PFG

Photogrammetrie Fernerkundung Geoinformation

Journal for Photogrammetry, Remote Sensing
and Geoinformation Science

Organ der Deutschen Gesellschaft für Photogrammetrie,
Fernerkundung und Geoinformation (DGPF) e. V.

Jahrgang 2012, Heft 5

Hauptschriftleiter:
Prof. Dr.-Ing. Wolfgang Kresse

Schriftleiter:
Prof. Dr.-Ing. Stefan Hinz, Prof. Dr. rer.nat. Carsten Jürgens,
Prof. Dr. rer.nat. Lars Bernard, Privatdozent Dr. techn. Franz
Rottensteiner und Dr.-Ing. Eckhardt Seyfert

E. Schweizerbart'sche Verlagsbuchhandlung
(Nägele u. Obermiller) Stuttgart 2012

Deutsche Gesellschaft für Photogrammetrie, Fernerkundung
und Geoinformation (DGPF) e. V.
Gegründet 1909

Die *Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation* (DGPF) e. V. unterstützt als Mitglieds- bzw. Trägergesellschaft die folgenden Dachverbände:

International Society
for Photogrammetry
and Remote Sensing

Deutsche Arbeits-
gemeinschaft für
Mustererkennung e.V.

GeoUnion
Alfred-Wegener-Stiftung

Herausgeber:

Präsident: Prof. Dr. Thomas Kolbe, Technische Universität Berlin, Institut für Geodäsie und Geoinformationstechnik, Straße des 17. Juni 135, 10623 Berlin, Germany, Tel.: +49-30-314-23274
Geschäftsstelle: Dr. Klaus-Ulrich Komp, c/o EFTAS Fernerkundung Technologietransfer GmbH, Oststraße 2–18, 48145 Münster, Germany, e-mail: klaus.komp@eftas.com

Published by: E. Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller), Johannesstraße 3A, 70176 Stuttgart, Germany, Tel.: +49-711 351456-0, Fax: +49-711 351456-99, e-mail: mail@schweizerbart.de
Internet: http://www.schweizerbart.de
♾ Gedruckt auf alterungsbeständigem Papier nach ISO 9706-1994

ISSN 1432-8364

Science Citation Index Expanded (also known as SciSearch®) Journal Citation Reports/Science Edition
Hauptschriftleiter: Prof. Dr.-Ing. Wolfgang Kresse, Hochschule Neubrandenburg, Fachbereich Landschaftsarchitektur, Geoinformatik, Geodäsie und Bauingenieurwesen, Brodaer Straße 2, 17033 Neubrandenburg, Germany, e-mail: kresse@hs-nb.de
Schriftleiter: Prof. Dr.-Ing. Stefan Hinz, Karlsruher Institut für Technologie – KIT, Institut für Photogrammetrie und Fernerkundung, Englerstraße 7, 76131 Karlsruhe, Germany, e-mail: stefan.hinz@ipf.uni-karlsruhe.de, Prof. Dr. rer. nat. Carsten Jürgens, Ruhr-Universität Bochum, Geographisches Institut, Gebäude NA7/133, 44780 Bochum, Germany, e-mail: carsten.juergens@rub.de, Prof. Dr. rer.nat. Lars Bernard, Technische Universität Dresden, Fachrichtung Geowissenschaften, Helmholtzstraße 10, 01062 Dresden, Germany, e-mail: lars.bernard@tu-dresden.de, Privatdozent Dr. techn. Franz Rottensteiner, Leibniz Universität Hannover, Institut für Photogrammetrie und GeoInformation, Nienburger Straße 1, 30167 Hannover, Germany, e-mail: rottensteiner@ipi.uni-hannover.de und Dr.-Ing. Eckhardt Seyfert, Landesvermessung und Geobasisinformation Brandenburg, Heinrich-Mann-Allee 103, 14473 Potsdam, Germany, e-mail: eckhardt.seyfert@geobasis-bb.de

Erscheinungsweise: 6 Hefte pro Jahrgang.
Bezugspreis im Abonnement: € 219,– pro Jahrgang. Mitglieder der DGPF erhalten die Zeitschrift kostenlos. Der Online-Zugang ist im regulären Subskriptionspreis enthalten.

Anzeigenverwaltung: E. Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller), Johannesstraße 3A, 70176 Stuttgart, Germany, Tel.: +49-711 351456-0; Fax: +49-711 351456-99.
e-mail: mail@schweizerbart.de, Internet: http://www.schweizerbart.de
Bernhard Harzer Verlag GmbH, Westmarkstraße 59/59a, 76227 Karlsruhe, Germany, Tel.: +49-721 944020, Fax: +49-721 9440230, e-mail: Info@harzer.de, Internet: www.harzer.de
Printed in Germany by Tutte Druckerei GmbH, 94121 Salzweg bei Passau, Germany.

PFG – Jahrgang 2012, Heft 5
Inhaltsverzeichnis

Editorial

Originalbeiträge

Beitrag aus Wissenschaft und Praxis

Mitteilungen

Berichte von Veranstaltungen

Persönliches

Hochschulnachrichten

VDV-Preis

Buchbesprechung

Zusammenfassungen der „Originalbeiträge“ und der „Beiträge aus Wissenschaft und Praxis“ (deutsch und englisch) sind auch verfügbar unter www.dgpf.de/neu/pfg/ausgaben.htm

PFG 2012 / 5, 0499–0500
Stuttgart, October 2012

Photogrammetric Image Analysis

Automated extraction of objects from remotely sensed data has become an important topic of research in photogrammetry, computer vision, remote sensing, and geoinformation science. The need to discuss recent developments in methodological research and the potential of various data sources has formed the background for the latest edition of the ISPRS Conference on *Photogrammetric Image Analysis (PIA)* which was organized by the Department of Photogrammetry and Remote Sensing, Technische Universität München (TUM), and held at the TUM main campus between October 5 and 7, 2011 (see also the workshop report by Boris Jutzi in PFG issue 1/2012). This was the fourth workshop of that series, following the successful events in 1999, 2003 and 2007, all of them hosted by the Department of Photogrammetry and Remote Sensing at TUM, which has become a constant in the calendar of photogrammetry and remote sensing and has continually attracted about 120 participants from all around the world.

In parallel to the workshop, a call for papers was distributed to experts in the field of photogrammetric image analysis to submit papers to this special issue of PFG. The topics of the conference and, consequently, of this special issue, were related to the four ISPRS working groups (WG) co-operating in the organisation of the scientific program of the conference, namely WG I/2 "LIDAR, SAR and Optical Sensors" WG III/1 "Pose Estimation and Surface Reconstruction", WG III/4 "Complex Scene Analysis and 3D Reconstruction" and WG III/5 "Image Sequence Analysis". As a result, five papers (out of 10 submitted contributions) which have undergone a rigorous peer review process are published in this special issue of PFG.

The first paper, submitted by Jochen Meidow, is about a method for efficient least squares adjustment of multiple loops in sequences. Dead reckoning, i.e., relative orientation for sequences, suffers from inevitable drift. The latter can often be strongly reduced by constraints induced by closing loops. Meidow's method can deal with multiple loops simultaneously in a statistically optimal way. Additionally, the employed minimal representation without singularities leads to an efficient implementation. Results for 2D image mosaicking and 3D trajectory determination clearly demonstrate the method's potential.

In the second paper, Dorota Iwaszczuk et al. describe a technique for the orientation of images using 3D building models as ground control. Their method is based on a matching process that detects correspondences between the edges of a building model and image edges, using random sample consensus (RANSAC) to distinguish correct matches from incorrect ones. The experiments presented by the authors show that their method delivers promising results for infrared images taken from a helicopter.

The third paper written by Hui Ju et al. deals with another aspect of image registration, namely with the precise alignment of airborne laserscanner (ALS) data and optical (aerial or satellite) imagery. Based on an analysis of the shortcomings of existing methods, they develop a new technique consisting of two steps: First, the optical and an ALS intensity image are coarsely aligned using a method based on the fast fourier transform. The transformation thus established is improved by homologous points obtained by scale and rotation invariant region descriptor matching. Initial results of the authors show that pixel level accuracy can be achieved by their method.

The fourth paper written by Daniel Muhle et al. focuses on identifying correspondences for matching 3D point clouds. The proposed algorithm is based on the basic principles of matching by utilizing distinctive feature descriptions and extends them in a way that they can be used to identify corresponding 3D points in sparsely populated and varying point clouds. The investigations show that a reliable matching of 3D points is possible and that the position uncertainty of a 3D point does not

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0133
1432-8364/12/0133 $ 0.50

seem to have a strong influence on the matching quality.

In the last contribution, Matthias Plaue et al. address the extraction of pedestrian trajectories and the according local density fields from videos in order to support the macroscopic modelling of human crowds. The results of an experiment conducted by the authors show that the proposed technique yields both spatially as well as temporally smooth density fields close to the standard definition of density at all scales.

In summary, both the workshop and the selection of papers of this special issue show a number of trends. (1) The use of image sequences, be it from airborne or from (mobile or fixed) terrestrial platforms, is gaining more and more attention in the community. This induces work both on the orientation of such sequences, which requires automation due to the sheer amount of images to be dealt with, and on tasks related to the detection of both static and moving objects. (2) The use of UAVs as aerial platforms capable of closing the "scale gap" between aerial and terrestrial image acquisition is becoming more and more common. As a consequence of the limitations of Micro-UAVs with respect to the payload they can carry, methods capable of dealing with images of lower quality than standard aerial imagery need to be investigated. (3) Methods conceptionally capable of delivering results in real time become important, especially in the field of video surveillance. (4) Modern photogrammetric research tries to tackle problems that were out of scope of 'traditional' photogrammetry, e.g. detection and tracking of pedestrians, or the derivation of behavioural patterns from tracked pedestrians in crowds. (5) Computer vision continues to have a strong impact on the current directions of photogrammetric research.

Finally, we would like to sincerely thank everybody involved in the preparation of this special issue. We are very grateful to all staff at Technische Universität München for their invaluable help in organizing the workshop, to the workshop participants who have made the meeting a success, and to the authors of this special issue for making available their excellent papers, and for keeping a tough timeline. We thank the editor-in-chief of PFG, Wolfgang Kresse, and his team for all the freedom we have had when assembling this special issue and for the help they extended to us in the technical preparation.

Uwe Stilla, Helmut Mayer,
Michael Schmitt, München,
Boris Jutzi, Karlsruhe &
Franz Rottensteiner, Hannover

This PFG-issue includes seven other articles that relate to all three topics of PFG, photogrammetry, remote sensing, and geoinformation science. Hernández-Lopez et al. present a study about the radiometric calibration of Leica's ADS40. Immitzer et al. evaluate the new spectral channels of WorldView-2 for tree specification (in German). Lausch et al. investigate the influence of scale on the classification of hyperspectral image data. Mannel & Price show that multi-seasonal imagery of a multi-spectral camera leads to better classification results than imagery data from a hyperspectral camera with significantly more bands, but recorded at the same time. Arroyo Ohori et al. address an often encountered issue of geometry data and present a new method for repairing errors in planar partitions like cadastral parcels or land use coverages. Cavegn & Nebiker demonstrate a new method for the automatic mapping of road signs (in German). A report about the new spatial data infrastructure of the Austrian civil air traffic control Austro Control finalizes the suite of articles (also in German).

Wolfgang Kresse

PFG 2012 / 5, 0501 – 0510
Stuttgart, October 2012

Article

Efficient Multiple Loop Adjustment for Computer Vision Tasks

JOCHEN MEIDOW, Ettlingen

Keywords: loop adjustment, image mosaicking, simultaneous localization and mapping

Summary: Applications performing dead reckoning generally suffer from the inevitable drift of the solutions due to measurement noise and other systematic errors. When dealing with moving sensors such as cameras or laser scanners, loop closing represents one of the most important strategies to compensate for these effects. Imposing loop constraints reduces the uncertainties that arise and permits the construction of consistent maps, mosaics, or scene reconstructions. We propose and demonstrate an approach to correct multiple loops simultaneously by least squares adjustments in 2D and 3D. The results are statistically optimal, since all individual uncertainties and correlations are considered. The use of minimal transformation representations without singularities enables efficient implementation. Furthermore, we address the selection of consistent loop constraints by exploiting the natural order of the sensor datasets given by the acquisition process. The feasibility and the efficiency of the approach are demonstrated using synthetic data reflecting an indoor exploration scenario and by examples for terrestrial and airborne video mosaicking.

Zusammenfassung: Die Lösungen von Anwendungen, die eine Koppelnavigation durchführen, weisen in der Regel eine unvermeidbare Drift aufgrund von Messrauschen und restsystematischen Fehlern auf. Konzeptionell stellen Schleifenschlüsse eine wichtige Strategie dar, um diese Effekte zu kompensieren. Die Schleifenbedingungen reduzieren die auftretenden Ungenauigkeiten und gestatten den Aufbau von konsistenten Karten, Mosaiken oder Szenenrekonstruktionen. Wir stellen einen allgemeinen Ansatz vor, um mehrere auftretende Schleifen gleichzeitig in 2D und 3D durch Ausgleichungen zur korrigieren. Die Ergebnisse sind statistisch optimal, da alle individuellen Unsicherheiten und Korrelationen berücksichtigt werden. Die verwendeten minimalen Repräsentationen für Transformationen in 2D bzw. 3D weisen keine Singularitäten auf und ermöglichen eine effiziente Umsetzung. Des Weiteren wird eine Strategie für die Auswahl konsistenter Bedingungsgleichungen vorgestellt, welche sich aufgrund der zeitlichen Reihenfolge der Daten im Erfassungsprozess ergibt. Die Durchführbarkeit und die Effizienz des Ansatzes wird anhand synthetischer Daten demonstriert, die eine Innenraumexploration simulieren, sowie mit Beispielen für die Mosaikierungen von terrestrisch und luftgetragen erfassten Videobildern.

1 Introduction

1.1 *Motivation*

Today, navigation, image mosaicking, and the exploration of unknown environments by simultaneous localization and mapping (SLAM) are common tasks in robotics, computer vision and photogrammetry. Exteroceptive sensors, such as cameras or laser scanners, measure the proximity of objects to the sensor's frame of reference. The use of moving sensors permits the acquisition of environments for which one builds or updates maps. Furthermore, the sensors enable the determination of ego-motion in 3D by establishing correspondences between consecutively captured datasets. The latter is of special interest when external references such as landmarks with known coordinates or reference signals such as GPS-signals are not available. In this situation, dead reckoning has to be applied, which is subject to inevitable drift due to measurement noise and

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0134 1432-8364/12/0134 $ 2.50

remaining systematic errors. The same holds for the mosaicking of video streams: Assuming planar scenes, the chaining of consecutive image pairs suffers from accumulation of random feature-tracking errors and systematic errors due to imperfect calibration. This becomes evident when loops are present: drift appears in discrepancies or gaps at the joints.

Conceptually, loop closing is one of the most important strategies to compensate for drift and to obtain more precise, globally consistent results. Whenever a system recognizes places already visited, the discrepancies that occur should be distributed over the covered path and the scene reconstruction.

1.2 *Related Work*

Contributions to loop adjustment are manifold and can be found in the robotics, photogrammetry, and computer vision literature. The formulation and solution of optimization or interpolation tasks, the use of appropriate parameterizations and the construction of consistent sets of constraints are all topics in which loop adjustment has been considered.

For single loops, the parameter corrections sought can be determined by distributing the updates proportionally over the loop trajectory. In 3D, this can be accomplished by considering the so-called minimal-length trajectory between two poses, which is given by a straight line in the corresponding tangent space. This interpolation can be controlled using scalar weights reasonably chosen proportional to the uncertainties at hand and leads to a minimal bending of the trajectory (Dubbelman et al. 2010). In the context of video mosaicking, an optimization procedure is proposed by Caballero et al. (2007) for instance. Applying an extendend Kalman filter, an optimization procedure is employed for updating the loop homographies. Within each step of the filter update a normalization of the homography matrices is necessary to fix the scale of the homogeneous representation.

A solution to reduce the errors in a 3D network of observed transformations is provided in Grisetti et al. (2007). The optimization is performed using a variant of gradient descent. In doing so, the updates are considered separately for the rotational component (spherical linear interpolation) and the translational component. Results for experiments with thousands of involved transformations are shown. When deriving constraints from a network of observations, the resulting set of equations may not necessarily be consistent. When the network is represented as a graph, this leads to a need for so-called fundamental cycle bases (Unnikrishnan & Kelly 2002b).

The work closest to ours is that of Estrada et al. (2005), which covers loop detection and adjustment for the 2D SLAM case. The squared Mahalanobis distance is calculated for hypothesis testing and nonlinear constrained least-squares optimization by sequential quadratic programming is used to impose multiple loop constraints simultaneously. However, this work does not address the need for finding a consistent set of independent constraints.

1.3 *Contribution*

Essentially, we extend the approach of Estrada et al. (2005) to the 3D case. We propose an efficient method to build large and globally consistent mosaics and sensor trajectories within the same framework. As proposed by Meidow (2011) in the context of video mosaicking, this is achieved by exploiting the power of algebraic projective geometry, by using minimal representations without singularities (Strasdat et al. 2010), and by compiling consistent sets of loop constraints (Unnikrishnan & Kelly 2002b). For the latter we exploit the natural order of the datasets (video images or laser scans) captured by a moving sensor. The approach allows the simultaneous adjustment of multiple loops in a batch process after solving the place recognition task. By considering image-to-image transformations or pose changes as observations, we chose an adjustment model with constraints for these transformations parameters only, which leads to small equation systems to be solved. The stochastic model involved in these tasks rigorously incorporates the uncertainties of the transformation parameters.

For the image alignment task, this approach assumes planar scenes or a fixed projection centre. We assume uncalibrated cameras with, however, straight-line preserving optics, i.e., we assume the lens distortion to be negligible.

2 Theoretical Background

2.1 *Notation and Preliminaries*

Homogeneous entities are denoted by upright boldface letters, e.g. $\mathbf{x}$ or $\mathbf{H}$, Euclidean vectors and matrices with italic boldface letters, e.g. $\boldsymbol{l}$ or $\boldsymbol{R}$. For homogeneous coordinates "=" means an assignment or an equivalence up to a common scale factor $\lambda \neq 0$.

For the minimal parameterizations of a 2D homography or 3D motion, we exploit the power series

$$\exp(\boldsymbol{K}) = \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{K}^k = I + \boldsymbol{K} + \frac{1}{2}\boldsymbol{K}^2 + \dots, \tag{1}$$

which is the power series for square matrices analogous to the power series for the scalar exponential function. For the analytical computation of Jacobians, we will frequently use the rule

$$\operatorname{vec}(\boldsymbol{ABC}) = \left(\boldsymbol{C}^\mathsf{T} \otimes \boldsymbol{A}\right)\operatorname{vec}(\boldsymbol{B}) \tag{2}$$

and its specializations. Here, the vec operator stacks all columns of a matrix and $\otimes$ denotes the Kronecker product. The skew-symmetric matrix $\boldsymbol{S}(\boldsymbol{a})$ built from a 3-vector $\boldsymbol{a}$ induces a cross-product.

2.2 *Parameterizations*

We use homogeneous coordinates to represent homographies in 2D and motion in 3D. In doing so, 3D motion can be considered as a special homography in 3D, which paves the way to a common framework to treat 2D and 3D tasks in the same way. Homographies form a group. Thus, one can "undo" a transformation by computing and applying the inverse transformation (matrix inversion). The concatenation or chaining of two or more transformations is carried out by matrix multiplication.

A planar projective transformation is a linear transformation on homogeneous 3-vectors represented by a non-singular 3×3 matrix $\mathbf{H} = (H_{ij})$, cf. (Hartley & Zisserman 2004):

$$\begin{pmatrix} u' \\ v' \\ w' \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix}, \tag{3}$$

or, more briefly, $\mathbf{x}' = \mathbf{H}\,\mathbf{x}$. This transformation is unique up to scale and has therefore eight degrees of freedom. It can be written in inhomogeneous form as

$$x' = \frac{H_{11}x + H_{12}y + H_{13}}{H_{31}x + H_{32}y + H_{33}} \quad \text{and}$$

$$y' = \frac{H_{21}x + H_{22}y + H_{23}}{H_{31}x + H_{32}y + H_{33}} \tag{4}$$

with $\mathbf{x} = [x, y, 1]^\mathsf{T}$ and $\mathbf{x}' = [x', y', 1]^\mathsf{T}$.

A rigid-body transform in $\mathbb{R}^3$ can be expressed by a 4×4 transformation matrix applied to homogeneous 4-vectors:

$$\mathbf{H} = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \mathbf{0}^\mathsf{T} & 1 \end{bmatrix} \text{ with } \boldsymbol{R} \in SO(3) \text{ and } \boldsymbol{t} \in \mathbb{R}^3, \tag{5}$$

where $SO(3)$ is the Lie group of rotation matrices. The motion matrices (5) form a smooth manifold and, therefore, the Special Euclidean group $SE(3)$. Its operator is matrix multiplication.

In homogeneous coordinates, all transformations are realized by multiplication. Thus, linearization in a multiplicative manner is straightforward (Förstner 2010). With a multiplicative expansion, a nonlinear update reads

$$\mathbf{H} = \Delta\mathbf{H} \cdot \mathbf{H}_0 \tag{6}$$

with an approximate homography $\mathbf{H}_0$ and the update $\Delta\mathbf{H}$ within the iteration sequence.

The power series for the matrix exponential (1) relates Euclidean updates to representations in the tangent space. For 2D homographies, this is the zero-trace matrix

$$\boldsymbol{K}(\boldsymbol{q}_{2\mathrm{D}}) = \begin{pmatrix} k_1 & k_4 & k_7 \\ k_2 & k_5 & k_8 \\ k_3 & k_6 & -k_1 - k_5 \end{pmatrix}, \tag{7}$$

which depends linearly on the eight correction parameters $\boldsymbol{q}_{2D} = [k_1, \ldots, k_8]^\mathsf{T}$ (Begelfor & Werman 2005). Requiring all matrices to have determinant one, thus constituting a Lie group $SL(3)$, leads to an easy linearization scheme (Förstner 2012). For 3D motion, the relation is given by the so-called twist representation

$$\boldsymbol{K}(\boldsymbol{q}_{3D}) = \begin{bmatrix} \boldsymbol{S}(\boldsymbol{r}) & \Delta \boldsymbol{t} \\ \mathbf{0}^\mathsf{T} & 0 \end{bmatrix} \tag{8}$$

with the six motion parameters $\boldsymbol{q}_{3D} = [\boldsymbol{r}^\mathsf{T}, \Delta\boldsymbol{t}^\mathsf{T}]^\mathsf{T}$ comprising three rotation parameters $\boldsymbol{r}$ and a translational update $\Delta\boldsymbol{t}$ (Bregler & Malik 1998).

