calibration. 271
KOCH, R.: Siehe EUGSTER, H., GOTTSMANN, F., KÄSER, C., NEBIKER, S. & KOCH, R.: 3DRoadTV-Service – Georeferenzierter 3D-Bilddatendienst für die Infrastrukturpflege und -verwaltung. 344
KÖNIG, H.: Siehe WALDHÄUSL, P., ABART, G., HANKE, K., HIERMANSEDER, M., KÖNIG, H., MANSBERGER, R., NAVRATIL, G., SCHENNACH, G. & TWAROCH, C.: Sind das landesweite Netzwerk von Grenzen und die Grenzsteine ein Weltkulturerbe? 534
KREMER, J.: Siehe FRITSCH, D., GRIMM, A., KREMER, J., ROTHERMEL, M. & WENZEL, K.: Bilddatenerfassung mit einem Gyrocopter – Erste Erfahrungen zur „Photogrammetrie nach Bedarf“. 470
KRÜGER, W.: Siehe SAUR, G., HEINZE, N. & KRÜGER, W.: Video-like exploitation of MiSAR sequences from small UAVs. 411
KRZYSTEK, P.: Siehe YAO, W., KRZYSTEK, P. & HEURICH, M.: A study on the detection of understory regenerations in temperate forest areas using airborne full-waveform LiDAR. 437
KUNTZ, S.: Siehe SCHLUND, M., VON PONCET, F., KUNTZ, S., HOEKMAN, D. H. & SCHMULLIUS, C.: TanDEM-X als Datengrundlage für Referenzkartierungen im Kontext von REDD. 418
LEINSS, B.: Siehe CRAMER, M., HAALA, N., ROTHERMEL, M., LEINSS, B. & FRITSCH, D.: UAV-gestützte Datenerfassung für Anwendungen der Landesvermessung – das Hessigheim-Projekt. 450
LEITERER, R., MORSDORF, F. & SCHAEPMAN, M. E.: Characterization of forest understory using multi-temporal full-waveform airborne laser scanning. 384
LINKIEWICZ, M.: Extraktion von senkrechten Fassadenebenen aus 3D-Punktwolken von Schrägluftbildern. 354
LINKIEWICZ, M.: Siehe WIEDEN, A. & LINKIEWICZ, M.: True- Oblique- Orthomosaik aus Schrägluftbildern zur Extraktion von 3D-Geoinformationen. 373
LUBER, A.: Siehe RUEß, D., LUBER, A. & REULKE, R.: Epipolar Geometry of Heterogeneous Stereo Camera Systems. 316
LUHMANN, T.: Siehe HASTEDT, H. & LUHMANN, T.: Untersuchungen zum Genauigkeitspotential eines hybriden Kamerasystems aus RIM- und RGB-Kamera. 308
MAKSYMIUK, O., BRENNER, A. & STILLA, U.: Detection of Stationary Vehicles in Airborne Decimeter Resolution SAR Intensity Images using Morphological Attribute Filters. 391
MAKSYMIUK, O.: Siehe FANG, L., MAKSYMIUK, O., SCHMITT, M. & STILLA, U.: Improvement of Motion Estimation of the Taku Glacier using Spaceborne SAR Images. 62
MAKSYMIUK, O.: Siehe NOVACK, T., MAKSYMIUK, O. & STILLA, U.: A Concept for Guiding the Learning of Conditional Random Fields for the Classification of Urban Areas in SAR Images. 152
MANSBERGER, R.: Siehe WALDHÄUSL, P., ABART, G., HANKE, K., HIERMANSEDER, M., KÖNIG, H., MANSBERGER, R., NAVRATIL, G., SCHENNACH, G. & TWAROCH, C.: Sind das landesweite Netzwerk von Grenzen und die Grenzsteine ein Weltkulturerbe? 534

MANTHEY, K.: About the impact of CCSDS image data compression on Image Quality. 116
MARTY, M.: Siehe GINZLER, C., MARTY, M. & BÜHLER, Y.: Grossflächige hochaufgelöste Schneehöhenkarten aus digitalen Stereoluftbildern. 71
MATTI, E. & NEBIKER, S.: Geometrie- und farbbasierte Punktwolkenklassifizierung von komplexen urbanen Laserscanningszenen mittels überwachter SOM-Klassifikation. 228