During the iterative estimation we do not update the homography parameters or the motion parameters in tangent space, but rather the approximate transformations $\mathbf{H}_0$ by mapping the updates onto the manifold $SL(3)$ or $SE(3)$, respectively

$$\mathbf{H}_0^{(\nu+1)} = \exp\big(\boldsymbol{K}(\boldsymbol{q})\big) \cdot \mathbf{H}_0^{(\nu)}, \tag{9}$$

where ν denotes the iteration step. Thus, we end up with the estimates $\hat{\mathbf{H}} = \mathbf{H}_0$, $\hat{\boldsymbol{q}} = \mathbf{0}$, and $\boldsymbol{\Sigma}_{\hat{q}\hat{q}}$ for the covariance matrix of the estimated transformation parameters $\boldsymbol{q}_{2D}$ and $\boldsymbol{q}_{3D}$, respectively.

2.3 *Sequential Links, Cross Links and Cumulative Transformations*

For establishing loops, we consider two different types of image alignment and motion: sequential links and cross links (Turkbeyler & Harris 2010). *Sequential links* are given by the chaining of consecutive, i.e., temporally adjacent, sensor data. For mosaicking and visual SLAM, these transformations are usually determined by tracking or matching corresponding image features and, for laser scanning data, by the registration of point clouds. For the efficient computation of Jacobians within the adjustment process, we will frequently use *cumulative transforms* of sequential links

$$^{k}\mathbf{H}_1 = \prod_{l=1}^{k-1} {}^{l+1}\mathbf{H}_l = {}^{k}\mathbf{H}_{k-1} \cdot \ldots \cdot {}^{3}\mathbf{H}_2 \cdot {}^{2}\mathbf{H}_1, \quad k = 2, \ldots, n \tag{10}$$

in analogy to cumulative sums and cumulative products. They start with the very first homography $^{2}\mathbf{H}_1$ of each sequence and result from the kinematic chains. The concatenation of motions is represented by multiplication from the left.

Cross links can be established whenever a system recognizes an already visited place or by visual inspection. Here, the search for correspondences is much more challenging and sophisticated methods such as SIFT (scale invariant feature transform) for images (Lowe 2004) or ICP (iterative closest point) for point clouds (Besl & McKay 1992) are needed to cope with changed sensor aspects.

3 Multiple Loop Adjustment

After the detection of loop closure events by place recognition and the determination of the corresponding cross link parameters, the mosaic loops or motion involved in these events need to be adjusted. Taking all uncertainties of the transformation parameters into account, we pursue a statistically rigorous and optimal approach (section 3.2). The approach is suboptimal compared to the joint estimation of all transformations together with the common 2D and 3D points, e.g. Szeliski (2006), Triggs et al. (2000), but enables the efficient treatment of large datasets. For multiple loop adjustment, the selection of a proper and consistent set of constraints is crucial (section 3.1).

3.1 *Choice of Cycle Basis*

The set of constraints used within the adjustment must not be redundant, i.e., the equations must be independent. This can be realized by considering a topological graph as a description of the links given by chained motions or image transforms. Fig. 1 shows an example for a sensor path with three established cross links. The vertices of the graph denote locally captured data elements, i.e., images or point clouds, and the edges denote transformations. Additional cross links have been marked by arrows. In this graph, it is possible to establish six different cycles (loops) that do not form an independent set. A complete cover of the

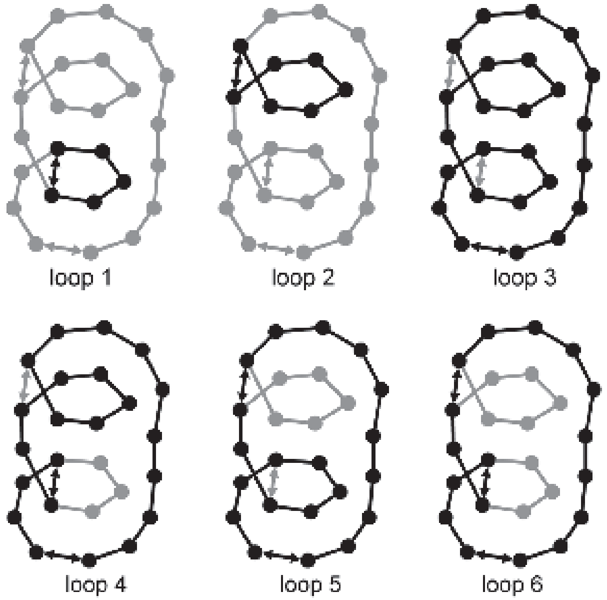


Fig. 1: Six possible loops for a given sensor path with three cross links marked by arrows. Loops 1, 2 and 6 constitute the smallest-size basis, whereas the loops 1, 2 and 3 form the proposed set of loops established along the sensor's path with one cross link each.

graph is given, for instance, by loops 1, 2 and 6. In graph theory, such an independent and complete cycle cover is called a fundamental cycle basis (UNNIKRISHNAN & KELLY 2002b).

The solution of the least squares problem proposed in section 3.2 does not depend on the choice of the cycle basis, since the linearized equations corresponding to any cycle basis form an independent set in the space of constraints provided the approximate values are the same. Choosing the cycle basis with the smallest size reduces the computational costs since the Jacobians involved are sparser. Here "smallest size" refers to the number of edges to be traversed in total (shortest loop lengths). However, the problem of finding a cycle basis of smallest size has been proven to be NP-hard for an arbitrary graph (THOMASSEN 1997). Therefore and for the sake of simplicity, we establish loops along the sensor path immediately whenever they occur, i.e., each loop contains exactly one cross link. This approach is simple and easy and exploits the natural temporal and spatial order of the data provided by the acquisition process. In Fig. 1 the loops 1, 2 and 6 constitute the smallest-size cycle basis, whereas the loops 1, 2 and 3 are loops established immediately whenever the sensor revisits a place.

3.2 *Imposing Loop Constraints*

The proposed loop adjustments are performed by solving a standard least squares problem. A solution to this problem is obtained by an adaption of sequential quadratic programming (TRIGGS et al. 2000). Considering the transformation parameters to be given but uncertain observations, the solution is equivalent to the adjustment with constraints for observations only (KOCH 1999, MCGLONE et al. 2004). For the sake of completeness, we briefly summarize here the adjustment model used and its corresponding procedure. Then the loop constraints and the corresponding Jacobians are presented.

With the given block-diagonal covariance matrix $\mathbf{\Sigma}_{pp}$ of the unconstrained (observed) parameters $[{}^2\boldsymbol{q}_1{}^\mathsf{T}, {}^3\boldsymbol{q}_2{}^\mathsf{T}, \dots {}^n\boldsymbol{q}_{n-1}{}^\mathsf{T}]^\mathsf{T}$ (set of sequential links) augmented by the observed cross link parameters, the optimization problem is given by

$$\Omega = \left(\boldsymbol{p} - \hat{\boldsymbol{p}}\right)^{\mathsf{T}} \boldsymbol{\Sigma}_{pp}^{-1} \left(\boldsymbol{p} - \hat{\boldsymbol{p}}\right) \;\rightarrow\; \min \tag{11}$$

subject to $\boldsymbol{c}(\hat{\boldsymbol{p}}) = \mathbf{0}$. Here, the sought constrained parameters are $\hat{\boldsymbol{p}}$ and the loop constraints are $\boldsymbol{c}(\hat{\boldsymbol{p}}) = \mathbf{0}$. Linearization of the constraints yields

$$\boldsymbol{c}(\hat{\boldsymbol{p}}) = \boldsymbol{c}(\boldsymbol{p}_0) + \boldsymbol{C}\boldsymbol{\Delta}\hat{\boldsymbol{p}} \tag{12}$$

with the Jacobian $\boldsymbol{C}$. Then the estimator is

$$\hat{\boldsymbol{p}}^{(\nu+1)} = \boldsymbol{p} + \boldsymbol{\Sigma}_{pp}\boldsymbol{C}^{\mathsf{T}} \left(\boldsymbol{C}\boldsymbol{\Sigma}_{pp}\boldsymbol{C}^{\mathsf{T}}\right)^{-1} \boldsymbol{c}_0 \tag{13}$$

with $\boldsymbol{C}\boldsymbol{\Sigma}_{pp}\boldsymbol{C}^{\mathsf{T}}$ being the covariance matrix of the contradictions $\boldsymbol{c}(\boldsymbol{p})$ and

$$\boldsymbol{c}_0 = \boldsymbol{C}\left(\hat{\boldsymbol{p}}^{(\nu)} - \boldsymbol{p}\right) - \boldsymbol{c}\left(\hat{\boldsymbol{p}}^{(\nu)}\right) \tag{14}$$

at each iteration v. The covariance matrix for the estimates (13)

$$\boldsymbol{\Sigma}_{\hat{p}\hat{p}} = \boldsymbol{\Sigma}_{pp} - \boldsymbol{\Sigma}_{pp}\boldsymbol{C}^{\mathsf{T}} \left(\boldsymbol{C}\boldsymbol{\Sigma}_{pp}\boldsymbol{C}^{\mathsf{T}}\right)^{-1} \boldsymbol{C}\boldsymbol{\Sigma}_{pp} \tag{15}$$

is fully occupied in general and therefore computable for problems of moderate size only.

The constraints for the chained homographies or integrated motion are simply

$${}^{j}\mathbf{H}'_{i} - \prod_{l=i}^{j-1} {}^{l+1}\mathbf{H}_{l} = \mathbf{O} \tag{16}$$

for a single loop with $j > i$. The constraints take one cross link (see section 3.1) into account, denoted by a single prime. Please note that all homography matrices are unambiguous due the determinant one constraint or the Euclidean motion parameterization. An alternative formulation of the loop constraints (16) is

$${}^{i}\mathbf{H}'_{j} \cdot \prod_{l=i}^{j-1} {}^{l+1}\mathbf{H}_{l} = \boldsymbol{I} \tag{17}$$

(Unnikrishnan & Kelly 2002a), whereby the Jacobians w.r.t. the cross link depend on the sequential links, too.

The constraints have to be fulfilled for the adjusted parameters. During the optimization process we update the approximate homographies according to (9). Thus, the approximate values for the parameters $\boldsymbol{p}$ are simply zeros and we can expand the constraints with the corresponding multiplicative updates $\Delta\mathbf{H} = \exp(\boldsymbol{K}(\boldsymbol{q})) = \boldsymbol{I}$. For the vectorization of the constraints (16), we consider this for the cross link (i, j) and a sequential link $(k, k+1)$

$${}^{j}\Delta\mathbf{H}'_{i} \cdot {}^{j}\mathbf{H}'_{i} - {}^{j}\mathbf{H}_{k+1}\left({}^{k+1}\Delta\mathbf{H}_{k} \cdot {}^{k+1}\mathbf{H}_{k}\right){}^{k}\mathbf{H}_{i} = \mathbf{O} \tag{18}$$

With $\Delta\mathbf{h} = \mathrm{vec}(\Delta\,\mathbf{H})$ and the rule (2), the vectorization of (18) yields nine equations for 2D homographies and 16 equations for 3D motion respectively:

$$\begin{aligned}\boldsymbol{c} &= \left({}^{j}\mathbf{H}'^{\mathsf{T}}_{i} \otimes \boldsymbol{I}\right){}^{j}\Delta\mathbf{h}'_{i} - \left({}^{k+1}\mathbf{H}^{\mathsf{T}}_{i} \otimes {}^{j}\mathbf{H}_{k+1}\right)\cdot {}^{k+1}\Delta\mathbf{h}_{k} \\ &= \mathbf{0}\end{aligned} \tag{19}$$

The parameters $\boldsymbol{q}$ of a single transformation are related to the corresponding homography $\mathbf{H}$ by $\Delta\mathbf{H} = \exp(\boldsymbol{K}(\boldsymbol{q})) \approx \boldsymbol{I} + \boldsymbol{K}$ for small values $\boldsymbol{q}$. Thus, the linearization is

$$\Delta\mathbf{h} \approx \mathrm{vec}(\boldsymbol{I}) + \mathrm{vec}\left(\boldsymbol{K}(\boldsymbol{q})\right) = \mathrm{vec}(\boldsymbol{I}) + \boldsymbol{G}\boldsymbol{q} \tag{20}$$

with the constant transformation matrix $\boldsymbol{G}$. For 2D homographies, $\boldsymbol{G}$ is simply the 9×8 matrix

$$\boldsymbol{G}_{2\mathrm{D}} = \begin{bmatrix} \boldsymbol{I}_8 \\ -1,0,0,0,-1,0,0,0 \end{bmatrix} \tag{21}$$

(Meidow 2011). For 3D pose changes, we apply a 16×6 transformation matrix $\boldsymbol{G}_{3\mathrm{D}}$ composed of ones, zeros and minus ones.

The constraints (19) are linearily dependent and a reduction to a set of six and eight constraints, respectively, is necessary. This can be achieved by deleting equations or by computing a linear combination. Applying the pseudo inverses of the matrices $\boldsymbol{G}$, yields reduced constraints $\bar{\boldsymbol{c}} = \boldsymbol{G}^{+}\boldsymbol{c}$.

The Jacobians are then

$$\boldsymbol{C}^{(c)} = \frac{\partial\bar{\boldsymbol{c}}(\hat{\boldsymbol{q}}')}{\partial\hat{\boldsymbol{q}}'} = \frac{\partial\bar{\boldsymbol{c}}(\hat{\boldsymbol{q}}')}{\partial\Delta\mathbf{h}'}\frac{\partial\Delta\mathbf{h}'}{\partial\hat{\boldsymbol{q}}'} = \boldsymbol{G}^{+}\left({}^{j}\mathbf{H}'^{\mathsf{T}}_{i} \otimes \boldsymbol{I}\right)\boldsymbol{G} \tag{22}$$

w.r.t. the parameters of the cross link and

$$\boldsymbol{C}^{(s)} = \frac{\partial \overline{\boldsymbol{c}}(\hat{\boldsymbol{q}})}{\partial \hat{\boldsymbol{q}}} = \boldsymbol{G}^{+}\left({}^{k+1}\mathbf{H}_i^{\mathsf{T}} \otimes {}^{j}\mathbf{H}_{k+1}\right)\boldsymbol{G}$$
$$= \boldsymbol{G}^{+}\left[\left({}^{i}\mathbf{H}_1^{-\mathsf{T}} \cdot {}^{k+1}\mathbf{H}_1^{\mathsf{T}}\right) \otimes \left({}^{j}\mathbf{H}_1 \cdot {}^{k+1}\mathbf{H}_1^{-1}\right)\right]\boldsymbol{G} \tag{23}$$

w.r.t. the parameters of a sequential link. Note that all Jacobians can be computed with a cross link or cumulative transforms only.

For multiple loops to be adjusted, the entire Jacobian $\boldsymbol{C}$ is then a matrix for the S sequential links

$$\boldsymbol{C}_{\text{seq}} = \begin{bmatrix} \boldsymbol{C}_{11}^{(s)} & \boldsymbol{C}_{12}^{(s)} & \cdots & \boldsymbol{C}_{1S}^{(s)} \\ \boldsymbol{C}_{21}^{(s)} & \boldsymbol{C}_{22}^{(s)} & \cdots & \boldsymbol{C}_{2S}^{(s)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{C}_{L1}^{(s)} & \boldsymbol{C}_{L2}^{(s)} & \cdots & \boldsymbol{C}_{LS}^{(s)} \end{bmatrix} \tag{24}$$

augmented by a usually much smaller block diagonal matrix

$$\boldsymbol{C}_{\text{cross}} = \operatorname{Diag}\left(\boldsymbol{C}_1^{(c)}, \boldsymbol{C}_2^{(c)}, \ldots, \boldsymbol{C}_L^{(c)}\right), \tag{25}$$

where L denotes the number of cross links, i.e., loops, and $D = 8$ parameters for 2D homographies and $D = 6$ parameters for 3D motion, respectively. Thus, the size of the Jacobian

$$\boldsymbol{C} = \left[\boldsymbol{C}_{\text{seq}} \mid \boldsymbol{C}_{\text{cross}}\right] \tag{26}$$

is $DL \times D(S+L)$. Its shape is usually very flat, since $S >> L$ holds. The sparseness of the Jacobian $\boldsymbol{C}_{\text{seq}}$ depends on the chosen cycle basis. The size of the matrix to be inverted in (13) depends only on the number of loops L. Note that no special treatment of the homographies not being part of any loop is necessary. No case-by-case analysis is necessary, since the corresponding estimates for the corrections in (13) will simply be zeros.

4 Experiments

In the following, we demonstrate the feasibility of the approach for the tasks of video mosaicking (terrestrial and airborne, section 4.1) and 3D exploration, the latter possibly by simultaneous localization and mapping (SLAM) with the help of video cameras or laser scanners (section 4.2).

4.1 *Video Mosaicking*

For each image of a video stream, salient points have been extracted by the Förstner operator (Förstner & Gülch 1987), and tracked in the corresponding subsequent image by the Lucas-Kanade tracker (Lucas & Kanade 1981). The respective positional uncertainties – represented by covariance matrices – have been determined for incorporation into the adjustment. The correspondences of the cross

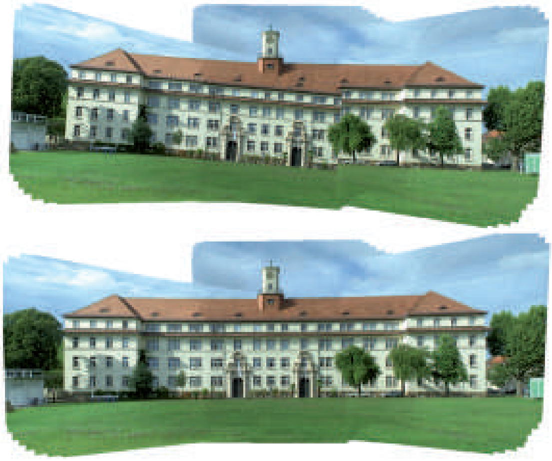


Fig. 2: Mosaicking of 355 video images of size 720 × 576 pixel. Top: Image alignment by considering consecutive links only. Bottom: Result after the adjustment of two loops.

links have then been established by applying the scale invariant feature transform SIFT (Lowe 2004) in combination with random sample consensus (Fischler & Bolles 1981).

Fig. 2 shows the result for the acquisition of a façade by an uncalibrated ordinary video camera before and after adjustment of two loops. The camera was moved in a pattern according to the ∞ sign. The mosaic obtained by applying the consecutive homographies only is shown in Fig. 2, top, and reveals numerous discrepancies due to drift. Subsequent adjustment yields a consistent result. The gauge freedoms for the mosaics have been fixed by selecting the mean cumulative homography (10) as reference transformation.

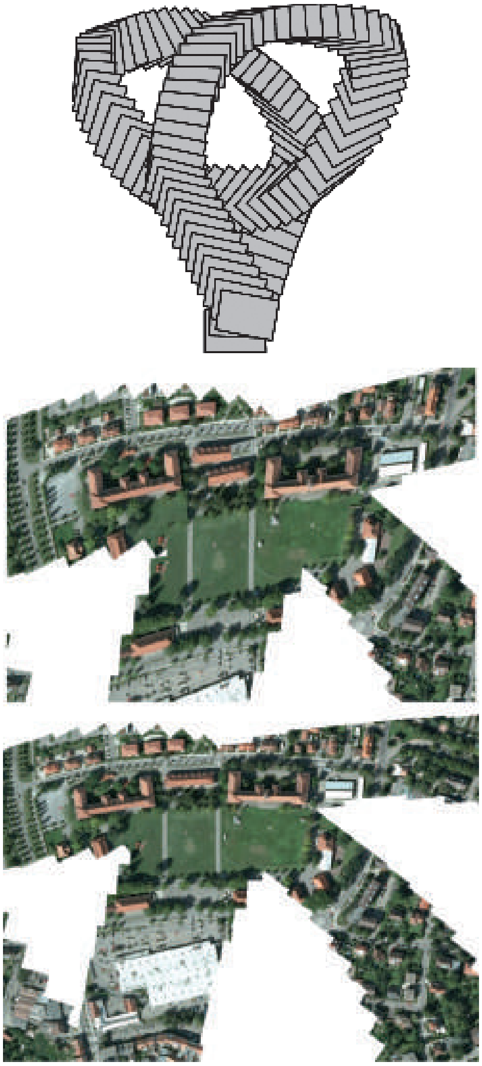


Fig. 3: Mosaicking of approx. 1,200 images with 480 × 320 pixel resolution. Top: Footprints of the adjusted images (every 10th drawn). Middle: Aerial mosaic built by using the estimated sequential links only (detail). Bottom: Mosaic after adjusting four loops (detail). Image data: ©GeoContent provided by Google.

For the provision of airborne imagery we abused a virtual globe system as an image source and camera simulator. This guarantees that the model assumptions are valid – all uncertainties stem from tracking and matching, respectively. The images have a rather poor resolution of 480 × 320 pixel. The given camera path features varying heights above the ground as well as changing roll, pitch, and yaw angles. During the flight, the camera carried out a 180-degree-turn around its optical axis, see Fig. 3. Four loop closing events have been identified by visual inspection. The computing time for the adjustment was 1.4 seconds on a standard CPU at 1.59 GHz with non-optimized Matlab code in four iterations.

4.2 *Simulated Indoor Exploration*

To the author's knowledge, there are no publicly available 3D data that contain sensor motion for sequential links and additional cross links with uncertainty information as well as ground truth for evaluation, neither derived from visual data nor from laser scans. Therefore, we restricted our investigation to synthetic data reflecting already derived navigation solutions obtained by dead reckoning. In doing so, we are omitting the effects of specific feature tracking and/or feature matching techniques, which are outside the scope of this investigation.

Fig. 4 summarizes the experimental setup, the simulated data and the results. The chosen sensor trajectory imitates indoor exploration by a forward looking moving video camera. The camera visits two rooms, moves along corridors, and finally revisits its starting point. The results could be provided by a visual SLAM algorithm, e.g. a sliding window bundle adjustment. Gaussian noise has been added to the successive relative motion. This results in the camera path depicted in Fig. 4, centre. The gap between the starting and end points of the path is huge. The result after closing and adjusting the three loops given by the fundamental cycle basis depicted in Fig. 1 (Loops 1, 2 and 3) is clearly more consistent. The computing time for the adjustment was 0.1 seconds at 1.59 GHz with a non-optimized Matlab implementation.

Most of the absolute sensor poses change considerably during the adjustment. However, the relative consecutive motions change only slightly due to the least squares approach. Locally, the trajectory is hardly affected while, globally, consistency is achieved. For the determination of the empirical accuracy of the adjustment result, we compute the test statistic

$$T = \frac{1}{R}\left(\hat{\boldsymbol{p}} - \boldsymbol{p}^{(\mathrm{ref})}\right)^{\mathsf{T}} \left(\boldsymbol{\Sigma}_{\hat{p}\hat{p}} + \boldsymbol{\Sigma}_{pp}^{(\mathrm{ref})}\right)^{-1} \left(\hat{\boldsymbol{p}} - \boldsymbol{p}^{(\mathrm{ref})}\right) \sim F_{R,\infty} \qquad (27)$$

based on the Mahalanobis distance w.r.t. reference values of the parameters, i.e., the ground truth (McGlone et al. 2004, Dickscheid et al. 2008). The test statistic is $F_{R,\infty}$-distributed with $R = DS$ degrees of freedom. We obtain $T = 1.0174$ which is lower than the 0.95-quantile 1.0176 of the F-distribution. Thus the accuracy potential has been fully exploited.

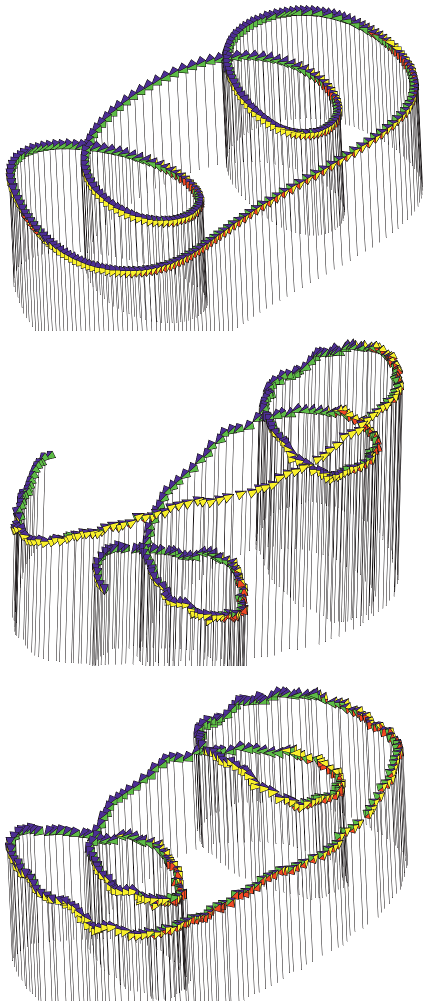


Fig. 4: Synthetic data with 3,000 camera poses (every 10th is plotted) simulating indoor exploration by a moving camera with subsequent place recognition and loop adjustment. Top: Ground truth with three loops. Middle: Noise and drift added to the consecutive motion. Bottom: Trajectory after adjustment of the loops.

5 Conclusions

We have proposed a statistically optimal approach to adjust multiple loops simultaneously in particular for image alignment in 2D and simultaneous localization and mapping in 3D. The adjustment model chosen and the parameterizations used enable an efficient implementation: Hundreds of images or poses can be adjusted within a second. To come up with a set of consistent loop constraints, we exploit the natural order of datasets given by the sequential acquisition process of moving mapping sensors.

Visual inspection of the test results clearly reveals the capability of the approach to compensate for inevitable drift, which indicates that the approach supports the generation of consistent mosaics and scene reconstructions. For path planning during exploration, such loop closing implies an obvious strategy: loops should be detected and closed as soon as possible and wherever possible.

Acknowledgements

The author would like to thank the anonymous reviewers for their assistance in improving the manuscript.

References

Begelfor, E. & Werman, M., 2005: How to put Probabilities on Homographies. – IEEE Trans. on Pattern Analysis and Machine Intelligence **27** (10): 1666–1670.

Besl, P.J. & McKay, N.D., 1992: A Method for Registration of 3-D Shapes. – IEEE Trans. on Pattern Analysis and Machine Intelligence **14** (2): 239–256.

Bregler, C. & Malik, J., 1998: Tracking People with Twists and Exponential Maps. – Computer Vision and Pattern Recognition (CVPR **’98**): 8–15.

Caballero, F., Merino, L., Ferruz, J. & Ollero, A., 2007: Homography Based Kalman Filter for Mosaic Building. Applications to UAV Position Estimation. – International Conf. on Robotics and Automation (ICRA **2007**): 2004–2009.

Dickscheid, T., Läbe, Th. & Förstner, W., 2008: Benchmarking Automatic Bundle Adjustment Results. – **21st** Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS) (B3a): 7–12.

Dubbelman, G., Esteban, I. & Schutte, K., 2010: Efficient trajectory bending with applications to loop closure. – Intelligent Robots and Systems (IROS **’10**): 1–7.

Estrada, C., Neira, J. & Tardós, J.D., 2005: Hierarchical SLAM : Real-Time Accurate Mapping of Large Environments. – IEEE Transactions on Robotics **21** (4): 588–596.

Fischler, M.A. & Bolles, R.C., 1981: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. – Communications of the ACM **24** (6): 381–395.

Förstner, W. & Gülch, E., 1987: A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features. – ISPRS Intercommission Workshop on Fast Processing of Photogrammetric Data: 281–305.

Förstner, W., 2010: Minimal Representations for Uncertainty and Estimation in Projective Spaces. – Asian Conf. on Computer Vision **II:** 619–633.

Förstner, W., 2012: Minimal Representations for Testing and Estimation in Projective Spaces. – PFG **2012** (3): 209–220.

Grisetti, G., Grzonka, S., Stachniss, C., Pfaff, P. & Burgard, W., 2007: Efficient Estimation of Accurate Maximum Likelihood Maps in 3D. – Intelligent Robots and Systems: 3472–3478, San Diego, CA, USA.

Hartley, R. & Zisserman, A., 2004: Multiple View Geometry in Computer Vision. – Second edition, Cambridge University Press, Cambridge.

Koch, K.-R., 1999: Parameter Estimation and Hypothesis Testing in Linear Models. – Second Edition, Springer, Berlin.

Lowe, D.G., 2004: Distinctive Image Features from Scale-Invariant Keypoints. – International Journal of Computer Vision **60** (2): 91–110.

Lucas, B.D. & Kanade, T., 1981: An Iterative Image Registration Technique with an Application to Stereo Vision. – Image Understanding Workshop: 212–130.

McGlone, J.C., Mikhail, E.M. & Bethel, J., 2004: Manual of Photogrammetry. – 5th ed., American Society of Photogrammetry and Remote Sensing.

Meidow, J., 2011: Efficient Video Mosaicking by Multiple Loop Closing. – U. Stilla et al. (eds.): Photogrammetric Image Analysis 2011, Lecture Notes in Computer Science **6952:** 1–12, Springer.

Strasdat, H., Montiel, J.M.M. & Davison, A.J., 2010: Scale Drift-Aware Large Scale Monocular SLAM. Robotics: Science and Systems.

Szeliski, R., 2006: Image alignment and stitching: A tutorial. – Foundations and Trends in Computer Graphics and Computer Vision **2** (1): 1–104.

Thomassen, C., 1997: On the complexity of finding a minimum cycle cover of a graph. – SIAM Journal on Computing **26** (3): 675–677.

Triggs, B., McLauchlan, P., Hartley, R. & Fitzgibbon, A., 2000: Bundle adjustment – a modern synthesis. – Triggs et al. (eds.): Vision Algorithms: Theory and Practice, Lecture Notes in Computer Science **1883:** 298–372, Springer.

Turkbeyler, E. & Harris, C., 2010: Mapping of Movement to Aerial Mosaic with Geo-Location Information. – Optro **2010**.

Unnikrishnan, R. & Kelly, A., 2002a: A Constrained Optimization Approach to Globally Consistent Mapping. – Intelligent Robots and Systems (IROS **’02**) 1: 564–569.

Unnikrishnan, R. & Kelly, A., 2002b: Mosaicing Large Cyclic Environments for Visual Navigation in Autonomous Vehicles. – IEEE International Conf. on Robotics and Automation 2002 (ICRA **’02**) 4: 4299–4306.

Address of the Author:

Dr.-Ing. Jochen Meidow, Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB, D-76275 Ettlingen, Tel.: +49 7243 992-117, Fax: -299, e-mail: jochen.meidow@iosb.fraunhofer.de

Manuskript eingereicht: März 2012
Angenommen: Juni 2012

PFG 2012 / 5, 0511 – 0522
Stuttgart, October 2012

Article

Line based Matching of Uncertain 3D Building Models with IR Image Sequences for Precise Texture Extraction

Dorota Iwaszczuk, Ludwig Hoegner, Michael Schmitt & Uwe Stilla, München

Keywords: line matching, infrared images, building model, texture

Summary: Thermal building textures are used for the detection of damaged or weak spots in the insulation of building hulls. These textures can be extracted from directly geo-referenced oblique airborne infrared (IR) image sequences by projecting a 3D building model into the images. However, the direct geo-referencing is often not sufficiently accurate and the projected 3D model does not match the structures in the image. Thus we present a technique with the main goal of finding the best fit between the 3D building model and the IR image sequence. For this purpose we correct exterior orientation via line based matching. We assign image lines to projected model lines based on the distance and angle between them. The maximum distance and maximum angle between assigned lines is given by the uncertainties in the projected lines, which is derived from the uncertainties in the 3D building model and from the uncertainties in the camera position and orientation by error propagation. Then we use the random sample consensus (RANSAC) to remove incorrect correspondences. The correspondences selected by RANSAC are adjusted using the least squares method. In the adjustment we consider both uncertainties in the model and in the image features. To evaluate the presented method we test it running the algorithm among the set of images and visually assess the improvement.

Zusammenfassung: *Linienbasierte Zuordnung von unsicheren 3D-Gebäudemodellen mit IR Bildsequenzen zur präzisen Texturgewinnung.* Thermische Gebäudetexturen werden für die Detektion von Schwach- und Schadstellen in der Isolation von Gebäudehüllen eingesetzt. Solche Texturen können aus den direkt georeferenzierten, infraroten (IR) Luftbildsequenzen gewonnen werden, indem das 3D Gebäudemodell ins IR Bild projiziert wird. Das direkte Georeferenzieren ist jedoch oft nicht genau genug und das projizierte 3D Modell stimmt nicht mit den Bildstrukturen überein. Deswegen wird hier eine Technik präsentiert, mit dem Ziel, die beste Anpassung zwischen dem IR Bild und dem 3D Gebäudemodell zu finden. Dafür korrigieren wir die Parameter der äußeren Orientierung mittels linienbasierter Zuordnung. Die Bildlinien werden den projizierten Modelllinien aufgrund der Entfernung und aufgrund des Winkels zwischen ihnen zugeordnet. Die maximale Distanz und der maximale Winkel zwischen zugeordneten Linien ergeben sich aus der Unsicherheit der projizierten Linien, die aus den Unsicherheiten der Kameraposition und Orientierung und aus den Unsicherheiten der 3D-Gebäudemodelle mittels Fehlerfortpflanzung berechnet werden. Danach verwenden wir RANdom SAmple Consensus (RANSAC), um die fehlerhaften Korrespondenzen auszusortieren. Die von RANSAC ausgewählten Korrespondenzen werden mit der Methode der kleinsten Quadrate ausgeglichen. Bei der Ausgleichung werden zwei Unsicherheiten berücksichtigt: die des 3D Gebäudemodells sowie die der Bildmerkmale. Um die präsentierte Methode zu evaluieren, wird der Algorithmus in einer Bildsequenz angewandt und die Verbesserung visuell beurteilt.

DOI: 10.1127/1432-8364/2012/0135 1432-8364/12/0135 $ 2.75

1 Introduction

Thermal inspections of buildings contribute to the detection of damaged and weak spots in the building structure. Three dimensional (3D) spatial referencing of the captured images helps the interpretation of the data, especially for large area inspection using images taken by a mobile mapping system. Façades seen from the street level can by captured by a camera mounted on a vehicle (Hoegner et al. 2007) and the roofs are imaged from a flying platform. Using multi aspect airborne oblique images, the missing walls in inner yards are captured. The spatial reference is achieved by combining infrared images with 3D building models via texture mapping. The existing 3D building models can be projected into the infrared (IR) images and the building textures can be extracted. For the projection the exterior and interior orientation parameters of the camera need to be known. These parameters can be determined directly from the navigation device and camera system calibration parameters (camera calibration, boresight and lever-arm calibration). Unfortunately, the direct geo-referencing is often not sufficiently accurate and the model does not match the structures in the image. To refine the registration a model-to-image matching should be carried out.

2 Related Work

In literature the model-to-image matching problem for airborne imagery is frequently addressed and many methods for solving the problem have been presented. Früh et al. (2004) propose line matching based on slope and proximity by testing different random camera positions. However, as Ding & Zakhor (2008) mentioned, this method requires high computational effort. Hsu et al. (2000) search for the best camera position by minimizing the disagreement between projected features and features detected in the image. Other authors propose methods for coarse orientation which use vanishing points (Ding & Zakhor 2008, Förstner 2010). These methods lead to faster results, but they assume so called "Manhattan scenes", where many horizontal and vertical lines can be detected in the image. In some works (Vosselman 1992, Eugster & Nebiker 2009) relational matching is applied, which does not only consider the agreement between an image feature and a model feature, but also takes the relations between features into account.

Methods can also be differentiated based on which image features they use for matching. Some authors propose points (Ding & Zakhor 2008, Avbelj et al. 2010), but most works consider lines as more natural for building structures and use them for co-registration (Debevec 1996, Früh et al. 2004, Schenk 2004, Eugster & Nebiker 2009). In some papers hybrid methods employing points and lines at the same time are presented (Zhang et al. 2005, Tian et al. 2008).

Only few authors take the uncertainty of the applied 3D models into account. Usually the models used for these kinds of matching are stored in a parameterized form (Sester & Förstner 1989, Lowe 1991, Debevec et al. 1996, Vosselman 1998), which is very useful for 3D reconstruction, because the parameterized models represent simple buildings or building primitives. However, reconstructed building models are frequently modelled by polyhedra and stored in a format supporting polyhedral models, e.g. CityGML. Thus, for matching images with existing 3D building models other methods should be developed.

In our research we present a method for matching polyhedral 3D building models with thermal IR imagery. Because we do not use the parametric model representation, we cannot directly apply the stochastic model proposed by Sester & Förstner (1989). In contrast to Hoegner et al. (2007) we do not use terrestrial image sequences but instead match images taken by a camera mounted on a flying platform. We extend our previous work (Avbelj et al. 2010) as we do not use the intersection points but apply a line based matching and consider both, the uncertainties in the 3D building model and those in the image features. The goal of this technique is an improvement of the camera position, so that the best fit between the 3D building model and the image structure is achieved.

3 Line based Matching

In the presented research the main goal is to find the best fit between the existing 3D building model and the IR image sequence. We assume a calibrated camera system with GPS/INS navigation, known interior orientation of the IR camera, and known lever-arm and bore-sight parameters. Accordingly, the camera position and orientation are determined quite accurately. This allows the projection of the 3D building models into the image using the collinearity equations. Unfortunately the projected 3D model often does not match the structures in the image. On one hand this is a consequence of the remaining error after the camera calibration and the camera pose (position and orientation) determination. On the other hand the mismatch can be related to the rolling shutter effect. Rolling shutter occurs in cameras with a line-wise readout system when the photographed objects or the camera are moving. Such systems are often used in IR cameras, thus the rolling shutter effect is particularly noticeable in IR images taken by a moving camera and can lead to a "shrinking" or "stretching" effect on the imaged objects. Another reason for the mismatches can be the unmodelled lens distortions of the IR camera or unmodelled vibrations of the camera. To reduce all these remaining errors we propose a matching procedure between the 3D models and each IR image frame.

Our method is based upon the least squares method. We use the stochastic model to embed the uncertainties of the extracted image lines and the uncertainty of the 3D building models. The uncertainty of the image lines is given by the uncertainty of the extraction process of the lines, and the uncertainty of the 3D building models is a result of the inaccurate extraction process and generalization. In this research we use 3D building models generated from aerial images. Therefore we assume the roofs to be more reliable, because they were directly measured during the 3D reconstruction. Besides, the radiometric properties of the ground (sidewalks and streets) in thermal IR are similar to the properties of the walls, so that the edges between have very low contrast and often cannot be extracted.

3.1 *Line Parameterization*

Polyhedral 3D building models are stored as a set of polygons defined by lines (edges) and points (corners). Lines are natural building structures, which can be detected in the image. To use these lines for co-registration a mathematical description of a line is needed. Typically a line in 3D is described by a direction vector $\boldsymbol{v}$ and a point $\boldsymbol{P}$. For this description any point $\boldsymbol{P}$ belonging to the line can be used, thus there is more than one set of parameters describing one line. To solve this problem ROBERTS (1988) introduced a line representation which is unique and unambiguous. This line representation was discussed, varied and applied in photogrammetric context by SCHENK (2004).

Roberts' line representation is based on two orientation parameters (α, θ) and two positional parameters (X_s, Y_s). The azimuth α and zenith θ can be deduced from the spherical coordinates of vector $\boldsymbol{v}$. (X_s, Y_s) are coordinates of the intersection point with the plane $X'Y'$, where $X'Y'Z'$ is the rotated original coordinate system XYZ, so that the Z'-axis is parallel to the line (Fig. 1a). All equations required to calculate these parameters are given by SCHENK (2004) and MEIERHOLD et al. (2008). Each point on the line can be expressed as

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} X_s \cos\alpha\cos\theta - Y_s \sin\alpha + t\cos\alpha\sin\theta \\ X_s \sin\alpha\cos\theta + Y_s \cos\alpha + t\sin\alpha\sin\theta \\ -X_s \sin\theta + t\cos\theta \end{pmatrix}. \tag{1}$$

As we can see, all lines, including vertical and horizontal ones, are defined using (1). This parameterization uses four parameters, which is the number of degrees of freedom of a 3D line. We use this representation of lines to express the edges of the 3D building model.

Similarly, we searched for a 2D line representation which uses the minimal number of parameters and is defined in all cases. For this purpose the representation with angle γ and distance p can be used:

$$x\cos\gamma + y\sin\gamma - p = 0\,, \tag{2}$$

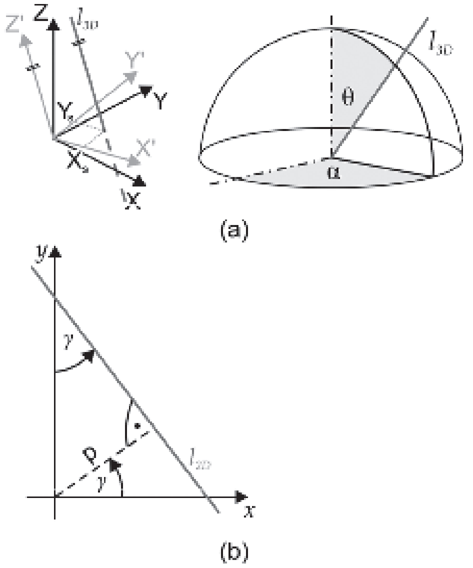


Fig. 1: Parametrization of a line – graphical representation a) in 3D space using 4 parameters; b) in 2D space using 2 parameters.

p denotes the shortest distance from the line to the origin of the coordinate system, and γ denotes direction angle of the normal vector to the line (Fig. 1b).

3.2 *Assignment of Corresponding Lines*

Assignment of corresponding lines is carried out in the 2D image space. The model lines are projected into the image using coarse exterior orientation parameters obtained by direct georeferencing, and for each model line potentially corresponding image lines are found. In this work we apply an assignment based on relative position and orientation. We calculate a buffer S_i around every projected visible model line segment. The width $2 \cdot \Delta p$ of S_i is given by $\Delta p = 3 \cdot \sigma_p$, where σ_p is the uncertainty of the parameter p (distance) of the projected model line, calculated by propagating the uncertainty of the camera position and the uncertainty of the model. For all image line segments within the buffer we calculate the angle difference $\Delta\gamma_{ij}$ (Fig. 2). All model line segments for which $\Delta\gamma_{ij}$ is smaller than a threshold $\Delta\gamma_{max}$ are accepted as correspondences. $\Delta\gamma_{max}$ is calculated as $\Delta\gamma_{max} = 3 \cdot \sigma_\gamma$, where σ_γ is the uncertainty of the parameter (angle) γ of the projected model line, which is calculated by propagating the uncertainty of the camera position and the uncertainty of the model. Formally we can write these conditions as:

$$l_j \to l_{mi} \quad \text{if } \left\{e_{1j}, e_{2j}\right\} \subseteq S_i \wedge \left|\Delta\gamma_{ij}\right| \le \Delta\gamma_{\max} \quad (3)$$

$$\Delta\gamma_{ij} = \left|\gamma_i - \gamma_j\right|, \quad (4)$$

where l_{mi} denotes the i^{th} model line, S_i is the search space for the i^{th} model line (buffer around l_{mi}), l_j denotes the j^{th} image line, e_{1j} and e_{2j} are the end points of the j^{th} image line, γ_i is the γ-parameter of the representation given by (2) for the projection of the i^{th} model line and γ_j is the γ-parameter for the j^{th} image line.

This search for correspondences is applicable in our case because we assume to know the exterior orientation of the camera from the GPS/INS path precisely enough for the projected model lines to be only a few pixels away from their corresponding image features. To ensure a reliable assignment a visibility checking algorithm is also required, so that hidden edges are not projected into the image and no correspondences for them are assigned.

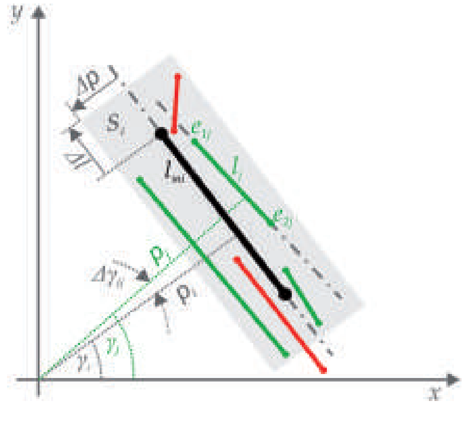


Fig. 2: Assignment of correspondences. l_{mi} is the i^{th} model line. For every l_{mi} (in black) a search space S_i (the buffer around l_{mi}) is defined. If both end points e_{1j} and e_{2j} of image line are within S_i and the angle difference $\Delta\gamma_{ij}$ between l_{mi} and l_j is smaller then a threshold, then the image line l_j is accepted as a correspondence for l_{mi} (in green). Otherwise l_j is rejected as a correspondence for l_{mi} (in red).