MAYER, H.: Siehe BRANDENBURGER, W., DRAUSCHKE, M., NGUATEM, W. & MAYER, H.: Detektion von Fensterverdachungen und Gesims in Fassadenbildern und Punktwolken. 202
MENZ, G.: Siehe THONFELD, F. & MENZ, G.: Entwicklung eines robusten Veränderungsdetektionsverfahrens für räumlich hoch auflösende Satellitendaten. 194
MICHAELIS, R. & HESE, S.: Objektbasierte spektrale Analyse und Klassifikation von ölkontaminierter Landoberfläche in Westsibirien mit hochaufgelösten RapidEye Daten. 127
MIRI, M., BAYER, S. & BUCHER, T.: Evaluation of high resolution digital surface models for single tree extraction approaches in mixed forests. 137
MIRI, M.: Siehe BAYER, S., MIRI, M. & BUCHER, T.: Automatisierte Einzelbaumerfassung auf Basis von hochaufgelösten TrueOrtho-Mosaiken und photogrammetrisch abgeleiteten Oberflächenmodellen. 26
MORSDORF, F.: Siehe LEITERER, R., MORSDORF, F. & SCHAEPMAN, M. E.: Characterization of forest understory using multi-temporal full-waveform airborne laser scanning. 384
MÜCKE, W., SCHROIFF, A., HOLLAUS, M. & PFEIFER, N.: Erkennen von liegendem Totholz in full-waveform Airborne Laserscanning Daten. 401
NAVRATIL, G.: Siehe WALDHÄUSL, P., ABART, G., HANKE, K., HIERMANSEDER, M., KÖNIG, H., MANSBERGER, R., NAVRATIL, G., SCHENNACH, G. & TWAROCH, C.: Sind das landesweite Netzwerk von Grenzen und die Grenzsteine ein Weltkulturerbe? 534
NEBIKER, S., KAMER, K. & EUGSTER, H.: Integrierte Georeferenzierung von luft- und bodengestützten Stereobilddaten im Straßenbereich – Strategien und Ergebnisse. 499
NEBIKER, S.: Siehe EUGSTER, H., GOTTSMANN, F., KÄSER, C., NEBIKER, S. & KOCH, R.: 3DRoadTV-Service – Georeferenzierter 3D-Bilddatendienst für die Infrastrukturpflege und -verwaltung. 344
NEBIKER, S.: Siehe MATTI, E. & NEBIKER, S.: Geometrie- und farbbasierte Punktwolkenklassifizierung von komplexen urbanen Laserscanningszenen mittels überwachter SOM-Klassifikation. 228
NETZBAND, M.: Change detection by means of multitemporal satellite imagery for the Indian Mega-City Hyderabad. 146
NGUATEM, W.: Siehe BRANDENBURGER, W., DRAUSCHKE, M., NGUATEM, W. & MAYER, H.: Detektion von Fensterverdachungen und Gesims in Fassadenbildern und Punktwolken. 202
NOVACK, T., MAKSYMIUK, O. & STILLA, U.: A Concept for Guiding the Learning of Conditional Random Fields for the Classification of Urban Areas in SAR Images. . 152
PETERS, S.: Siehe ADLER, P. & PETERS, S.: Automatische Klassifizierung von Fichten in digitalen 4-Kanal Luftbildern. 9
PFEIFER, N.: Siehe MÜCKE, W., SCHROIFF, A., HOLLAUS, M. & PFEIFER, N.: Erkennen von liegendem Totholz in full-waveform Airborne Laserscanning Daten. 401
POELCHAU, M.: Siehe WULF, G., POELCHAU, M. & KENKMANN, T.: Structural analysis of Martian impact craters using several remote sensing data. 263
REIDELSTÜRZ, P.: Siehe DRAUSCHKE, M. & REIDELSTÜRZ, P.: Klassifikation von Fernerkundungsdaten zur Früherkennung von Borkenkäferschäden im Bayerischen Wald. 54
REULKE, R.: DIN 18714-8 − Anforderungen an die Bildqualität. 526

REULKE, R.: Siehe RUEß, D., LUBER, A. & REULKE, R.: Epipolar Geometry of Heterogeneous Stereo Camera Systems. 316
REULKE, R.: Siehe TRISTRAM, S. & REULKE, R.: 3D-Bildanalyse mit Scene Flow-Algorithmen. 237
ROESSNER, S.: Siehe BEHLING, R., ROESSNER, S., WETZEL, H.-U. & KAUFMANN, H.: Automated Multi-temporal Landslide Identification in Kyrgyzstan Using Satellite Remote Sensing Data. 244
ROTHERMEL, M.: Siehe CRAMER, M., HAALA, N., ROTHERMEL, M., LEINSS, B. & FRITSCH, D.: UAV-gestützte Datenerfassung für Anwendungen der Landesvermessung – das Hessigheim-Projekt. 450
ROTHERMEL, M.: Siehe FRITSCH, D., GRIMM, A., KREMER, J., ROTHERMEL, M. & WENZEL, K.: Bilddatenerfassung mit einem Gyrocopter – Erste Erfahrungen zur „Photogrammetrie nach Bedarf". 470
RUB, M., FARKAS, E. & WIEDEMANN, A.: Dynamische Erfassung linearer Infrastrukturen unter und über der Erde. 509
RUEß, D., LUBER, A. & REULKE, R.: Epipolar Geometry of Heterogeneous Stereo Camera Systems. 316
SAČKOV, I. & KARDOŠ, M.: Point Clouds from Airborne Laser Scanning and Aerial Images for Forestry Application. 160
SAUERBIER, M.: 3D-Gebäudeaufnahme mittels UAV-basierter Luftbilder. 169
SAUR, G., HEINZE, N. & KRÜGER, W.: Video-like exploitation of MiSAR sequences from small UAVs. 411
SCHAEPMAN, M. E.: Siehe LEITERER, R., MORSDORF, F. & SCHAEPMAN, M. E.: Characterization of forest understory using multi-temporal full-waveform airborne laser scanning. 384
SCHENNACH, G.: Siehe WALDHÄUSL, P., ABART, G., HANKE, K., HIERMANSEDER, M., KÖNIG, H., MANSBERGER, R., NAVRATIL, G., SCHENNACH, G. & TWAROCH, C.: Sind das landesweite Netzwerk von Grenzen und die Grenzsteine ein Weltkulturerbe? . 534
SCHLICHTING, A., BRENNER, C. & SCHÖN, S.: Bewertung von inertialen Messsystemen mittels Laserscannern und bekannter Landmarken. 326
SCHLUND, M., VON PONCET, F., KUNTZ, S., HOEKMAN, D. H. & SCHMULLIUS, C.: TanDEM-X als Datengrundlage für Referenzkartierungen im Kontext von REDD. . 418
SCHMITT, M., BRÜCK, A., SCHÖNBERGER, J. & STILLA, U.: Potential of Airborne Single-Pass Millimeterwave InSAR Data for Individual Tree Recognition. 426
SCHMITT, M.: Siehe FANG, L., MAKSYMIUK, O., SCHMITT, M. & STILLA, U.: Improvement of Motion Estimation of the Taku Glacier using Spaceborne SAR Images. 62
SCHMULLIUS, C.: Siehe BACH, E., FRÖHLICH, B., WALDE, I., HESE, S., SCHMULLIUS, C. & DENZLER, J.: Landbedeckungsklassifikation von Satellitenaufnahmen unter Verwendung von Kontextinformation mit minimaler menschlicher Interaktion. 17
SCHMULLIUS, C.: Siehe SCHLUND, M., VON PONCET, F., KUNTZ, S., HOEKMAN, D. H. & SCHMULLIUS, C.: TanDEM-X als Datengrundlage für Referenzkartierungen im Kontext von REDD. 418
SCHÖN, S.: Siehe SCHLICHTING, A., BRENNER, C. & SCHÖN, S.: Bewertung von inertialen Messsystemen mittels Laserscannern und bekannter Landmarken. 326
SCHÖNBERGER, J.: Siehe SCHMITT, M., BRÜCK, A., SCHÖNBERGER, J. & STILLA, U.: Potential of Airborne Single-Pass Millimeterwave InSAR Data for Individual Tree Recognition. 426
SCHROIFF, A.: Siehe MÜCKE, W., SCHROIFF, A., HOLLAUS, M. & PFEIFER, N.: Erkennen von liegendem Totholz in full-waveform Airborne Laserscanning Daten. 401