3.3 *Line Based Least Squares Adjustment*

The correspondence between 3D coordinates and their perspective projection into the image is given by the collinearity equations. The collinearity equations can be combined with (1) and the parameters m and d from the line representation

$$l: y = mx + d \tag{5}$$

can be calculated. Detailed equations needed to express m and d in terms of camera position parameters are given by MEIERHOLD et al. (2008). The authors also mention the problem of vertical image lines, which cannot be expressed using (4) and propose to change the line representation to

$$l: x = m'y + d', \text{ where } m' = \frac{1}{m},\ d' = \frac{d}{m}. \tag{6}$$

The problem during the adjustment is that some lines can change from being non-vertical to being vertical in the iteration process, so that the Jacobian matrix would have to be re-designed. To avoid this problem we use (2), express γ and p in terms of camera orientation parameters and use them as observations.

$$\gamma = \arctan\left(-\frac{1}{m}\right), \quad p = \frac{d}{\sqrt{m^2+1}} \tag{7}$$

For the adjustment we use the least squares method with the model:

$$\gamma_j + \hat{v}_{\gamma j} = f(\hat{X}_0, \hat{Y}_0, \hat{Z}_0, \hat{\omega}, \hat{\phi}, \hat{\kappa}, \alpha_i, \theta_i, X_{Si}, Y_{Si}), \tag{8}$$

$$p_j + \hat{v}_{pj} = f(\hat{X}_0, \hat{Y}_0, \hat{Z}_0, \hat{\omega}, \hat{\phi}, \hat{\kappa}, \alpha_i, \theta_i, X_{Si}, Y_{Si}, c, x_0, y_0). \tag{9}$$

In (8) and (9) the hat "^" is used for unknowns to be estimated. **f** denotes the functional model derived from the modified collinearity equations (7) based on the exterior orientation of the camera (X_0, Y_0, Z_0, ω, φ, κ), the interior orientation parameters (c, x_0, y_0) and the parameters (α_i, θ_i, X_{Si}, Y_{Si}) of the 3D lines. The interior orientation parameters and the parameters of the 3D lines are assumed to be constant. $\hat{v}_{\gamma j}$ and $\hat{v}_{pj}$ denote the corrections we seek to minimize in the least squares context:

$$\left(\sum \hat{v}_{\gamma j}{}^2 + \sum \hat{v}_{pj}{}^2\right) \rightarrow \min. \tag{10}$$

Writing the model in a vector form, where

$$\mathbf{b} = [\gamma_1, p_1, \gamma_2, p_2, \ldots\ldots, \gamma_n, p_n]^T, \tag{11}$$

$$\hat{\mathbf{v}} = [\hat{v}_{\gamma 1}, \hat{v}_{p1}, \hat{v}_{\gamma 2}, \hat{v}_{p2}, \ldots\ldots, \hat{v}_{\gamma n}, \hat{v}_{pn}]^T, \tag{12}$$

$$\hat{\mathbf{x}} = [\hat{X}_0, \hat{Y}_0, \hat{Z}_0, \hat{\omega}, \hat{\phi}, \hat{\kappa}] \tag{13}$$

we get:

$$\mathbf{b} + \hat{\mathbf{v}} = \mathbf{f}(\hat{\mathbf{x}}) = \mathbf{f}(\dot{\mathbf{x}} + \Delta\hat{\mathbf{x}}), \quad \hat{\mathbf{x}} = \dot{\mathbf{x}} + \Delta\hat{\mathbf{x}}. \tag{14}$$

Here **b** denotes the observation vector for n-correspondences, $\hat{\mathbf{x}}$ denotes the vector of estimated unknowns, and $\dot{\mathbf{x}}$ denotes the vector of approximated values for unknowns. Applying a first order approximation using a Taylor series the Jacobian matrix **A** is calculated as

$$\mathbf{A} = \frac{\partial \mathbf{f}(\dot{\mathbf{x}})}{\partial \dot{\mathbf{x}}}. \tag{15}$$

Then $\Delta\hat{\mathbf{x}}$ and $\hat{\mathbf{v}}$ are estimated using

$$\Delta\hat{\mathbf{x}} = \left(\mathbf{A}^T \mathbf{P}_{bb} \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{P}_{bb} \left(\mathbf{b} - \mathbf{f}(\dot{\mathbf{x}})\right), \tag{16}$$

$$\hat{\mathbf{v}} = \mathbf{A}\Delta\hat{\mathbf{x}} - \left(\mathbf{b} - \mathbf{f}(\dot{\mathbf{x}})\right). \tag{17}$$

where $\mathbf{P}_{bb}$ is the weighting matrix and

$$\mathbf{P}_{bb} = \mathbf{Q}_{bb}{}^{-1} = \sigma_0^2 \cdot \mathbf{C}_{bb}{}^{-1}. \tag{18}$$

$\mathbf{Q}_{bb}$ denotes the weight coefficient matrix, $\mathbf{C}_{bb}$ denotes the covariance matrix and σ_0^2 denotes the variance factor.

3.4 *Uncertainty of the Building Model*

The uncertainty of the building model is related to the inaccuracies of creation and generalization. Many building models are created from aerial imagery, where the roof corners and the height to the ground are measured. Hence often the roof overlap is not modelled, and the wall edges are less accurate (less reliable) than the roof edges. We assume different accuracy values for the roof points and for the wall/ground points in the model, which are presented graphically as ellipses in Fig. 3. The *Z*-coordinate is assumed to be less accurate than *X*- and *Y*-coordinates. In Fig. 3 it can be observed that in the case of oblique airborne images not only the *X*- and *Y*-components, but also the *Z*-component of the uncertainty has a strong influence on the position error of the projected point.

The model uncertainty can be embedded into the model presented in section 3.3 extending the vector of unknowns with the model line parameters:

$$\hat{\mathbf{x}} = [\hat{X}_0, \hat{Y}_0, \hat{Z}_0, \hat{\omega}, \hat{\phi}, \hat{\kappa}, \hat{\alpha}_i, \hat{\theta}_i, \hat{X}_{Si}, \hat{Y}_{Si}]\ . \tag{19}$$

Also the functional model **f** (14) has to be extended with equations for the model line parameters as follows:

$$\begin{aligned} \alpha_i + \hat{v}_{\alpha i} &= \hat{\alpha}_i \\ \theta_i + \hat{v}_{\theta i} &= \hat{\theta}_i \\ X_{Si} + \hat{v}_{Xsi} &= \hat{X}_{Si} \\ Y_{Si} + \hat{v}_{Ysi} &= \hat{Y}_{Si} \end{aligned} \tag{20}$$

Accordingly the Jacobian matrix **A** is extended with *4n* observations and *4n* unknowns. Unfortunately the uncertainties of the parameters of the 3D line models (α_i, θ_i, X_{Si}, Y_{Si}) are not directly known. Usually the position accuracy of the 3D building models created from aerial imagery is given for building corners. Therefore the uncertainty in the parameters of the 3D line models (α_i, θ_i, X_{Si}, Y_{Si}) has to be calculated using error propagation law:

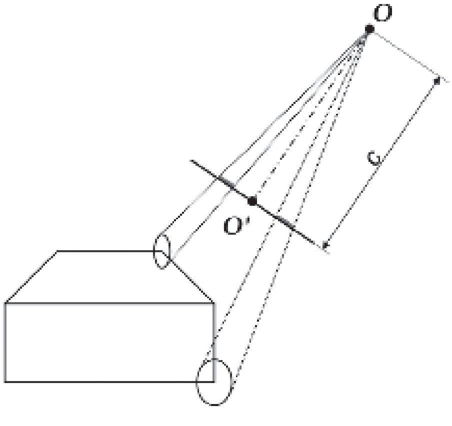


Fig. 3: Projection of model point uncertainty into the image in oblique geometry.

$$\mathbf{C}_{yy} = \mathbf{F}\mathbf{C}_{xx}\mathbf{F}^T\ . \tag{21}$$

$\mathbf{C}_{xx}$ denotes the covariance matrix for 3D coordinates of the corners, **F** is the Jacobian for the function transforming the *XYZ* coordinates to line parameters α, θ, X_S, Y_S, and $\mathbf{C}_{yy}$ denotes the covariance matrix for the line parameters. The uncertainty of X_S and Y_S depends on the coordinate system. Therefore all calculations should be carried out in the local coordinate system.

3.5 *Reliability and Uncertainty of Image Lines*

Edge extraction is carried out using the Canny algorithm. The edges are detected with different extraction parameters. By varying the minimum edge strength required for a feature to be accepted as an edge, different results are achieved. Lowering this parameter results in multiple detections, because low-contrast edges are also included. However, these edges are less reliable as building edges. Setting a high minimum edge strength means "stronger" edges can be detected, but there may not be a sufficient amount to fit the model. Thus we propose edge detection with varying minimum edge strength. As a result we get three sets of detected edges. We then combine all three edge sets using the minimum strength to weight them.

Additionally, we assume that building edges correspond to long image lines. Short lines more likely correspond to other objects or to noise. Therefore we also use the length of the detected lines for weighting. The weights of the lines are calculated as follows:

$$g_j = \frac{1}{2}\left(\frac{l_j}{d_{\max}} + \frac{a_j}{255}\right) \tag{22}$$

where g_j denotes the weight for the j^{th} image line, l_j denotes the length of the j^{th} line, $a_j \in [0,255]$ denotes the threshold for the minimum edge strength used for the extraction of the j^{th} line and d_{max} is the length of the diagonal

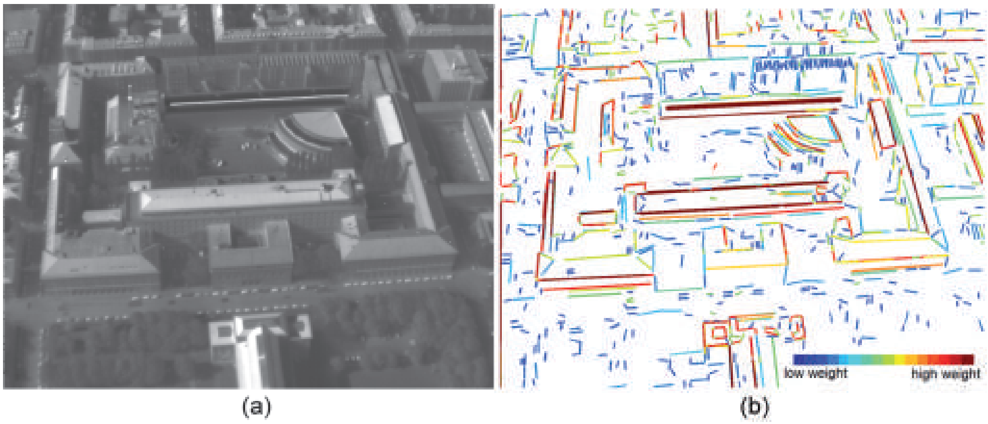


Fig. 4: An exemplary IR frame (a) and the weighting of the lines extracted in this frame (b). Three different minimum edge strengths were used for the extraction.

of the entire image. An exemplary weighting of the lines is presented in Fig 4.

We use weighting to identify the reliable edges and do not consider those with low weights for the assignment of correspondences.

For the image features the position uncertainty of the end points is given. Thus, we calculate the $\mathbf{P}_{bb}$ matrix for γ and p using error propagation as shown in (21).

4 Elimination of Incorrect Correspondences Using RANSAC

The method of assigning correspondences described in section 3.2 allows for the selection of multiple image lines corresponding to one model line. This leads to many incorrect correspondences, which have to be eliminated or reduced. For this purpose we apply the random sample consensus – RANSAC (Fischler & Bolles 1981). From the set of all hypothetical correspondences selected in the procedure from section 3.2, we randomly select three correspondences from different parts of the model and calculate exterior orientation parameters without redundancy. We then check how many of the remaining correspondences fit the randomly estimated exterior orientation parameters. This procedure is repeated k-times, and k is calculated as

$$k = \frac{\log(1-0.99)}{\log(1-(1-\varepsilon)^r)}, \tag{23}$$

where r is the number of necessary observations, $\varepsilon \in (0,1)$ is the outliers rate and the probability that RANSAC makes at least one error free selection is 99 %. We estimate ε as

$$\varepsilon = N - N_{Mod}, \tag{24}$$

where N is the number of hypothetical correspondences selected by the assignment algorithm and N_{Mod} is the number of model lines which have at least one assigned image line. The RANSAC algorithm results in new exterior orientation parameters and a set of correct correspondences. These data are taken as the input for the adjustment procedure described in section 3.3.

5 Data Description

For our experiments we used a test dataset captured in a densely built-up area in the centre of Munich, Germany. The thermal images were taken using the IR camera AIM 640 QLW FLIR with a frame rate of 25 images per second. The camera was mounted on a platform carried by a helicopter. The flying height was approximately 400 m above ground level. The camera was forward-looking with an oblique view of approximately 45°. The size of the images is 640 × 512 pixels. According to the flying height, camera orientation and camera parameters, the ground resolution of the IR images varies from about 0.5 m in the foreground to about 1.4 m in the background of the oblique images.

For geo-referencing we use data acquired by an Applanix POS AV 510 GPS/INS system. To correct the INS drift a Kalman filter is applied and an extended bundle adjustment is carried out (Kolecki et al. 2010). The corrected exterior orientation parameters are used for the model projection.

The 3D building models were created semi-automatically using commercial software for 3D building reconstruction from nadir view aerial images and stored in the *CityGML* format. In this format the buildings are stored face-wise as sequences of 3D points. Accordingly, all the lines are stored twice (once in each of the two faces intersecting at that line), and points even occur three or more times. We therefore reorganize the model to reduce double lines. We collect all points (corners) and remove repetitions. Then we store lines and faces as references to the points.

6 Results

We carried out tests with the data described in section 5. An exemplary frame with the image lines (cyan, blue and magenta) and the projected model (yellow and green) is presented in Fig. 5a. The model lines which have correspondences are printed in green and the model lines without correspondences in yellow. The image lines without correspondences are presented in cyan; image lines which were assigned to model lines, but refused by RANSAC are depicted in blue, and image lines which were accepted by RANSAC and taken as input for the adjustment are depicted in magenta. Fig. 5b shows the same frame with the projected model before (red) and after (green) adjustment.

In Fig. 5b an improvement of the position of the projected model can be observed. Projected building structures match the IR image very well. Thanks to this, thermal building textures can be extracted precisely. Nevertheless, accuracy of the exterior orientation parameters estimation is not very high. For the exemplary frame from Fig. 5 the standard de-

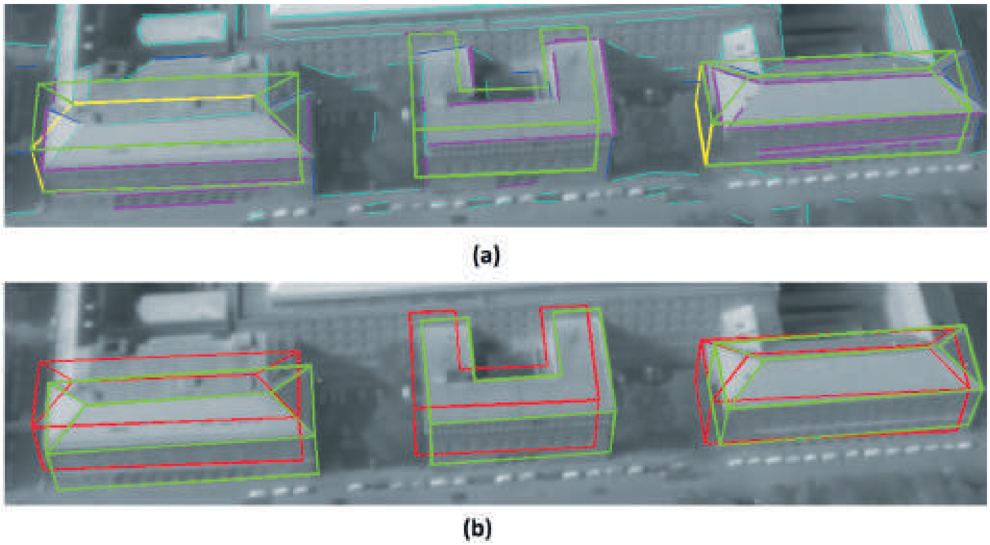


Fig. 5: Exemplary frame with extracted lines and a projected model: (a) before matching, where cyan = image lines without correspondences, blue = image lines with correspondences but refused by RANSAC, magenta = accepted image lines, green = model lines with correspondences, yellow = model lines without correspondences; and (b) after matching, where red = 3D building model projected with initial exterior orientation parameters, green = 3D building model projected with adjusted exterior orientation parameters.

viations are $\sigma_X = 3.8$ m, $\sigma_Y = 6.5$ m, $\sigma_Z = 4.5$ m, $\sigma_\omega = 0.60°$, $\sigma_\varphi = 0.69°$, and $\sigma_\kappa = 0.39°$. This is related to the low resolution of the images and the low accuracy of the extracted edges.

To evaluate the method and to investigate the sensitivity of the method with respect to changes in the initial exterior orientation we selected one frame and carried out a test in which we generated normally distributed random numbers with mean $\mu = 0$ and standard deviation $\sigma_{xyz} = 1$ m, $\sigma_{\omega\varphi\kappa} = 0.1°$ and used them to degrade the initial exterior orientation parameters. We successively repeated the test with increasing σ_{xyz} and $\sigma_{\omega\varphi\kappa}$. Results of this investigation are presented in the first row of Tab. 1. For comparison we conducted the same test without RANSAC (second row of Tab. 1).

Tab. 1: Percentage of successfully matched samples with downgraded initial exterior orientation. **σ** denotes the standard deviation used for the generation of normally distributed random numbers, with $\sigma = \sigma_{xyz}$, $\sigma_{\omega\varphi\kappa}$, where $\sigma_{xyz} = 1$ m, and $\sigma_{\omega\varphi\kappa} = 0.1°$. First row: results using RANSAC; second row: results based on the assignment of correspondences without outlier detection.

Successfully matched samples when downgrading the exterior orientation with normally distributed numbers using mean μ = 0 standard deviation (%)					
	σ	3**σ**	4**σ**	5**σ**	7**σ**
(1)	96	68	65	61	46
(2)	98	43	20	16	0

Tab. 1 shows that our method works for well geo-referenced images. The search for correspondences proposed in section 3.2 delivers good results if the camera position and rotations are known with an accuracy of 1–2 m and 0.1–0.2°, respectively. As we mentioned in section 3.2, the size of the search space S_i and the angular threshold are calculated based on the expected displacement and rotation of the projected model lines, which are in turn calculated from the propagation of errors in the 3D building model and the exterior orientation parameters. If we downgrade the initial exterior orientation by 4σ, we get the width parameter of the buffer $\Delta p = 3 \cdot \sigma_p$ in range of about 40 pixels and the angular threshold $\Delta\gamma = 3 \cdot \sigma_\gamma$, which corresponds to about 8°. The values are so high, because σ_p and σ_γ are calculated as propagation of the uncertainty of the camera position and orientation and the uncertainty of the 3D building model. This leads to many incorrect correspondences with $\varepsilon \approx 85$ %. In case of downgrading with 7σ we get $\varepsilon \approx 93$ %. The typical least squares adjustment method cannot cope with such a large number of outliers. However, using RANSAC the algorithm is robust even if the camera is "shifted" from the initial position by few meters (e.g. Fig. 6). For 7σ we still get 46 % successfully matched samples.

7 Discussion and Future Work

Line based model-to-image matching has high potential for co-registration of buildings with oblique airborne images. Edges are the most representative features for building structures and can be easily detected in the image using standard image processing algorithms. Considering the uncertainty of image lines and of the building model, as proposed in this paper, a better fit between the building model and the image structures is achieved. However, esti-

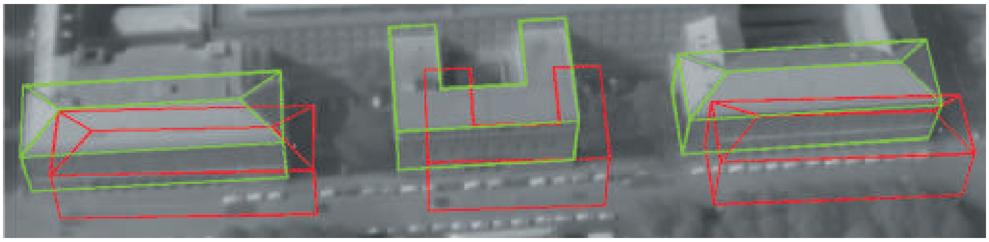


Fig. 6: Exemplary frame with the 3D building models projected with exterior orientation parameters downgraded by normally distributed numbers with μ and 7σ (red) and projected 3D building models after adjustment (green).

mated exterior orientation should not be understood as the true position of the camera, but only as parameters needed for a better initial projection in order to obtain a sufficient number of correct correspondences. In the future, neighbouring frames should also be employed in the adjustment to improve the relative orientation of the sequence. We expect that this will stabilize the model-to-image matching, in particular by reducing the "jumping" of the camera.

The uncertainty of the building models allows the calculation of the corrections for the model parameters. However, these corrections cannot be applied to improve the geometry of the model, because in the presented adjustment no constrains for planes are implemented. Therefore, some coplanar 3D building lines which originally belonged to one plane can be non-coplanar after adjustment. Nevertheless, the corrected position in the image can be used for texture extraction.

Applying the RANSAC algorithm we get good results not only for very well geo-referenced data but also for mid-quality geo-referencing. However, further improvements in the search for candidate correspondences are needed in case of low-quality input. Some already existing methods, e.g. vanishing points or relational matching, can be applied to find coarse exterior orientation parameters.

As an alternative or complement to RANSAC an outlier detector within the adjustment can be implemented using the cofactor matrix for the corrections Q_{vv}. However, only incorrect image lines can be efficiently detected. Errors in the model lines are difficult to recognize because their redundancy components are low, and therefore the influence of the errors on the correction for each parameter is very small.

In IR images depicting urban scenes it is often difficult to extract building edges on the ground. Often pavement appears similar to walls. Accordingly, in some frames a very good fit between the 3D building models and roof structure was achieved, while in the walls some remaining displacements occurred.

In the future more attention should be paid to the rolling shutter effect. This phenomenon should also be taken into consideration in the adjustment. Typically the IR cameras read out the measured radiation row-wise. If the camera is moving each row is read at a different point in time, and therefore at a different position, which should be accounted for in the geometrical model used for estimation.

Acknowledgements

The authors would like to thank FGAN-FOM, Ettlingen, for providing images of the flight campaign.

References

Avbelj, J., Iwaszczuk, D. & Stilla, U., 2010: Matching of 3D wire-frame building models with image features from infrared video sequences taken by helicopters. – International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences **38** (3B): 149–154.

Debevec, P.E., Taylor, C.J. & Malik, J., 1996: Modelling and rendering architecture from photographs: a hybrid geometry- and image-based approach. – **23rd** Annual Conference on Computer Graphics and Interactive Techniques: 11–20.

Ding, M. & Zakhor, A., 2008: Automatic registration of aerial imagery with untextured 3D LiDAR models. – IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR).

Eugster, H. & Nebiker, S., 2009: Real-time Georegistration of Video Streams from Mini or Micro UAS using Digital 3d City Models. – **6th** International Symposium on Mobile Mapping Technology, Presidente Prudente, Sao Paulo, Brazil.

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

Förstner, W., 2010: Optimal vanishing point detection and rotation estimation of single image from a legoland scene. – International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences **38** (3A): 157–162.

Früh, C., Sammon, R. & Zakhor, A., 2004: Automated Texture Mapping of 3D City Models With Oblique Aerial Imagery. – **2nd** International Symposium on 3D Data Processing, Visualization, and Transmission (DPVT **'04**).

Hoegner, L., Kumke, H., Meng, L. & Stilla, U., 2007: Automatic extraction of textures from infrared image sequences and database integration for 3D building models. – PFG **2007** (6): 459–468.

Hsu, S., Samarasekera, S., Kumar, R. & Sawhney, H.S., 2000: Pose Estimation, Model Refinement, and Enhanced Visualization Using Video. – Proceedings of CVPR **00 I:** 488–495.

Kolecki, J., Iwaszczuk, D. & Stilla, U., 2010: Calibration of an IR camera system for automatic texturing of 3D building models by direct georeferenced images. – EuroCOW 2010: on CD.

Lowe, D.G., 1991: Fitting Parameterized Three-Dimensional Models to Images. – IEEE Transactions on Pattern Analysis and Machine Intelligence **13** (5): 441–450.

Meierhold, N., Bienert, A. & Schmich, A., 2008: Linebased referencing between images and laser scanner data for image-based point cloud interpretation in a CAD environment. – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **37** (B5): 437–444.

Roberts, K.S., 1988: A new representation for lines. – IEEE Proceedings of Computer Vision and Pattern Recognition: 635–640.

Schenk, T., 2004: From point-based to feature-based aerial triangulation. – ISPRS Journal of Photogrammetry and Remote Sensing **58:** 315–329.

Sester, M. & Förstner, W., 1989: Object Location Based on Uncertain Models. – Burkhardt, H., Höhne, K.H. & Neumann, B. (eds.): Mustererkennung **1989** (11): 457–464, DAGM-Symposium, Springer-Verlag London, UK, ISBN 3-540-51748-0.

Tian, Y., Gerke, M., Vosselman, G. & Zhu, Q., 2008: Automatic edge matching across an image sequence based on reliable points. – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **37** (B3b): 657–662.

Vosselman, G., 1992: Relational matching. – Lecture Notes in Computer Science, Springer.

Vosselman, G., 1998: Interactive Alignment of Parameterised Object Models to Images. – International Archives of Photogrammetry and Remote Sensing **32** (3/1): 272–278.

Zhang, Y., Zhang, Z., Zhang, J. & Wu, J., 2005: 3D Building Modelling with Digital Map, LIDAR Data and Video Image Sequences. – Photogrammetric Record **20** (111): 285–302.

Address of the Authors:

Dorota Iwaszczuk, M.Sc., Dipl.-Inf. Ludwig Hoegner, Dipl.-Ing. Michael Schmitt & Univ.-Prof. Dr.-Ing. Uwe Stilla, Technische Universität München, Photogrammetry and Remote Sensing, D-80333 München, Tel.: +49-89-289-22637,-22680,-22672,-22671, Fax: +49-89-289-23202, e-mail: {dorota.iwaszczuk}{ludwig.hoegner}{michael.schmitt}{stilla}@bv.tum.de

Manuskript eingereicht: März 2012
Angenommen: Juni 2012

PFG 2012 / 5, 0523–0534
Stuttgart, October 2012

Article

A New Approach to Robust LiDAR/Optical Imagery Registration

Hui Ju, Charles Toth & Dorota A. Grejner-Brzezinska, Columbus, Ohio, USA

Keywords: LiDAR intensity image, aerial/satellite image, registration, LPFFT, MI, SIFT

Summary: Most of the image registration/matching methods are applicable to images acquired by either identical or similar sensors from various positions. Simpler techniques assume some object space relationship between sensor orientations, such as near parallel image planes, certain overlap, and comparable radiometric characteristics. More robust high-level feature-based methods allow for larger variations in image orientation and texture; for example, SIFT (scale invariant feature transform), a highly robust registration technique for wide baseline images. Nevertheless, registration between LiDAR (light detection and ranging) intensity and optical (satellite and aerial) images is still a big challenge, as substantial differences do exist in their radiometric characteristics. Reviewing and testing popular multiple domain image registration techniques, such as feature-based SIFT, intensity-based MI (mutual information), and frequency-based LPFFT (log-polar fast fourier transform), it is realized that no single technique could solve LiDAR intensity and optical image registration completely. Alternatively, a new approach to robust LiDAR/optical imagery registration, taking advantages of feature-, intensity-, and frequency-based methods, is proposed. Initial testing with a few datasets showed good performance of the new method, achieving pixel-level accuracy for the registration.

Zusammenfassung: Die meisten Methoden der Bildregistrierung/Bildanpassung können auf Bilder angewandt werden, die entweder mit identischen oder ähnlichen Sensoren aus verschiedenen Positionen aufgenommen werden. Einfachere Techniken basieren auf der Annahme eines speziellen Zusammenhangs im Objektraum, wie z.B. annähernd parallele Bildebenen, eine bestimmte Überlappung oder ähnliche radiometrischen Eigenschaften. Robustere high-level merkmalsbasierte Methoden ermöglichen größere Variationen in Bildorientierung und -textur, z.B. SIFT (Scale Invariant Feature Transform), eine sehr robuste Technik für die gegenseitige Registrierung von Bildpaaren mit langen Basislinien. Dennoch ist die Registrierung von LiDAR (Light Detection And Ranging) Intensitätsbildern und optischen (satelliten- oder flugzeuggestützten) Bildern noch immer eine große Herausforderung, da erhebliche Unterschiede in ihren radiometrischen Eigenschaften bestehen. Ein Test mehrerer beliebter Techniken zur Registrierung von Bildern aus unterschiedlichen Domänen wie z.B. das merkmalsbasierte Verfahren auf Basis von SIFT, das intensitätsbasierte Verfahren mit Hilfe von MI (Mutual Information) sowie das frequenzbasierte Verfahren LPFFT (Log-Polar Fast Fourier Transform), haben gezeigt, dass keine Technik einzeln das Problem der Registrierung von LiDAR Intensitätsbildern und optischen Bildern vollständig lösen kann. Als Alternative wird eine neue Bildregistrierungsmethode für LiDAR Intensitäts- und optische Bilder vorgeschlagen, welche die Vorteile der merkmalsbasierten, intensitätsbasierten und frequenzbasierten Methoden verbindet. Erste Tests des Verfahrens mit wenigen Datensätzen lieferten gute Ergebnisse mit Genauigkeiten in der Größenordnung von einem Pixel.

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0136
1432-8364/12/0136 $ 3.00

1 Introduction

1.1 *Motivation*

Image registration is a core task for various applications in digital photogrammetry, computer vision, remote sensing, and vision-aided navigation. Its purpose is to estimate the geometric transformation using an adequate number of correspondences between images acquired at different times, perspectives or even from different sensors. Image matching methods, computing those correspondences, are typically applicable to images acquired by either identical or similar sensors from various positions, and in the past few years, much effort has been devoted to develop automatic tie point extraction methods (Remondino & Ressl 2006).

As a growing number of various image sensors provide multiple image coverage worldwide, the need for registering imagery acquired from different airborne and spaceborne platforms is growing. Several satellite systems deliver high resolution imagery in short repeat time, large-format aerial digital cameras provide multispectral imagery at unprecedented resolution, LiDAR (light detection and ranging) systems collect both range and intensity images at local scale, while IfSAR (interferometric synthetic aperture radar) data are acquired from spaceborne and airborne platform at global scale, etc. All those data should be accurately registered for data fusion to support better geospatial data and information extraction.

The motivation for this study comes from three applications: terrain-based navigation, improving the geo-referencing of satellite imagery by using ground control, and developing a new man-made object modelling methods via fusing LiDAR and aerial images. In all those applications, LiDAR intensity and optical image registration plays an important role. It should be noted that data collected only from airborne/spaceborne platforms is considered in this study.

Registering LiDAR intensity and optical images is a particularly difficult task due to their substantially different characteristics, such as different sensing methodology (e.g. wavelength, passive/active image acquisition), geometric and radiometric differences, etc. In a LiDAR system, electromagnetic pulses in the visible and/or infrared bands are emitted from a transmitter, and besides the range measurement, the strength of the reflected pulse is recorded, which is known as the intensity value. A LiDAR intensity image is typically generated by rasterizing the intensity values of the point cloud. With increasing laser point density, e.g. 8–15 points per m^2, it is possible to obtain high-resolution LiDAR intensity images; nevertheless, they are still poor in quality in comparison with optical images. The main reason is the problem of rasterizing the irregularly distributed point cloud. After several rasterization tests, 1 m GSD is selected based on our data.

1.2 *Review of Multiple-Domain Image Registration Methods*

Multiple-domain image registration, also known as multi-modal image registration, has been investigated for decades, and, in general, can be classified into three major categories: feature-based, intensity-based, and frequency-based.

Feature-based registration methods use the similarity between features from the image pair to determine the transformation parameters. Low-level feature-based techniques use low-level features, such as points, corners and edges extracted from images. Unfortunately, the identification of conjugate corners or edges is difficult in the LiDAR intensity and optical image pair due to the irregular and sparse nature of LiDAR points at break lines. High-level feature-based techniques use high-level features such as regional descriptors and shape descriptors. SIFT (Lowe 1999, 2004) can be regarded as a complex descriptor, which could provide good registration results between aerial and satellite images, but failed in LiDAR intensity and optical image domains (Ju et al. 2011, Toth et al. 2010). The reason is that the substantial differences between the LiDAR intensity and optical image make the key points quite different in the two domains. Even for those key points extracted from similar locations, their descriptors can be still quite different, leading to mismatches. In contrast, fusing LiDAR and optical imagery for modelling of building facades is different. More importantly, the terrestrial laser scanner is much closer to the building, which can provide much denser and nearly regularly distributed points on the facades. Consequently, features, like corners of windows and doors, are easy to identify on the facades. Therefore, SIFT works rather well for terrestrial laser scanner and optical camera data (Böhm & Becker 2007, Becker & Haala 2008). For the airborne data, alternatively, other primitives, such as 3D straight lines and surface patches extracted from LiDAR data, are generally considered to be used to fuse optical images (Habib et al. 2004, Kim & Habib 2009). Note that in those approaches, LiDAR intensity is hardly considered.

Intensity-based registration methods usually define an intensity-based similarity measure between the templates (reference/target) or images, and then perform an optimization over allowed transformations to maximize this measure. For example, LSTM (least square template matching) is used to estimate the template-to-template transformation, which is normally an affine model (Grün 1985). Once enough correspondences are found via LSTM, the transformation between the image pair can be determined, and thus, image registration is achieved. For the multiple domain image registration, the method based on MI (mutual information) (Viola & Wells 1997) is one of the most popular ones, widely used in medical imaging applications and proved to be very effective. The basic concept of MI-based image registration comes from the information theory. Each image is regarded as a 2D discrete signal, carrying information. If two images are matched, their mutual information should be large and their joint entropy should be small. The transformation parameters are solved via maximizing the MI value based target function, in which different kinds of constraints to describe feature characteristics or spatial information can be introduced to improve the registration results. Although a few studies applying MI-based registration methods on multiple domain imagery in the fields of photogrammetry and remote sensing are reported, such as registration between TerraSAR-X and IKONOS images (Suri & Reinartz 2010) as well as between terrestrial camera image and infrared image (Liu et al. 2010), they have not been applied to LiDAR intensity and optical images. According to our limited datasets, NMI (normalized mutual information) can be used to find the transformation parameters, in the latter case; preliminary test results are discussed in the section 2.2.

Frequency-based registration methods use characteristics, such as the phase in the spectral domain to determine the transformation parameters between two images. Note that these techniques are typically restricted to handle images with limited surface-induced distortions. A popular frequency-based method is LPFFT (log-polar fast fourier transform) which estimates the shift or similarity transformation between image pairs without any feature detection (Reddy & Chatterji 1996, Wolberg & Zokai 2000, Zokai & Wolberg 2005). Based on our experiences, it is difficult to achieve reliable results applying the traditional LPFFT to our data.

Experimental results of using SIFT, MI and LPFFT on LiDAR intensity and optical images are given in section 2.

1.3 *Proposed Method*

The main contribution of this paper is to propose a hybrid multiple domain image registration method using a coarse-to-fine strategy, which largely refines our previous approach (Toth et al. 2011). First, a modified LPFFT with an internal validation module is used to estimate the coarse similarity transformation between LiDAR intensity and optical image pair. Next, strong HCs (Harris corners) in both images are generated and transformed to the other image via the estimated coarse transformation, and, subsequently, scale- and rotation-invariant PDF (probability density function) mean-shift matching (Comaniciu et al. 2003) is performed to find the correct correspondences. Finally, the RANSAC (random sample consensus) (Fischler & Bolles 1981) scheme is used to remove outliers and estimate the parameters of an affine transformation.

1.4 *Data*

To support this study, two datasets were used. The 1 m GSD orthorectified satellite images by GeoEye, acquired in January 2010, and 1 m GSD intensity images from airborne LiDAR data by Fugro-EarthData from 2009 covering the San Diego, California, USA area, represent a typical mix of terrain topography and landscape, including residential areas, roads, and vegetated areas. The 0.2 m GSD high resolution DMC aerial imagery and 1 m GSD intensity image from LiDAR data by ODOT (the Ohio Department of Transportation) cover the corridor area of highway I-70 in the Belmont County and highway 161 in Franklin County, Ohio, USA.

2 Experiences with SIFT, MI and LPFFT

In this section, test results of applying SIFT, MI and LPFFT to LiDAR intensity and optical image registration are discussed. These experiences directly inspired us to seek an alternative approach.

2.1 *SIFT*

In our early study, the baseline SIFT implementation was used for multiple domain image registration. It was found that SIFT can provide reliable matching results between satellite and aerial images. However, based on our limited data, SIFT matching between LiDAR intensity and optical images is not reliable. Fig. 1 shows typical SIFT matching results for a LiDAR intensity and satellite image pair. The number of the matched features is pretty small, and, more importantly, none of the matches is correct for this image pair, though there are keypoints extracted at similar locations. The main reason for the failure of SIFT matching is that it is very difficult to find similar keypoints in both images due to the substantial differences in radiometric characteristics and, in some extent, spatial resolution.

2.2 *MI*

To assess the MI-based registration performance on LiDAR intensity and optical image pairs, only simplified test scenarios are used.

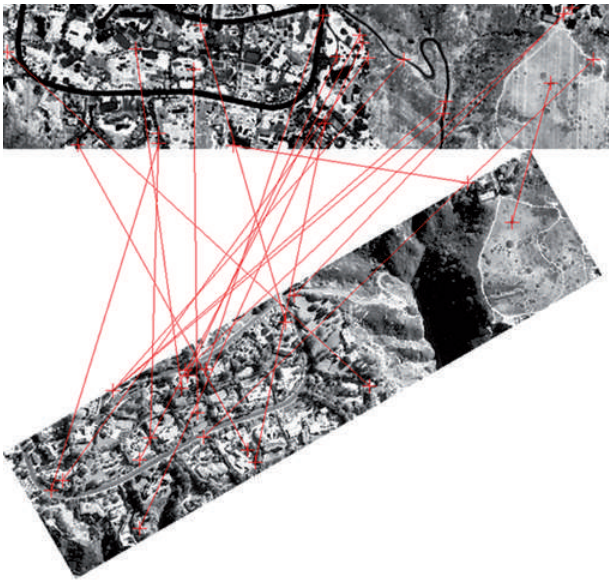


Fig. 1: SIFT matching results between LiDAR intensity image (top) and satellite image (bottom).

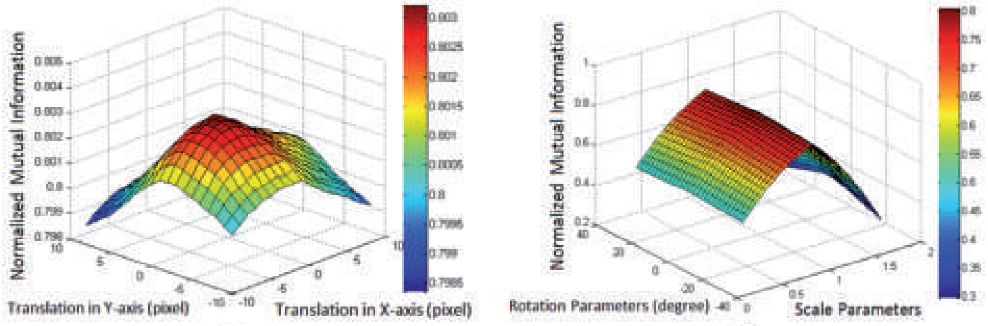
Also, the LiDAR intensity image is manually aligned with the optical image. Then, different scale, rotation and translation parameters are computed through expanding the correct parameters (s_0, φ_0, t_{x0}, t_{y0}) by a combination of incremental values (Δs, $\Delta\varphi$ and Δt) via (1). Those parameters are applied to the optical image, and then, the NMI (normalized mutual information) is computed for each image pair. 3D surfaces extended by (s, φ, NMI) and (t_x, t_y, NMI) are used to visualize the performance. The NMI surface should peak at the location of the correct parameters (s_0, φ_0) and (t_{x0}, t_{y0}), respectively. For our tests, $\Delta s = 0.05$, $\Delta\varphi = 1°$ and $\Delta t = 1$ pixel. Fig. 2 illustrates the typical NMI surface w.r.t. translation parameters (a) and w.r.t. scale and rotation parameters (b). Clearly, NMI can be used to locate the correct parameters in a given searching space. Fig. 2b shows that NMI is more sensitive to scale than to rotation angle, though, the correct scale and rotation parameters can be identified in the search space, see Fig. 3.

$$\begin{aligned} &s{:}= \{s|s_i = s_0 \pm i\cdot\Delta s\} \\ &\varphi{:}=\{\varphi|\varphi_i = \varphi_0 \pm i\cdot\Delta\varphi\} \qquad (1) \\ &t_x{:}= \{t_x|t_{x0} \pm i\cdot\Delta t\},\ t_y{:}= \{t_y|t_{y0} \pm i\cdot\Delta t\} \end{aligned}$$

NMI is a reliable indicator to locate the correct transformation parameters in a given (limited) search space; however, the determination of the right search space can be difficult.

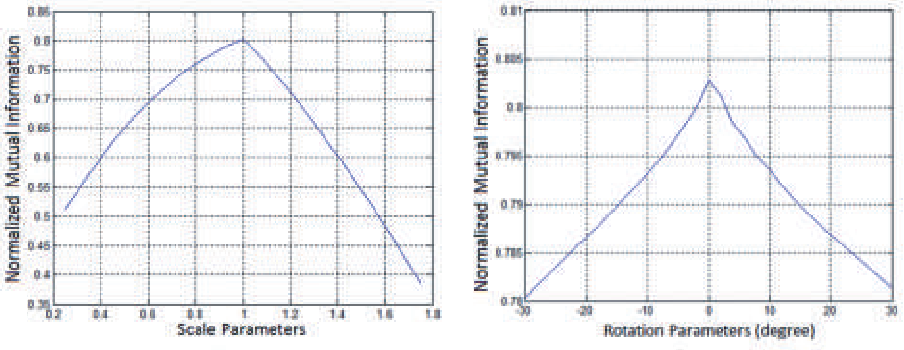
2.3 *LPFFT*

Using the LPFFT registration method, scale and rotation parameters of the similarity model are estimated in the first step. Next, the sec-



(a) (b)

Fig. 2: (a) NMI surface expanded by translation parameters and (b) NMI surface expanded by rotation and scale parameters.



(a) (b)

Fig. 3: (a) plot of NMI versus scale parameters and (b) plot of NMI versus rotation parameters.

ond image is transformed based on the estimated scale and rotation values, so the images have comparable orientation and scale. Finally, the translation parameters are estimated by NCC (normalized cross correlation). For efficient processing, FFT-accelerated NCC (FFT NCC) is used. In our testing, the cross phase correlation, i.e. the response to iFFT (inverse fast fourier transform) of the phase difference turned out to be very noisy. Simply considering the maximum response, which would indicate the correct scale and rotation parameters, is just not reliable. Therefore, a validation of the scale and rotation based on a Monte Carlo test is proposed, and provided good results.

According to our tests, using FFT NCC to compute the translation parameters is generally not reliable. Therefore, to estimate the translation parameters, a different approach was proposed. First, both images are converted into binary edge images; and then, a number of reference patches are automatically generated in the reference image. Next, the reference patches are matched in the second image based on template NCC matching. Ideally, image coordinate differences between all reference and matched patches should be identical, representing the common shift between the image pair. Our results confirmed the feasibility of the proposed procedure.

3 Methodology

Based on our experiences, the registration of the LiDAR intensity and optical image pair by the tested registration methods alone is just not reliable. Therefore, the coarse-to-fine hybrid multiple domain image registration approach is proposed; the workflow is illustrated in the Fig. 4. In the following subsections, the main components will be discussed in detail.

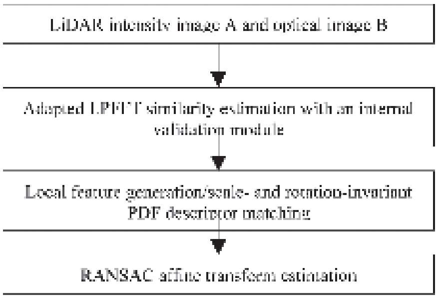


Fig. 4: Proposed multiple domain image registration workflow.

3.1 *Similarity Transformation Estimation*

The workflow of the adapted LPFFT similarity estimation approach is given in the Fig. 5. First, using the LPFFT, two parameters (scale and rotation) of the similarity model are estimated. LPFFT could provide a number of possible scale and rotation parameters from which the correct ones have to be identified through a validation. The validation of the scale and rotation parameters is achieved via a Monte Carlo test; more specifically, a Monte Carlo test is performed for a set of scale and rotation values computed using (1), where the two originally estimated parameters (s_0, φ_0) are perturbed by a combination of increment values (Δs, $\Delta\varphi$). Next, the second image is transformed using each scale and rotation combination in the set. If the scale and rotation parameters are correct, the image pair should have comparable orientation and scale. FFT-accelerated NCC, an efficient NCC computation method, is used to estimate the translation parameters for each image pair based on the maximum NCC values. For correct scale and rotation parameters, the maximum NCC values of all image pairs should fall in a significantly high range, which means that small scale and rotation changes around the correct scale and rotation still lead to a high NCC value. If the estimated scale and rotation are wrong, the maximum NCC values of all image pairs should be small. Fig. 6 shows a typical NCC surface based on wrong and correct parameters, (s_0, φ_0), respectively.