SEIFFERT, U.: Siehe KNAUER, U. & SEIFFERT, U.: An Approach for Automated Registration of Hyperspectral Images for Boresight Calibration. 271
SEITZ, R.: Siehe STRAUB, C., STEPPER, C., SEITZ, R. & WASER, L.T.: Möglichkeiten einer stratifizierten Schätzung von forstlichen Kenngrößen unter Verwendung von amtlichen Stereo-Luftbildern, Orthophotos und Laserscannerdaten. 186
STEBNER, K.: Siehe WIEDEN, A. & STEBNER, K.: Referenzorientierung für Bilddaten aus Mehrkopfkamerasystemen. 518

STEPPER, C.: Siehe STRAUB, C., STEPPER, C., SEITZ, R. & WASER, L.T.: Möglichkeiten einer stratifizierten Schätzung von forstlichen Kenngrößen unter Verwendung von amtlichen Stereo-Luftbildern, Orthophotos und Laserscannerdaten. 186
STERN, O.: Abbildende und nichtabbildende Geländespektrometrie zur Untersuchung von Stressphänomenen an Buchenpflanzen. 176
STILLA, U.: Siehe DARR, T., GÖTZ, C., TUTTAS, S., BECKER, T. & STILLA, U.: Genauigkeit der inversen Navigation bei hoher Passpunktsichtbarkeit zur Absolutpositionierung eines geführten Messsystems. 279
STILLA, U.: Siehe FANG, L., MAKSYMIUK, O., SCHMITT, M. & STILLA, U.: Improvement of Motion Estimation of the Taku Glacier using Spaceborne SAR Images. 62
STILLA, U.: Siehe HANEL, A., GÖTZ, C., TUTTAS, S., BECKER, T. & STILLA, U.: Orientierungsgenauigkeit eines Multi-Kamerasystems bei Roboterbewegungen in einer Fahrzeugkarosserie. 298
STILLA, U.: Siehe HOEGNER, L., WEINMANN, M., JUTZI, B., HINZ, S. & STILLA, U.: Co-registration of Time-of-Flight (TOF) camera generated 3d point clouds and thermal infrared images (IR). 481
STILLA, U.: Siehe IWASZCZUK, D., HOEGNER, L. & STILLA, U.: Zuordnung von 3D Gebäudemodellen und IR-Videosequenzen mit Linienverfolgung. 220
STILLA, U.: Siehe MAKSYMIUK, O., BRENNER, A. & STILLA, U.: Detection of Stationary Vehicles in Airborne Decimeter Resolution SAR Intensity Images using Morphological Attribute Filters. 391
STILLA, U.: Siehe NOVACK, T., MAKSYMIUK, O. & STILLA, U.: A Concept for Guiding the Learning of Conditional Random Fields for the Classification of Urban Areas in SAR Images. 152
STILLA, U.: Siehe SCHMITT, M., BRÜCK, A., SCHÖNBERGER, J. & STILLA, U.: Potential of Airborne Single-Pass Millimeterwave InSAR Data for Individual Tree Recognition. 426
STILLA, U.: Siehe TUTTAS, S. & STILLA, U.: Rekonstruktion von Fenstern aus Schrägsicht-ALS Punktwolken zur Anreicherung von Gebäudemodellen. 363
STRAUB, C., STEPPER, C., SEITZ, R. & WASER, L.T.: Möglichkeiten einer stratifizierten Schätzung von forstlichen Kenngrößen unter Verwendung von amtlichen Stereo-Luftbildern, Orthophotos und Laserscannerdaten. 186
THONFELD, F. & MENZ, G.: Entwicklung eines robusten Veränderungs-detektionsverfahrens für räumlich hoch auflösende Satellitendaten. 194
TRISTRAM, S. & REULKE, R.: 3D-Bildanalyse mit Scene Flow-Algorithmen. 237
TUTTAS, S. & STILLA, U.: Rekonstruktion von Fenstern aus Schrägsicht-ALS Punktwolken zur Anreicherung von Gebäudemodellen. 363