Next, the second image, B, is transformed using the estimated (correct) scale and rotation parameters, at which translation difference may still exist between the two images. To estimate the translation parameters, both images are converted into binary edge images. Then, a number of reference patches are automatically generated in the reference image, and finally, those reference patches are matched in the second image based on the template NCC

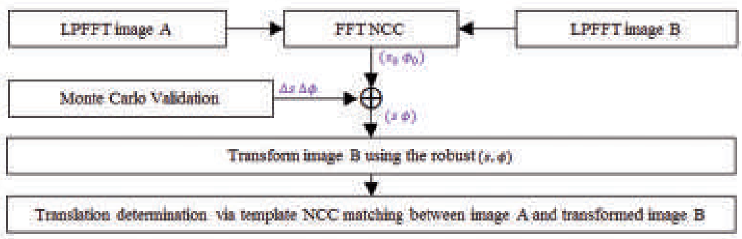
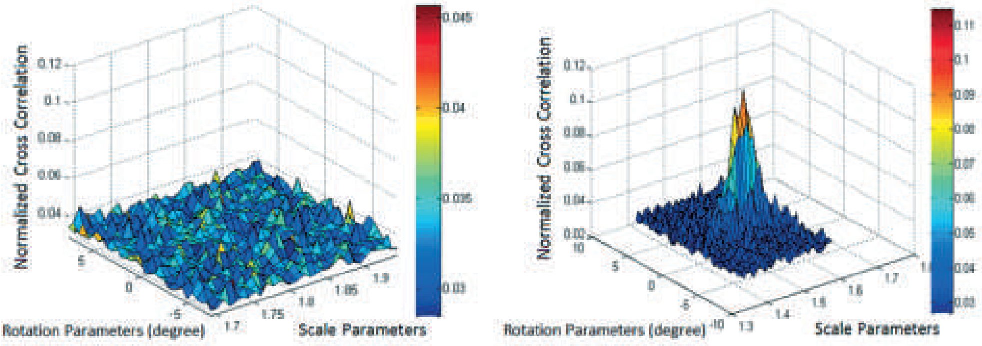
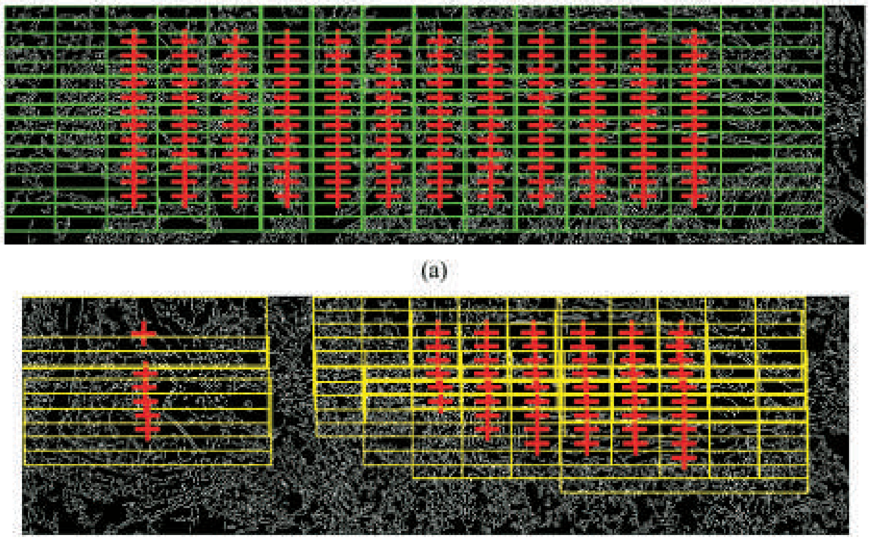


Fig. 5: Proposed similarity transform estimation method.



(a) (b)

Fig. 6: (a) NCC value surface centred at wrong scale and rotation and (b) at correct scale and rotation.



(a)

(b)

Fig. 7: (a) Reference patches in the LiDAR intensity edge image and (b) matched patches in the satellite edge image.

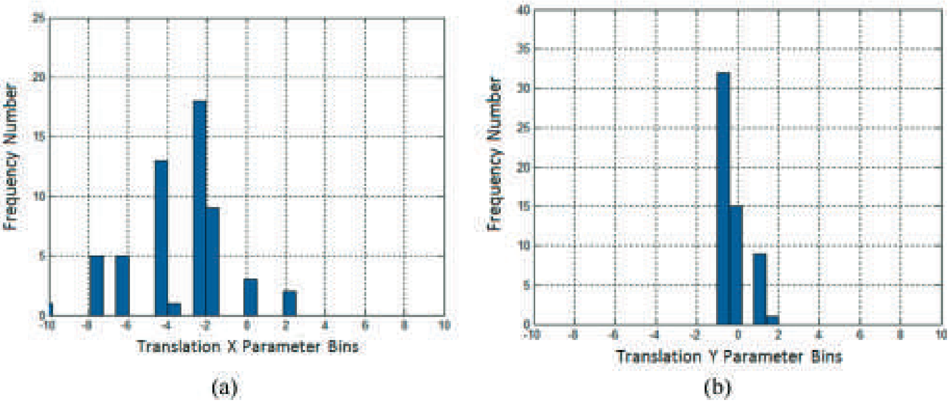


Fig. 8: (a) Histogram of column difference, translation in x and (b) row difference, translation in y.

matching. Fig. 7a shows the reference patches in the LiDAR intensity image A. Fig. 7b represents the matched templates in the second image. The template size is empirically determined, e.g. 0.3 times the image height and width. Ideally, image coordinate differences between all reference and matched patches should be identical. In reality, mismatches cannot be ruled out, and thus, the correct translation parameters are determined based on a statistical analysis of all column and row differences. For example, as shown in Fig. 8, most x-translations, t_x, fall between -3 and 0 pixels, and their accumulated total count over the interval is 28 out of 57. Similarly, the majority of y-translations t_y are between -1 and 0 pixels, totalling 47 out of 57. Thus, the average values t_x= -2.5, t_y= -1 pixels are accepted as the translation parameters.

3.2 *Scale-, Rotation-invariant Regional PDF Descriptor Matching*

The proposed feature generation and matching approach is illustrated in the Fig. 9. First, the HC detector is used to extract local feature points. As HCs in the two images are different, HCs from one image are transformed to the other image via the estimated similarity model from the adapted LPFFT. Square regions centred on those HCs are created in both images. Next, rotation-invariant kernel based PDF (probability density function) descriptor is created by applying a circular Epanechnikov kernel to the square region centred at the HC. The scale factor is known from the adapted LPFFT, hence the PDF descriptor is scale-invariant by adjusting the Epanechnikov

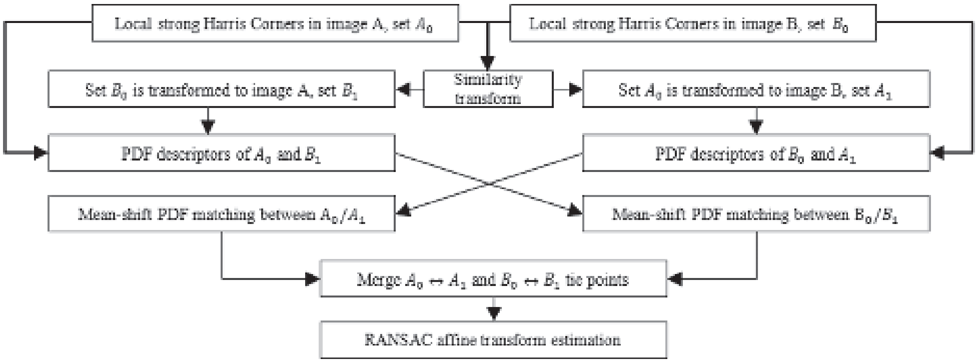


Fig. 9: Proposed affine transform estimation method.

kernel. The intensity PDF of the circular region is approximated by the normalized histogram; note for the 8-bit input image, the feature descriptor is 256-dimensional. Although the histogram is not the best nonparametric density estimation, it was proven to be sufficient for PDF matching purpose (COMANICIU et al. 2003). The similarity between two PDF descriptors is computed via the Bhattacharyya coefficient. Selecting PDF as feature descriptor is based on our earlier evaluation of multiple-domain image matching based on different feature spaces (JU et al. 2011).

If the similarity transformation is adequate (and properly estimated), the transformed feature locations should be close to the correct positions, and consequently, PDF mean-shift matching can fast reach the local maximum. PDF mean-shift matching is an efficient and robust object tracking method which can track objects under different illumination conditions and perspectives. In short, it maximizes the Bhattacharyya coefficient by finding the mode (peak) of the density in the local neighbourhood using mean-shift to recursively move to a new location (update). Feature region size can influence the PDF mean-shift matching performance and it is selected empirically. In our data, 50–110 pixels are appropriate. Fig. 10a shows an original feature in the LiDAR intensity image and Fig. 10b depicts the matched feature in the satellite image. Fig. 10c shows the PDF similarity score curve versus number of iterations; the local maximum is reached at the 4th iteration. Fig. 10d is the comparison between the reference PDF (blue) and the matched PDF (red).

All original features in the image pair are transformed to the other image; original features in the image A (set A_0) are transformed to the image B to form the set A_1. The same operation is applied to the original features in the image B. The PDF mean-shift matching is performed on features between A_0 and A_1 as

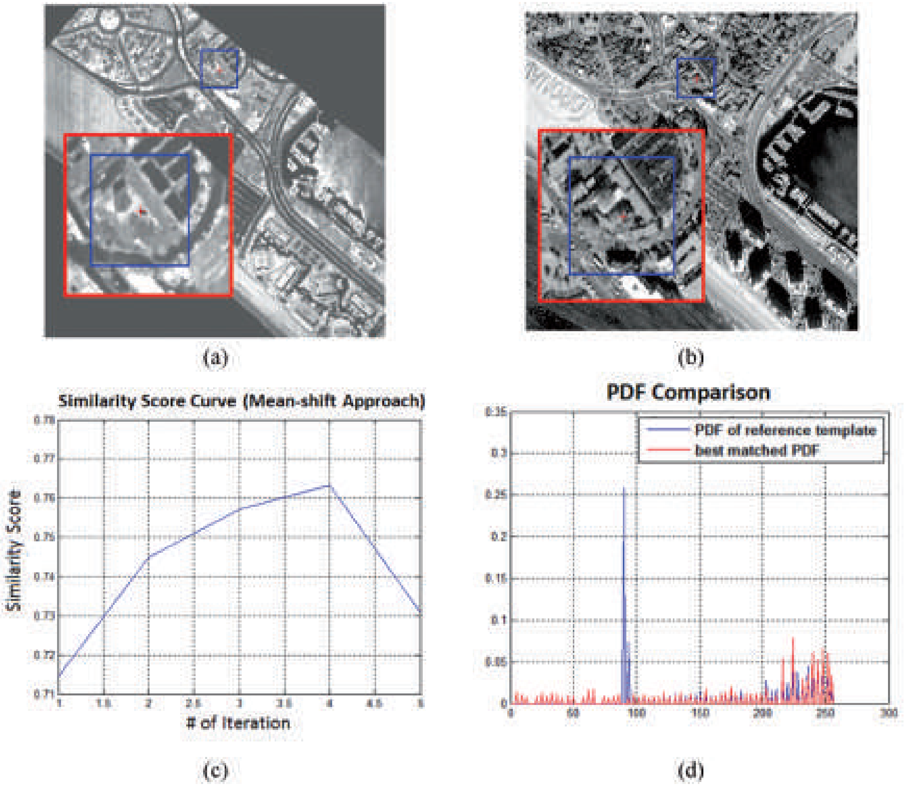


Fig. 10: (a) Feature location in the LiDAR intensity image, (b) matched feature location in the satellite image, (c) PDF similarity score curve and (d) comparison of reference PDF and matched PDF.

well as between B_0 and B_1 to find matched features with a similarity score larger than 0.7. Correspondences from $A_0 \leftrightarrow A_1$ and $B_0 \leftrightarrow B_1$ are merged as the complete set of tie points. Depending on object space characteristics, either affine and collinearity models, or more sophisticated models can be formed based on the matched feature locations. In all cases, blunder detection is necessary, which is based on RANSAC. In our test, RANSAC affine model estimation specifying a 0.5 pixel threshold for residual errors is used.

4 Experimental Results and Precision Analysis

The proposed approach was implemented in MATLAB and tested using the data introduced in section 1.4. Four aerial/LiDAR and satellite/LiDAR intensity image pairs were selected to evaluate the registration performance. The overlap is more than 90 % in the aerial/LiDAR and 100 % in the satellite/LiDAR image pairs. The extents of the overlap areas of the test image pairs are shown in Tab. 1. After RANSAC affine model estimation, the number of inliers is more than enough to determine the 6 parameters of the affine transformation in all tests. The RMSE (root-mean-square error) of position errors is used to judge the registration precision. Similarly to the re-projection error, the position error is computed as the position difference between the matched and transformed points in the optical image. The RMSE is computed on a pixel basis. As seen in Tab. 1, pixel level registration precision is obtained.

5 Conclusion

Feature-based SIFT registration, intensity-based MI registration, and frequency-based LPFFT registration methods were tested in this paper. Due to very different characteristics of LiDAR intensity and optical images, SIFT is unable to provide acceptable results based on our somewhat limited dataset. MI-based methods show good performance if the correct search range is given, which is a hard task. The traditional LPFFT has difficulty with finding the correct scale and rotation parameters from a set of candidates, and, in addition, the translation parameter determination is not reliable using FFT-accelerated NCC. Therefore, a hybrid method is proposed which is based on a two-step approach. The adapted LPFFT with a Monte Carlo validation check for the scale and rotation parameters, and estimating translation parameters based on the template NCC matching, is used to estimate an initial similarity transformation. Then, the scale- and rotation-invariant circular PDF descriptors centred at local strong HCs are created in each image, and then they are transformed to the other image via the estimated similarity transformation. The transformed feature location is the starting search position of a mean-shift PDF matching. In the final step, RANSAC affine model estimation is applied to the matched correspondences.

Tab. 1: Registration precision and performance and size of the test areas.

Aerial/LiDAR	**3–11**	**3–12**	**4–11**	**4–12**
Position RMSE (pixel)	0.95	1.16	0.99	1.13
Inlier/matched	16/30	11/25	8/16	15/26
Overlap size (m^2) Width(E) × Height(N)	463×811	465×804	473×823	458×818
Satellite/LiDAR	**06**	**07**	**08**	**09**
Position RMSE (pixel)	1.15	1.29	1.30	1.25
Inlier/matched	17/29	28/57	28/54	13/28
Overlap size (m^2) Width(E) × Height(N)	663×331	860×1426	723×970	326×575

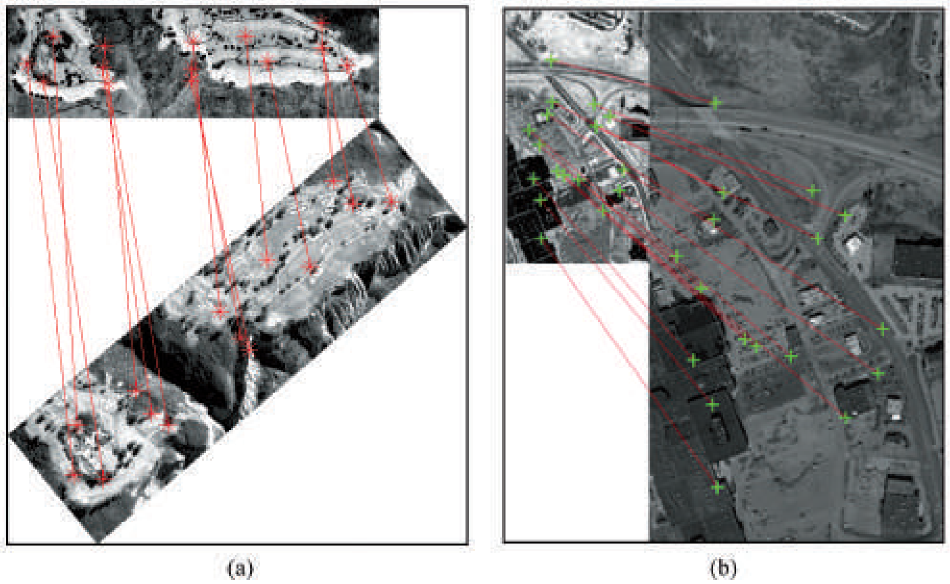


Fig. 11: (a) Registration between LiDAR intensity and satellite image pair, (b) beween LiDAR intensity and aerial image pair.

This registration method is applied to LiDAR intensity and optical images. The results on a few image pairs have shown good performance, as pixel level registration precision was obtained.

Acknowledgements

The authors thank the GeoEye Fundation, Fugro-EarthData and ODOT for the data provided for this research.

References

Becker, S. & Haala, N., 2008: Integrated LiDAR and Image Processing for the Modeling of Building Facades. – PFG **2008** (2): 65–81.

Böhm, J. & Becker, S., 2007: Automatic Marker-Free Registration of Terrestrial Laser Scans using Reflectance Features. – 8th Conference on Optical 3D Measurement Techniques: 338–344, Zürich, Switzerland.

Comaniciu, D., Ramesh, V. & Meer, P., 2003: Kernel-based Object Tracking. – IEEE Transactions on Pattern Analysis and Machine Intelligence **25** (5): 564–577.

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 ACM **24:** 381–395.

Grün, A.W., 1985: Adaptive Least Squares Correlation: A Powerful Image Matching Technique. – South African Journal of Photogrammetry, Remote Sensing and Cartography **14:** 175–187.

Habib, A.F., Ghanma, M.S. & Tait, M., 2004: Integration of LiDAR and Photogrammetry for Close Range Applications. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **35** (B5).

Ju, H., Toth, C.K. & Grejner-Brzezinska, D.A., 2011: Evaluation of Multiple-domain Imagery Matching based on Different Feature Spaces. – 2011 ASPRS Fall Conference, Herndon, VA, USA, CD-ROM.

Kim, C. & Habib, A.F., 2009: Object-based Integration of Photogrammetric and LiDAR Data for Automated Generation of Complex Polyhedral Building Models. – Sensors **9** (7): 5679–5701.

Liu, Z.Y., Zhou, F.G., Bai, X.Z., Wang, H. & Tan, D.J., 2010: Multi-modal Image Registration by Mutual Information based on Optimal Region Selection. – International Conference on Information, Network and Automation: 249–253, Kunming, China.

LOWE, D., 1999: Object Recognition from Local Scale Invariant Features. – International Conference on Computer Vision: 1150–1157, Corfu, Greece.

LOWE, D., 2004: Distinctive Image Feature from Scale-invariant Keypoints. – International Journal of Computer Vision **60** (2): 91–110.

REDDY, B.S. & CHATTERJI, B.N., 1996: An FFT-based Technique for Translation, Rotation and Scale-Invariant Image Registration. – IEEE Transactions on Image Processing **5** (8): 1266–1271.

REMONDINO, F. & RESSL, C., 2006: Overview and Experiences in Automated Markerless Image Orientation. – IAPRSSIS **36** (3): 248–254.

SURI, S. & REINARTZ, P., 2010: Mutual Information Based Registration of TerraSAR-X and IKONOS Imagery in Urban Areas. – IEEE Transactions on Geoscience and Remote Sensing **48** (2): 939–949.

TOTH, C.K., JU, H. & GREJNER-BRZEZINSKA, D.A., 2010: Experience with using SIFT for Multiple Image Domain Matching. – **2010** ASPRS Annual Conference, San Diego, CA, USA, CD-ROM.

TOTH, C.K., JU, H. & GREJNER-BRZEZINSKA, D.A., 2011: Matching between Different Image Domains. – PIA 2011, Lecture Notes in Computer Science **6952:** 37–47, Munich, Germany.

VIOLA, P. & WELLS, W.M., 1997: Alignment by Maximization of Mutual Information. – International Journal of Computer Vision **24** (3): 137–154.

WOLBERG, G. & ZOKAI, S., 2000: Robust Image Registration using Log-Polar Transform. – 2000 International Conference on Image Processing **1:** 493–496.

ZOKAI, S. & WOLBERG, G., 2005: Image Registration using Log-Polar Mapping for Recovery of Large-Scale Similarity and Projective Transform. – IEEE Transactions on Image Processing **14** (10): 1422–1434.

Address of the Authors:

Ph.D. candidate HUI JU, Dr. CHARLES TOTH, Prof. Dr. DOROTA A. GREJNER-BRZEZINSKA, The Ohio State University, Department of Civil, Environmental and Geodetic Engineering, 470 Hitchcock Hall, 2070 Neil Avenue, Columbus, OH 43213, USA, Tel.: +1-614-292-7681, Fax: +1-614-292-2957, e-mail: ju.32@osu.edu, toth@cfm.ohio-state.edu, dbrzezinska@osu.edu

Manuskript eingereicht: März 2012
Angenommen: Juni 2012

PFG 2012 / 5, 0535–0545
Stuttgart, October 2012

Article

Identifying Correspondences in Sparse and Varying 3D Point Clouds using Distinctive Features

Daniel Muhle, Hannover, Steffen Abraham, Hildesheim, Manfred Wiggenhagen & Christian Heipke, Hannover

Keywords: photogrammetry, matching, point cloud

Summary: In a wide range of applications stereo systems are used to extract geometric information from the scene observed with the stereo cameras. One possible solution to reconstruct the motion of such a system is to establish correspondences between points of the point clouds generated from stereo matching of image features at different epochs. There exists a large variety of approaches to establish correspondences between image or 3D data. A special group of algorithms, mostly inspired by the work of Lowe (2004), is based on the notion of distinctive feature descriptions. These algorithms assume the existence of a dense neighbourhood changing not too much over time. But the prevalence of untextured regions or computational constraints hindering the use of computationally expensive dense stereo matching approaches often result in only sparse point clouds and thus these approaches cannot be used for the registration of sparse 3D data. In our work we present a new approach that uses the basic principles of distinctive feature descriptions and extends them in a way that they can be applied to identify corresponding points between sparse 3D point clouds. Furthermore, an evaluation is given investigating the advantages and limitations of our approach. The results clearly show the effectiveness of the presented distinctive features to establish point matches between sparse 3D point clouds.

Zusammenfassung: In vielen unterschiedlichen Anwendungsbereichen werden Stereosysteme verwendet, um geometrische Informationen über die aufgenommene Szene zu extrahieren. Eine dabei anfallende Teilaufgabe ist das Identifizieren von Korrespondenzen zwischen Punkten einer 3D Punktwolke, die zu unterschiedlichen Zeitpunkten durch das Stereomatching von Bildmerkmalen entstanden ist. Inspiriert durch die Arbeit von Lowe (2004) sind für die Suche nach korrespondierenden Punkten eine ganze Reihe von Ansätzen entstanden, die auf charakteristischen Beschreibungen aufsetzen. Alle diese Verfahren setzen das Vorhandensein einer dicht besetzten Nachbarschaft voraus, die sich über die Zeit hinweg nicht zu stark ändert. Allerdings führen untexturierte Bereiche oder Echtzeitanforderungen, die den Einsatz von rechenintensiven dense-matching Ansätzen verbieten, zu dünn besetzten 3D Punktwolken, so dass die bekannten Verfahren nicht unmittelbar verwendet werden können. In unserer Arbeit wird ein neuartiger Ansatz vorgestellt, der auf den Grundprinzipien der charakteristischen Beschreibungen aufbaut und diese so erweitert, dass sie für die Punktzuordnung in dünn besetzten 3D Punktwolken geeignet sind. Darüber hinaus wird eine Untersuchung vorgestellt, die die Vorteile und Grenzen des entwickelten Ansatzes aufzeigt. Die Ergebnisse zeigen deutlich die Leistungsfähigkeit der entwickelten charakteristischen Beschreibung für die Zuordnung von dünn besetzten 3D Punktwolken.

---

1 Introduction

One task towards reconstructing the motion of a stereo system, e.g. used as a robot's eyes as it traverses through its environment, is to establish correspondences between points of the point clouds reconstructed from stereo matching of image features at different epochs. There exists a variety of approaches to establish correspondences between image or 3D data. A special group of algorithms, mostly inspired by the work of Lowe (2004), is based on the notion of distinctive feature descriptions. Typical examples of these approaches

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0137 1432-8364/12/0137 $ 3.00

are SIFT (Lowe 2004) and Spin Images (Johnson & Hebert 1999) for the registration of 2D and 3D data. Most of these approaches cannot be used directly for the registration of sparse 3D data, though, as they assume the existence of a dense neighbourhood that does not significantly change over time. Sparse and varying point clouds must be expected if such a system is used in a man-made environment with untextured regions on floors and walls or if real-time constraints hinder the use of computationally expensive dense stereo matching.

In our work we present an approach that uses the basic principles of distinctive feature descriptions and extends them in a way that they can be applied to identify point matches between sparse point clouds. The resulting distinctive feature vector is sparse and discrete and can be used to establish correspondences efficiently. After discussing the state-of-the-art in section 2, a detailed description of the proposed algorithm is given in section 3. The performance of the presented approach is tested on different sequences of a stereo system moving along a corridor. The established correspondences between points are used to reconstruct the motion of the system. Furthermore, in section 4 an evaluation is given investigating the overall performance and limitations of our approach. The results clearly show the effectiveness of the presented distinctive features for the matching between sparse 3D point clouds. Section 5 presents some ideas for future extensions.

2 Related Work

The point clouds that are generated at the epochs $t = i$ and $t = j$, while e.g. a stereo system mounted on a robot platform traverses down a hallway, are related by a rigid transformation ${}^{j}\mathbf{T}_{i}$, where the symbol ${}^{j}\mathbf{T}_{i}$ combines the rotation ${}^{j}\mathbf{R}_{i}$ and the translation ${}^{i}\mathbf{t}_{j}$. The rotation ${}^{j}\mathbf{R}_{i}$ rotates a point ${}^{i}\mathbf{x}$, defined in the coordinate system at $t = i$ into the coordinate system at $t = j$. The translation ${}^{i}\mathbf{t}_{j}$ between the epochs is defined w.r.t. to the coordinate system at $t = i$. Given the point clouds from two epochs one seeks to find the transformation ${}^{j}\mathbf{T}_{i}$ minimizing the error e given in (1) for all n corresponding points.

$$e = \sum_{k=1}^{n} \left\| {}^{i_k}\mathbf{x} - {}^{i}\mathbf{R}_{j}\, {}^{j_k}\mathbf{x} + {}^{i}\mathbf{t}_{j} \right\| \tag{1}$$

A prerequisite for the solution of this task is to identify point correspondences between the point clouds at $t = i$ and $t = j$. If the correspondences are known, established approaches like ICP (iterative closest point) (Besl & McKay 1992) and its variants (Rusinkiewicz & Levoy 2001) or least squares matching (Grün & Akca 2005) can be used to find the optimal solution for ${}^{j}\mathbf{T}_{i}$ that minimizes the error given in (1). As the focus of our work is an algorithm to establish point correspondences between sparse and varying point clouds we first give in section 2.1 a short overview about the generally applied matching workflow, and show in section 2.2 the general ideas behind the use of distinctive features for the task of identifying correspondence either in 2D or 3D data. In section 2.3 we give a detailed explanation of the contribution of our work. For the experiments presented in section 4 signalized targets are used that can be identified easily in the stereo image pairs to create a sparse 3D point cloud. Approaches that allow the derivation of a sparse point cloud from dense but irregularly sampled point clouds can be found e.g. in Stückler & Behnke (2011) and Novatnack & Nishino (2007).

2.1 *Common Matching Workflow*

Overviews of approaches for the matching of a large variety of input data can be found e.g. in Brown (1992), Seeger & Laboureux (2000), Zitová & Flusser (2003) or McGlone et al. (2004, chap. 6.3). In general, approaches used for the matching of a variety of input data, follow a similar scheme:

Defining the feature space

The feature space defines the input data used for matching. Typical examples for the 2D and 3D case are grey values, gradients, corners and 3D point clouds. A feature is an element taken from the set of the input data.

Defining the parameter space

The dimension of the parameter space is defined by the choice of transformation used for mapping of the input data. Typical transformations that are often used in the context of feature matching are homographies, affine and Euclidean transformations.

Establishing assignments using a similarity measure

Matched features are identified by defining a similarity measure that is computed either for all combinations or a reasonable subset of the used/extracted features. Two features can for instance be matched if a) their respective similarity is the largest and b) the similarity is discriminative, i.e. it is above and the ratio between the second-best and the best match is below a pre-defined threshold.

2.2 *Matching with Distinctive Features*

Recently, feature representations that are unique and distinctive are widely used in the area of photogrammetry and computer vision. For these representations two different terms are used in the literature: descriptor and signatures. The differentiation between these terms is not always clear and sometimes they are used ambiguously. In LOWE (2004) a descriptor is defined as a distinctive and compressed representation of the original input data. In contrast, CALONDER et al. (2008) define a signature as the result of a mapping $F : \mathfrak{R}^D \rightarrow \mathfrak{R}^d$ that transforms the input data $\mathbf{x}_k \in \mathfrak{R}^D, \forall k = 1 \ldots n$ with the dimension D into another space with the dimension d. To avoid any ambiguities, we will use the term distinctive (feature) description in the following. The most important properties of distinctive descriptions independent of their actual realization are:

- The distinctive description is invariant w.r.t. a variety of changes of the input data. Typical changes comprise e.g. illumination changes, translation, rotation and/or scaling.
- The computation of a similarity measure between any two descriptions can be done using simple distance metrics like the Euclidean distance.

Examples for distinctive description used to match two-dimensional image data are the SIFT and SURF descriptors (LOWE 2004, BAY et al. 2008). For the identification of corresponding points in point clouds derived from either range scanner like e.g. the Microsoft Kinect system or terrestrial laser scanner, FLINT et al. (2008), WANG & BRENNER (2008), LO & SIEBERT (2009), BARNEA & FILIN (2010) and WEINMANN et al. (2011) directly apply variants of the SIFT algorithm that consider the special characteristics of the available 3D data. All these approaches require that the irregular 3D data must be resampled to a regular two-dimensional grid. Furthermore, the point signatures (CHUA & JARVIS 1997), spin images (JOHNSON & HEBERT 1999), the approach of GELFAND et al. (2005) and the NARF (normal aligned radial feature) developed by STEDER et al. (2011) are more examples of algorithms that use distinctive feature descriptions to establish matches between points from different point clouds. All these approaches have in common, that they incorporate points or information from a densely sampled local neighbourhood to define a unique local frame of reference. The definition of the local reference frame is usually the first step to achieve invariance against rotation and translation of the input data. A well-designed computation of the distinctive description allows the usage of simple distance metrics for the matching step and makes it robust against other changes of the input data like scaling or change of illumination. Furthermore, the developed feature descriptions simplify the matching step and established approaches like clustering or binary space-partitioning trees can be used to accelerate the search for matching descriptions (e.g. WINKELBACH & WAHL 2008, NISTÉR & STEWENIUS 2006).

2.3 *Contribution of Our Work*

Most of the known approaches that rely on distinctive descriptions to find matching point pairs between point clouds or surface data require the existence of a densely sampled and regular neighbourhood that does not change too much over time. Typical examples for data fulfilling such a requirement are the results from dense stereo matching or data acquired with a laser scanner. For these types of input data it is relatively easy to establish a unique frame of reference to be robust against rotation and translation. For systems generating sparse point clouds with local neighbourhoods that change over time as new points become visible and other points move out of the field of view, the known approaches using distinctive descriptions for matching cannot be used directly. The algorithm presented in section 3 extends the existing approaches for matching 3D points and presents solutions to:

- achieve invariance against rotation and translation of the input data for sparse and varying point clouds,
- compute a distinctive description that is just as sparse and varying as the input data,
- efficiently compute a similarity measure for the sparse and varying distinctive description.

3 Matching between Sparse and Varying Point Clouds

The proposed scheme to find matching point pairs in sparse and changing 3D point clouds follows the workflow of the existing approaches that use distinctive descriptions for matching. The first step is the definition of a local frame of reference to achieve invariance against rotation and translation of the point cloud. In the second step a sparse 2D distinctive description $\mathbf{D}$ is computed from selected points in a local neighbourhood. For the third step the sparseness of the 2D distinctive description is exploited to derive a compact 1D description that allows an efficient computation of similarity between two distinctive descriptions.

3.1 *Identification of Locally Planar Neighbourhoods*

The first step for the computation of the distinctive description for a point $\mathbf{x}$ in a sparse and varying point cloud is identical to the approaches mentioned in section 2.2: the estimation of a plane normal $\mathbf{n}$ from points in a local neighbourhood. To increase the probability that the normal does not change when neighbouring points disappear or new ones enter the camera's field of view only those points are selected for the computation of $\mathbf{n}$ that lie approximately on the same plane as $\mathbf{x}$. For the estimation of $\mathbf{n}$ and the identification of points lying in the same plane as $\mathbf{x}$ a brute-force approach using the RANSAC algorithm (Fischler & Bolles 1981) is applied. The functional model is the Hessian normal form given by (2), where d represents the distance of a point $\mathbf{x}$ to the plane with its normal $\mathbf{n}$.

$$\mathbf{n} \cdot \mathbf{x} = d \qquad (2)$$

The simple brute-force approach used here is sensible as the processed point clouds contain only few points lying mostly on the planar walls of a hallway. For these constraints the RANSAC algorithm needs only a small number of iterations to find a good solution. For the selection of the local neighbours we define the following parameters: a) maximum number n_{max} of neighbours considered, b) minimum number n_{min} of neighbours considered, and c) a maximum radius r_{max} used in the nearest neighbour search. As a result of this first step we have associated every point $\mathbf{x}$ of a point cloud with a locally planar neighbourhood, a normal vector $\mathbf{n}$ and a set $\mathbb{P} = \{x_1 \ldots x_n\}$ of its n neighbours. Those points with a non-planar neighbourhood are not considered further in the matching process.

3.2 *The 2D Distinctive Description*

The computation of the entries $\mathbf{d}_k$ for the two dimensional description $\mathbf{D}$ is done in three steps. First, for a neighbour $\mathbf{x}_i$ a local frame of reference is defined where the x-axis is given by the direction from $\mathbf{x}$ to $\mathbf{x}_i$ projected into the local plane and the z-axis is given by the local normal $\mathbf{n}$. The y-axis is computed from the cross product of the x- and z-axis. Second, as shown in Figs. 1(a) and 1(b) all neighbours $\mathbf{x}_j : \forall j = 1 \ldots n$ are projected into the x, y-plane of the local reference frame and transformed into two dimensional polar coordinates. Fig. 1(b) shows that for all neighbours $\mathbf{x}_j$ the radial distance d_j w.r.t. x and the angle θ_j w.r.t. the x-axis of the local reference frame are used as entries into **D**. Third, the steps one and two are repeated for all remaining neighbours. In contrast to e.g. the point signatures (Chua & Jarvis 1997), that use exactly one reference direction to define a local frame of reference, the approach presented here defines a local frame of reference for every neighbour. Such a strategy is advantageous if the structure of a local neighbourhood changes over time. If that is the case it is not advisable to select one of the neighbours as a reference direction, because that point might disappear from the stereo system's field of view and then a new reference direction must be selected and thus the distinctive description changes completely. The proposed computation of the entries $\mathbf{d}_k : \forall k = 1 \ldots n^2$ for the distinctive description **D** also fulfils the requirement of invariance w.r.t. rotation and translation of the point cloud. The computed entries (relative angles and distances) are not changed by a rotation around the z-axis and the use of a local frame of reference eliminates the influence of a translation. The distinctive description is not invariant w.r.t. scaling that changes the length of distances. Scale invariance can be achieved, however, when ratios of distances are used instead of distances.

3.3 *A Compact Distinctive Description*

It is clear from the algorithm in section 3.2 and Fig. 1(b) that the two dimensional distinctive description is still sparsely populated. This special structure of **D** allows, analogously to the idea given in Calonder et al. (2009), to design a more compact variant of the distinctive description. For a compact version of **D**, the non-discrete entries $\mathbf{d}_k : \forall k = 1 \ldots n^2, \mathbf{d}_k \in \mathbf{D}$ are mapped into an integer scalar $s_k : s_k \in \{1, 2, \ldots 2p\}$, where p is a parameter of the mapping and influences the discretisation. Finally, the s_k are pooled in an ordered result set $\mathbb{S}$ containing only unique values. As shown in section 3.4, the usage of integer values is advantageous as it allows an efficient implementation for the comparison of two distinctive descriptions. The s_k are computed by applying the mapping $\mathrm{T} : \mathbf{d}_k \rightarrow s_k$ to all entries $\mathbf{d}_k$ in **D**. For T the quad tree index (Finkel & Bentley 1974) is used that recursively divides the 2D space into discrete grids and gives an integer index for a two dimensional point. The only parameter of the quad tree index is p defining the number of quadrants used for the partitioning of the two dimensional space. This parameter controls the loss of accuracy caused by the discretization. For a chosen value of $p = 16$ the x-axis (radial distances d_j) and the y-axis (angle θ_j) will be partitioned into $2^{p/2} = 256$ bins. The value of

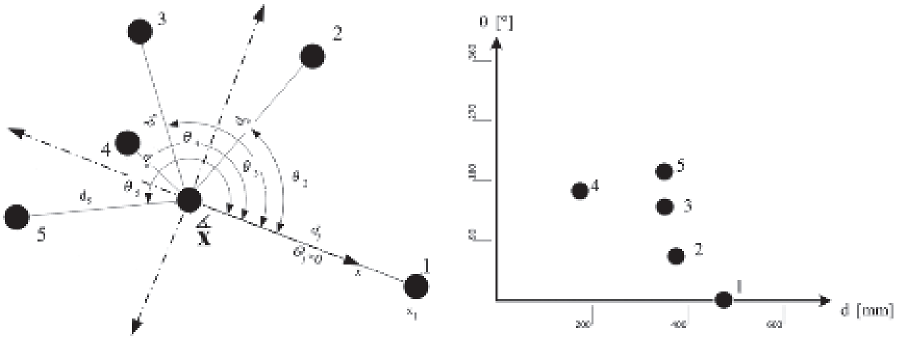


Fig. 1: Example for one iteration of the computation of the sparsely populated distinctive description **D**. (a) Definition of a local frame of reference. The *x*-axis is given by the direction to point 1; the *z*-axis points toward the reader. The *n* neighbours are transformed into the local frame and converted to polar coordinates (distance and direction). (b) Part of the descriptor **D** computed from the neighbourhood shown in (a). The entries into **D** are the polar coordinates of all neighbours for the local frame of reference.

p must be chosen w.r.t. to the density of the point cloud and must be large enough to avoid the event that neighbouring points fall into the same bin and are thus assigned the same index.

3.4 *Matching of 3D Points*

In order to determine the change of orientation between two epochs, it is necessary to establish correspondences between individual points of both point clouds. To find matched points a similarity measure $d_{i,j}$ is computed for all possible combinations of the n respectively m compact distinctive descriptions $\mathbb{S}_i$ and $\mathbb{S}_j$ for $t = i$ and $t = j$, where $n = |\mathbb{S}_i|$ and $m = |\mathbb{S}_j|$. The chosen similarity measure should support reliable matching of distinctive descriptions even if they match only partially and have a different number of entries. From section 3.3 it is clear that the proposed distinctive description encodes the structure of the local neighbourhood in a one-dimensional vector of unique integers: that means if two descriptions have identical entries it is very probable that they encode the structure of the same neighbourhood. Thus a possible similarity measure for the identification of matched points can be defined by the overlap of the two ordered and unique sets $\mathbb{S}_i$ and $\mathbb{S}_j$.

The first step to compute the overlap consists in determining the intersection $\mathbb{S}_\cap = \mathbb{S}_i \cap \mathbb{S}_j$ and the union $\mathbb{S}_\cup = \mathbb{S}_i \cup \mathbb{S}_j$ for the distinctive descriptions $\mathbb{S}_i$ and $\mathbb{S}_j$. Both the intersection and the union can be found efficiently as we are using sets of integer values for which a comparison of two values is very fast.

The similarity $d_{i,j} : d_{i,j} \in [0\ldots1]$ of two distinctive descriptions is then given by (3).

$$d_{i,j} = \frac{|\mathbb{S}_\cap|}{|\mathbb{S}_\cup|} = \frac{|\mathbb{S}_\cap|}{|\mathbb{S}_i| + |\mathbb{S}_j| - |\mathbb{S}_\cap|}. \tag{3}$$

Two descriptions are matched if their matching score $d_{i,j}$ is the highest (greedy approach) and is above a pre-defined threshold t_i. The threshold t_l can be computed by defining a minimum number n_{min} of neighbours that must be visible at $t = i$ and $t = j$ for a successful match. A formula for the computation of t_l is given in (4) where $n = |\mathbb{S}_i|$ and $m = |\mathbb{S}_j|$.

$$t_l = \frac{n_{min}^2}{n + m - n_{min}^2} \tag{4}$$

At the end of section 3.3 it is mentioned that it is important to select a value of p according to the point density of the observed 3D point clouds. If p is too small two neighbouring points might fall in the same bin and are assigned the same index. Such an event reduces the total number of entries in the distinctive description $\mathbb{S}$, because we allow only unique values to be present. As a result the identification of matching points might fail because the minimum number of identical entries (see (4)) is not reached.

3.5 *Robust Filtering of Matches*

Applying the matching scheme described above to point clouds generated by a stereo system at $t = i$ and $t = j$ results in a set of m candidate pairs $\{({}^{i_k}\mathbf{x}, {}^{j_l}\mathbf{x})\}$ for point correspondences. Following the insights of Sattler et al. (2009) we use a RANSAC based approach to eliminate wrong correspondences and to estimate the change of orientation between two epochs. The functional model used by the RANSAC algorithm is given in Weng et al. (1992) and needs at least three non-collinear correspondences to compute the change of orientation ${}^j\mathbf{T}_i$ between the two epochs $t = i$ and $t = j$.

In every iteration of the algorithm three pairs of correspondences are chosen randomly. In order to evaluate the quality of each hypothesis all points at $t = j$ are transformed into the epoch $t = i$ using ${}^j\mathbf{T}_i$ and a best match for every transformed point is identified with a nearest-neighbour search. The quality of the current hypothesis e_{hyp} is given by (5) where n is the number of matches identified by the nearest-neighbour search and $\|\cdot\|$ is the L2-norm of a vector.

$$e_{hyp} = \sum_{k=1}^{n} \left\| {}^{i_k}\mathbf{x} - {}^{i}\mathbf{R}_j\, {}^{j_k}\mathbf{x} + {}^{i}\mathbf{t}_j \right\| \tag{5}$$

The orientation ${}^i\mathbf{T}_j$ is computed from all correspondences found by the nearest-neighbour search, of the best hypothesis according to the functional model given in Weng et al. (1992).

4 Experimental Results

To evaluate the performance of the proposed matching scheme three different image sequences were recorded with a multi-stereo system traversing down a hallway. The two stereo systems A and B were mounted on a mobile platform in a way that the respective fields of view faced the opposite walls. The performed motion patterns were a pure planar motion without any rotation (E1), a planar motion where a rotation was only possible around the normal of the ground plane (E2) and a free motion with rotation around all axis (E3). An example of the planar motion of case E2 is given in Fig. 2. The sequences were recorded with a frame rate of 15 Hz and the stereo systems moved at a speed of approximately $2\,\frac{\mathrm{m}}{\mathrm{s}}$ between 2 and 4 metres down the hallway. The total number of frames captured in the different experiments is given in Tab. 1.

For the experiments signalized circular targets that can be identified reliably in the images were fixed on the walls of the hallway. The position of the centre of these targets is found using an ellipse measurement algorithm (LUHMANN 1986). The result of this feature extraction step is a list of image coordinates for the targets in both images of a stereo pair. With the known epipolar geometry of the stereo system corresponding targets can be identified easily using the approach in OTEPKA et al. (2002), where the fact is exploited that matching points in both images of the stereo system

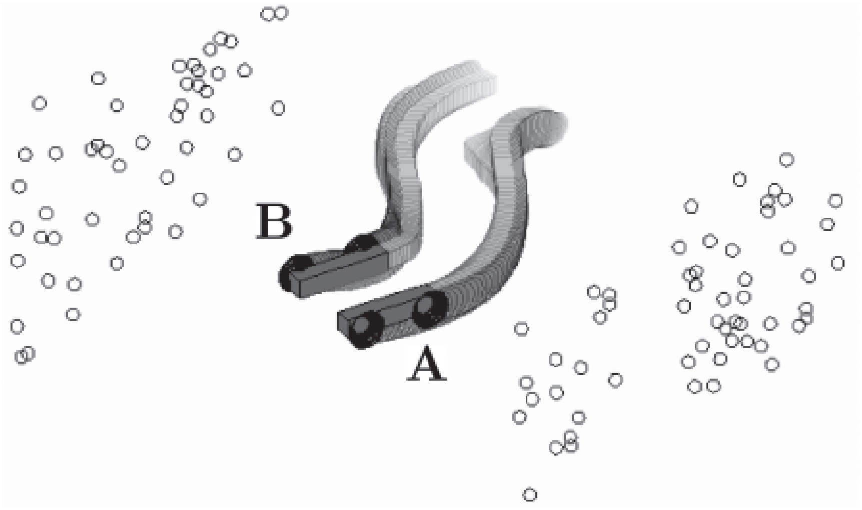


Fig. 2: Planar motion of two rigidly connected stereo systems **A** and **B** traversing down a hallway and observing signalized target on the opposite walls.