TUTTAS, S.: Siehe DARR, T., GÖTZ, C., TUTTAS, S., BECKER, T. & STILLA, U.: Genauigkeit der inversen Navigation bei hoher Passpunktsichtbarkeit zur Absolutpositionierung eines geführten Messsystems. 279
TUTTAS, S.: Siehe HANEL, A., GÖTZ, C., TUTTAS, S., BECKER, T. & STILLA, U.: Orientierungsgenauigkeit eines Multi-Kamerasystems bei Roboterbewegungen in einer Fahrzeugkarosserie. 298
TWAROCH, C.: Siehe WALDHÄUSL, P., ABART, G., HANKE, K., HIERMANSEDER, M., KÖNIG, H., MANSBERGER, R., NAVRATIL, G., SCHENNACH, G. & TWAROCH, C.: Sind das landesweite Netzwerk von Grenzen und die Grenzsteine ein Weltkulturerbe? 534
VON PONCET, F.: Siehe SCHLUND, M., VON PONCET, F., KUNTZ, S., HOEKMAN, D. H. & SCHMULLIUS, C.: TanDEM-X als Datengrundlage für Referenzkartierungen im Kontext von REDD. 418
WALDE, I.: Siehe BACH, E., FRÖHLICH, B., WALDE, I., HESE, S., SCHMULLIUS, C. & DENZLER, J.: Landbedeckungsklassifikation von Satellitenaufnahmen unter Verwendung von Kontextinformation mit minimaler menschlicher Interaktion. 17
WALDHÄUSL, P., ABART, G., HANKE, K., HIERMANSEDER, M., KÖNIG, H., MANSBERGER, R., NAVRATIL, G., SCHENNACH, G. & TWAROCH, C.: Sind das landesweite Netzwerk von Grenzen und die Grenzsteine ein Weltkulturerbe? 534

WASER, L.T.: Siehe STRAUB, C., STEPPER, C., SEITZ, R. & WASER, L.T.: Möglichkeiten einer stratifizierten Schätzung von forstlichen Kenngrößen unter Verwendung von amtlichen Stereo-Luftbildern, Orthophotos und Laserscannerdaten. ……..... 186
WEINMANN, M.: Siehe HOEGNER, L., WEINMANN, M., JUTZI, B., HINZ, S. & STILLA, U.: Co-registration of Time-of-Flight (TOF) camera generated 3d point clouds and thermal infrared images (IR). ……. 481
WENZEL, K.: Siehe FRITSCH, D., GRIMM, A., KREMER, J., ROTHERMEL, M. & WENZEL, K.: Bilddatenerfassung mit einem Gyrocopter – Erste Erfahrungen zur „Photogrammetrie nach Bedarf“. …… 470
WETZEL, H.-U.: Siehe BEHLING, R., ROESSNER, S., WETZEL, H.-U. & KAUFMANN, H.: Automated Multi-temporal Landslide Identification in Kyrgyzstan Using Satellite Remote Sensing Data. ……..... 244
WETZEL, H.-U.: Siehe FRANZKE, H. J. & WETZEL, H.-U.: Schlussfolgerungen aus der Anwendung der Geo-Fernerkundung in geowissenschaftlichen Projekten in SW-Deutschland. ……..... 256
WIEDEMANN, A.: Siehe RUB, M., FARKAS, E. & WIEDEMANN, A.: Dynamische Erfassung linearer Infrastrukturen unter und über der Erde. ……..... 509
WIEDEN, A. & LINKIEWICZ, M.: True- Oblique- Orthomosaik aus Schrägluftbildern zur Extraktion von 3D-Geoinformationen. ……. 373
WIEDEN, A. & STEBNER, K.: Referenzorientierung für Bilddaten aus Mehrkopfkamerasystemen. … 518
WULF, G., POELCHAU, M. & KENKMANN, T.: Structural analysis of Martian impact craters using several remote sensing data. …….. ... 263
YAO, W., KRZYSTEK, P. & HEURICH, M.: A study on the detection of understory regenerations in temperate forest areas using airborne full-waveform LiDAR. … . 437
ZAKŠEK, K.: Siehe BECHTEL, B., ZAKŠEK, K. & BÖHNER, J.: Hochauflösende Datensätze urbaner Oberflächentemperaturen zur Beobachtung der städtischen Wärmeinsel. . 36