Tab. 1: Aggregated numbers used for the performance evaluation of point matching between two epochs.

	E1 A	E1 B	E2 A	E2 B	E3 A	E3 B
# stereo pairs	121	120	200	160	290	290
# connected epochs	3173	3247	5313	4964	19215	23273
# possible matches	64197	64436	91085	95359	361526	469029
# established matches	49580	50481	78121	72811	214285	314917
# wrong matches	0	0	0	674	60	5
match quality $q_{i,j}$	77 %	79 %	86 %	76 %	59 %	67 %

have identical epipolar angles. The epipolar angle for a point is given as the intersection angle of its image ray and an epipolar plane defined e.g. by the epipole and the optical axis. Finally, the 3D point cloud is given by a forward intersection for all identified stereo correspondences.

4.1 *Evaluation Criteria*

The evaluation of the performance of the proposed matching algorithm follows the scheme presented in Mikolajczyk et al. (2005) where it is used to compare different image matching approaches. While Mikolajczyk et al. (2005) analyse the performance of the detection and the matching step, we concentrate only on the matching. In our case the detection has been performed by the stereo reconstruction of the extracted image features (signalized targets) and will not be examined any further.

For the evaluation the true number of possible matches between the point clouds of any two epochs must be known. In order to provide such information a bundle adjustment for all six sequences was performed (three different motion patterns for two stereo systems **A** and **B**). Within the adjustment the change of orientations ${}^{i}\mathbf{T}_0$ w.r.t. to a reference epoch t_0 and the 3D coordinates of the point cloud, defined in the global frame of reference, were estimated. The inner orientation for all cameras, the relative orientations of the stereo systems and the lengths of their baselines were determined in advance and used as fixed parameters in the adjustment. Given the adjusted ${}^{i}\mathbf{T}_0$ the transformation ${}^{j}\mathbf{T}_i$ for any combinations of two epochs $t = i$ and $t = j$ can be computed. Then the point cloud of $t = i$ is transformed into the epoch $t = j$ and all possible matches are identified using a nearest-neighbour search. This number is used as ground truth for the evaluation.

To differentiate between different aspects influencing the performance of the matching we first compute a quality measure $q_{i,j}$ and then three values $o_{i,j}$, $v_{i,j}$ and $s_{i,j}$ for every combination of epochs characterizing three possible sources that affect the proposed algorithm:

Overall match quality

The match quality $q_{i,j}$ is computed as ratio $q_{i,j} = \frac{\left|\mathbb{P}_{m_{i,j}}\right|}{\left|\mathbb{P}_{i\cap j}\right|}$, where $\left|\mathbb{P}_{i\cap j}\right|$ is the ground truth for the number of possible matches between $t = i$ and $t = j$ and $\left|\mathbb{P}_{m_{i,j}}\right|$ is the number of matches established by the proposed algorithm.

Overlap of point clouds

The overlap $o_{i,j}$ of the point clouds at $t = i$ and $t = j$ can be computed with (6) and is identical to the computation of the similarity measure given in section 3.4.

$$o_{i,j} = \frac{\left|\mathbb{P}_{i\cap j}\right|}{\left|\mathbb{P}_{i}\right| + \left|\mathbb{P}_{j}\right| - \left|\mathbb{P}_{i\cap j}\right|} \tag{6}$$

Change of view direction

The change of the view direction $v_{i,j}$ is computed w.r.t. to the normal direction of the points in the point cloud (see section 3.1). From the set $\left|\mathbb{P}_{i\cap j}\right|$ of points visible at both epochs, those matched pairs $\{{}^{i_k}x, {}^{j_k}\mathrm{x}\colon \forall k = 1\ldots n\}$ are selected for which a normal direction $\mathbf{n}$ could be computed. For all n pairs $\{{}^{i_k}n, {}^{j_k}n\}$ the angular difference v_k is given by $v_k = \arccos({}^{ik}n \cdot {}^{jk}n)$, where ${}^{ik}n \cdot {}^{jk}n$ is the dot product of two vectors and the function $\arccos(\cdot)$ returns an angle in the interval $[0\ldots\pi]$. From all v_k the median $\bar{v}_k$ is determined and finally $v_{i,j}$ is given by $v_{i,j} = \bar{v}_k$.

Scale change

To assess the influence of scale changes, i.e. different distances between the stereo system and the point cloud, on the matching process, the number $s_{i,j}$ is computed. First, the centroids of the point clouds at $t = i$ and $t = j$ are computed according to (7).

$$
\begin{aligned}
{}^{i}\hat{\mathbf{x}} &= \frac{1}{n}\sum_{k=1}^{n} {}^{i_k}\mathbf{x} : {}^{i_k}\mathbf{x} \in \left|\mathbb{P}_{i\cap j}\right| \\
{}^{j}\hat{\mathbf{x}} &= \frac{1}{n}\sum_{k=1}^{n} {}^{j_k}\mathbf{x} : {}^{j_k}\mathbf{x} \in \left|\mathbb{P}_{i\cap j}\right|
\end{aligned} \tag{7}
$$

The number $s_{i,j}$ is then given as a relative change $s_{i,j} = \frac{\min\left(\left\|{}^{i}\hat{\mathbf{x}}\right\|, \left\|{}^{j}\hat{\mathbf{x}}\right\|\right)}{\max\left(\left\|{}^{i}\hat{\mathbf{x}}\right\|, \left\|{}^{j}\hat{\mathbf{x}}\right\|\right)}$ of the distances at $t = i$ and $t = j$ w.r.t. to the centroids, where the functions min(·,·) and max(·,·) return the minimum and maximum values of their respective arguments.

The values $o_{i,j}$, $v_{i,j}$ and $s_{i,j}$ reflect the two major influences on the matching process. On the one hand, the overlap $o_{i,j}$ can be used to assess the influence of changing neighbourhoods on the matching process, because point clouds with a low overlap usually show large changes in the local neighbourhood of a point as well. On the other hand, uncertainties in the 3D position of points may lead to different quad tree indices during the computation of the distinctive description. The biggest influences on the point uncertainty result from large differences in the view directions and distance changes.

4.2 *Evaluation Results*

For the computation of the distinctive description for every point of every point cloud the parameters defined in section 3 must be set according to the point density of the point cloud. For the experiments the following values are chosen: maximum radius of local neighbourhood r_{max} = 750 mm, maximum number of selected neighbours n_{max} = 12, minimum number of neighbours n_{max} = 4 and total number of bins (2^p) for the computation of the quad tree index with $p = 64$.

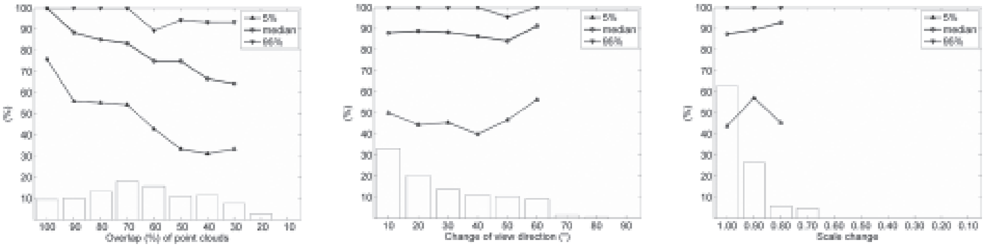
A first impression of the performance of the proposed matching scheme is given by the results presented in Tab. 1. Here the number of connected epochs are those combinations of epochs where the respective point clouds have at least three identical points, the possible matches represent the true number of matches derived from the results of the bundle adjustment (see section 4.1) and the wrong matches are the false positives before robust filtering. The differences in the number of recorded frames and connected epochs are caused by the different motion pattern. In the first experiment the stereo systems were moved in a straight and direct line down the hallway and thus only a smaller number of frames is necessary to capture the entire scene. In the last experiment, the systems were carried by hand and moved forward and backward along the corridor with changing rotations and translations. Such a motion pattern led to a larger number of frames and the continuously changing view directions resulted in more overlapping fields of view than for the translational motion and thus a larger number of connected epochs.

Tab. 1 shows that in most of the six experiments about 70–80 % of all possible matches were identified based on their respective distinctive descriptions. The most remarkable fact is that, at least w.r.t. to the large number of true matches, almost no false matches were established. None of these false positives were used to compute the change of orientation ${}^{j}\boldsymbol{T}_{i}$ between two epochs, because they were all successfully eliminated by the robust filtering process given in section 3.5.

For the detailed analysis of the performance of our proposed matching algorithm in Fig. 3, the match qualities $q_{i,j}$ for any combination of epochs $t = i$ and $t = j$ are grouped w.r.t. to the different causes that possibly affect the matching process, i.e. the computed $o_{i,j}$, $v_{i,j}$ and $s_{i,j}$. For the visualization in Fig. 3 the different groups are plotted on the x-axis and the distribution of the match quality for every group is plotted on the y-axis.

For clarity of the representation the distribution of the $q_{i,j}$ for every group is represented by the 5 %- and 95 %-quantiles and the median. The graphs in the Fig. 3 show a representative subset of the results for all six image sequences. No result for the first experiment (pure translation) is shown, as that motion pattern did not give enough variety for the change of view directions and scale.

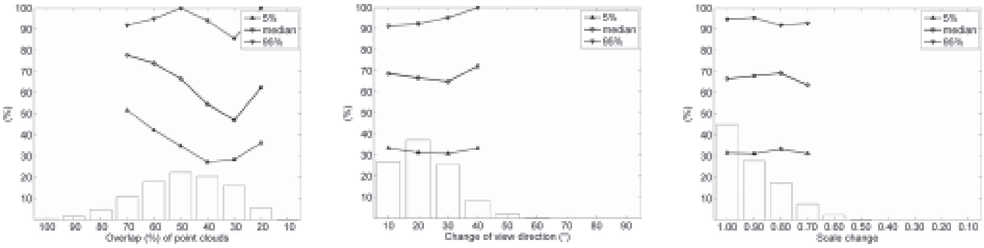
For the graphs the groups with a very low relative frequency are statistically not relevant and have been omitted. Figs. 3(a)–3(f) clearly show that the major source of influence is the change of the local neighbourhood indicated by a low overlap between the point clouds. This is supported by the observation that the median of the match quality is decreasing and the distribution is broadening with a reduced overlap. The Figs. 3(a) and 3(d) show that up to an overlap of 70 % for 90 % of all connected epochs, i.e. those between the 5 % and the 95 % quantiles, a match quality in the range of 60 %–100 % (Fig. 3a) and 55 %–95 % (Fig. 3(d)) is achieved. For an overlap of 40 % the median goes down to 70 % (Fig. 3(a)) and



(a) Distribution of match quality $q_{i,j}$ w.r.t. overlap of point clouds (E2 **A**)

(b) Distribution of match quality $q_{i,j}$ w.r.t. change of view direction (E2 **A**)

(c) Distribution of match quality $q_{i,j}$ w.r.t. change of scale (E2 **A**)



(d) Distribution of match quality $q_{i,j}$ w.r.t. overlap of point clouds (E3 **B**)

(e) Distribution of match quality $q_{i,j}$ w.r.t. change of view direction (E3 **B**)

(f) Distribution of match quality $q_{i,j}$ w.r.t. change of scale (E3 **B**)

Fig. 3: The graphs show the dependency between the achieved match qualitites $q_{i,j}$ and the different sources influencing the matching. The bars at the bottom represent the relative frequency of the respective group. The lines in the graph represent the distribution of the grouped matching quality using the 5 %-, 95 %-quantiles and the median. Groups with a relative frequency smaller than 5 % are not considered.

50 % (Fig. 3(d)) respective 70 % and 90 % of all connected epochs achieve a match quality in the interval from approximately 30 %–95 %.

The position uncertainty of a 3D point does not seem to have a strong influence on the match quality. Figs. 3(b), 3(c), 3(e) and 3(f) show that a change of the respective influence parameter does not change the median or the shape of the distribution significantly.

Larger changes in the local neighbourhood, caused by missed detections or points moving in or out of the stereo system's field of view, obviously have the effect that the similarity measure computed for two actually matching points is below the lower threshold t_l defined in section 3.4 and as a result that match is rejected incorrectly.

5 Conclusions and Future Work

A new approach to establish correspondences between points of sparsely populated and varying point clouds is presented in this paper. The identified matches can be used e.g. to estimate the change of orientation ${}^{j}\mathbf{T}_i$ between two epochs $t = i$ and $t = j$. The proposed algorithm is based on the known basic principles of matching using distinctive feature descriptions and extends them in a way that they can be used to identify corresponding 3D points in sparse and varying point clouds. The algorithm is an extension of the spin images (JOHNSON & HEBERT 1999). It exploits the fact that applying the spin image algorithm to sparse point clouds gives a sparse 2D distinctive description that can be compressed further. The resulting compact 1D description is designed in a way that it allows an efficient matching of two descriptions.

The evaluation of image sequences recorded by two stereo systems shows that our approach allows an efficient and reliable matching of 3D points. The number of true positives is mostly above 70 % and the number of false positives is much smaller than 1 %. The false positives are all eliminated successfully by robust filtering of the established matches.

A limiting prerequisite of the presented matching scheme is that the point clouds at different epochs $t = i$ and $t = j$ must have the same scale as absolute distances are used in the computation of the distinctive description. For a more general variant of the description it would be possible to use ratios of distances that are invariant against scale changes. Such a variant of the compact distinctive description could be used to extend existing image based matching approaches like SIFT (e.g. Lowe 2004). The identified correspondences between two images can be used to compute 3D model coordinates. Based on these model coordinates a scale invariant distinctive description could be used differently:

Application for checking image based matching

Possible matches with further images are first established using the known image based matching algorithms and then they are additionally verified using a scale invariant distinctive description for 3D points. Only matches passing both approaches are accepted and the number of false positives might be reduced.

Application for connecting images with large perspective change

Usually the image based approaches have problems to correctly identify correspondences for larger perspective changes. Here a distinctive description for 3D points might be helpful to find additional correspondences.

It becomes apparent that it might be advantageous to combine image based distinctive descriptions with the proposed description for 3D data. Such a combination of 2D and 3D data is also presented in Wu et al. (2008).

A further extension would be the lifting of the constraint that the neighbourhood used for the computation of the distinctive description must be approximately planar. The basic principles used here can be transferred easily to the case of arbitrary 3D neighbourhoods. It remains to be investigated how much the changing 3D neighbourhood effects the results.

References

Barnea, S. & Filin, S., 2010: Geometry-Image-Intensity Combined Features for Registration of Terrestrial Laser Scans. – International Archives of Photogrammetry and Remote Sensing **38** (3a): 145–150.

Bay, H., Ess, A., Tuytelaars, T. & Van Gool, L., 2008: Speeded-up robust features (SURF). – Computer Vision and Image Understanding **110** (3): 346–359.

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): 239–256.

Brown, L., 1992. A Survey of Image Registration Techniques. – ACM Computing Surveys **24:** 325–376.

Calonder, M., Lepetit, V. & Fua, P., 2008: Keypoint signatures for fast learning and recognition. – Forsyth, D., Torr, P. & Zisserman, A. (eds.): Computer Vision – ECCV 2008, LNCS **5302:** 58–71, Springer.

Calonder, M., Lepetit, V., Fua, P., Konolige, K., Bowman, J. & Mihelich, P., 2009: Compact signatures for high-speed interest point description and matching. – 2009 IEEE **12th** International Conference on Computer Vision: 357–364.

Chua, C. & Jarvis, R., 1997: Point Signatures: A New Representation for 3D Object Recognition. – International Journal of Computer Vision **25** (1): 63–85.

Finkel, R. & Bentley, J., 1974: Quad trees: a data structure for retrieval on composite keys. – Acta informatica **4** (1): 1–9.

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

Flint, A., Dick, A. & van den Hengel, A., 2008: Local 3D structure recognition in range images. – Computer Vision, IET **2** (4): 208–217.

Gelfand, N., Mitra, N.J., Guibas, L.J. & Pottmann, H., 2005: Robust global registration. – **Third** Eurographics Symposium on Geometry processing: 197–206.

Grün, A. & Akca, D., 2005: Least squares 3D surface and curve matching. – ISPRS Journal of Photogrammetry and Remote Sensing **59** (3):151–174.

Johnson, A. & Hebert, M., 1999. Using spin images for efficient object recognition in cluttered 3D scenes. – IEEE Transactions on Pattern Analysis and Machine Intelligence 1999 **21** (5): 433–449.

Lo, T.-W.R. & Siebert, J.P., 2009: Local feature extraction and matching on range images: 2.5D SIFT. – Computer Vision and Image Understanding **113** (12): 1235–1250.

Lowe, D., 2004: Distinctive Image Features from Scale-Invariant Keypoints. – International Journal of Computer Vision **60** (2): 91–110.

Luhmann, T., 1986: Automatic point determination in a réseau scanning system. – Commission V, Symposium Real-time Photogrammetry – a new challenge **26:** 400–408.

McGlone, C., Mikhail, E. & Bethel, J., 2004: Manual of photogrammetry. – **5th** ed., American Society for Photogrammetry and Remote Sensing, Bethesda, MD, USA.

Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T. & Van Gool, L., 2005: A comparison of affine region detectors. – International Journal of Computer Vision **65** (1–2): 43–72.

Nistér, D. & Stewenius, H., 2006: Scalable recognition with a vocabulary tree. – 2006 IEEE Conference on Computer Vision and Pattern Recognition **2:** 2161–2168.

Novatnack, J. & Nishino, K., 2007: Scale-Dependent 3D Geometric Features. – 2007 IEEE **11th** International Conference on Computer Vision: 1–8.

Otepka, J., Hanley, H. & Fraser, C., 2002: Algorithm developments for automated Off-Line Vision Metrology. – International Archives of Photogrammetry and Remote Sensing **34** (5): 60–67.

Rusinkiewicz, S. & Levoy, M., 2001: Efficient variants of the ICP algorithm. – **Third** International Conference on 3-D Digital Imaging and Modeling: 145–152.

Sattler, T., Leibe, B. & Kobbelt, L., 2009: Scramsac: Improving ransac's efficiency with a spatial consistency filter. – 2009 IEEE **12th** International Conference on Computer Vision: 2090–2097.

Seeger, S. & Laboureux, X., 2000: Feature extraction and registration: An overview. – Girod, B., Greiner, G. & Niemann, H. (eds.), Principles of 3D Image Analysis and Synthesis, The Springer International Series in Engineering and Computer Science **556:** 153–166.

Steder, B., Rusu, R., Konolige, K. & Burgard, W., 2011: Point feature extraction on 3d range scans taking into account object boundaries. – 2011 IEEE International Conference on Robotics and Automation: 2601–2608.

Stückler, J. & Behnke, S., 2011: Interest point detection in depth images through scale-space surface analysis. – 2011 IEEE International Conference on Robotics and Automation: 3568–3574.

Wang, Z. & Brenner, C., 2008: Point based registration of terrestrial laser data using intensity and geometry features. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **37** (B5): 583–590.

Weinmann, M., Weinmann, M., Hinz, S. & Jutzi, B., 2011: Fast and automatic image-based registration of TLS data. – ISPRS Journal of Photogrammetry and Remote Sensing **66** (6): 62–70.

Weng, J., Cohen, P. & Herniou, M., 1992: Camera calibration with distortion models and accuracy evaluation. – IEEE Transactions on Pattern Analysis and Machine Intelligence **14** (10): 965–980.

Winkelbach, S. & Wahl, F.M., 2008: Pairwise matching of 3d fragments using cluster trees. – International Journal of Computer Vision **78:** 1–13.

Wu, C., Clipp, B., Li, X., Frahm, J. & Pollefeys, M., 2008: 3D model matching with Viewpoint-Invariant Patches (VIP). – 2008 IEEE Conference on Computer Vision and Pattern Recognition: 1–8.

Zitová, B. & Flusser, J., 2003: Image registration methods: a survey. – Image and Vision Computing **21** (11): 977–1000.

Addresses of the Authors:

Dr.-Ing. Daniel Muhle, Dr.-Ing. Manfred Wiggenhagen & Prof. Dr.-Ing. Christian Heipke, Leibniz Universität Hannover, Institut für Photogrammetrie und GeoInformation, Nienburger Str. 1, D-30167 Hannover, Tel.: +49-511-762-2729, Fax: +49-511-762-2483, e-mail: {muhle}{wiggenhagen}{heipke}@ipi.uni-hannover.de

Dr.-Ing. Steffen Abraham, Robert Bosch GmbH, Robert-Bosch-Str. 200, 31139 Hildesheim.

Manuskript eingereicht: März 2012
Angenommen: Juni 2012

PFG 2012 / 5, 0547–0556
Stuttgart, October 2012

Article

Multi-View Extraction of Dynamic Pedestrian Density Fields

MATTHIAS PLAUE, MINJIE CHEN, GÜNTER BÄRWOLFF & HARTMUT SCHWANDT, Berlin

Keywords: density estimation, human crowd analysis, pedestrian flows, velocimetry, video analysis

Summary: In the framework of macroscopic models of human crowds, pedestrian dynamics are described via local density and flow fields. In this paper, we expand our previous work on the extraction of pedestrian trajectories and density fields from video recordings of crowd experiments in two ways. Firstly, we include data from different video cameras in order to cover a larger observation area. Secondly, we improve our previous density estimation method by introducing a new kernel function which (a) yields density fields that are also differentiable functions in time and (b) models the influence of multiple neighbouring pedestrians on the personal space of an individual.

We apply this density computation method to pedestrian trajectories extracted from video data of a crowd experiment conducted by us, and compare the results with other common methods for density computation in this context: a technique based on Voronoi diagrams, and a fixed-bandwidth estimator. We come to the conclusion that the technique proposed by us combines advantages from both alternative methods, yielding spatio-temporally smooth density fields close to the standard definition of density at all scales.

Zusammenfassung: *Multiperspektivische Erfassung dynamischer Dichtefelder von Fußgängern.* Makroskopische Modelle zur Beschreibung von Personenbewegungen greifen auf Konzepte wie lokale Dichte- oder Flussfelder zurück. In einer früheren Arbeit haben wir ein Verfahren beschrieben, durch das individuelle Trajektorien sowie Dichtefelder mithilfe von Videoaufnahmen von Experimenten mit Personenströmen erfasst werden können. Dieses Verfahren wurde in zweierlei Hinsicht erweitert bzw. verbessert: Zum einen können Videodaten aus mehreren Kameras dazu verwendet werden, einen größeren Beobachtungsbereich abzudecken. Zum anderen kann unsere Methode der Dichteschätzung durch Verwendung einer anderen Kernfunktion verbessert werden. Der so erhaltene Schätzer liefert Dichtefelder welche (a) bzgl. des Zeitparameters differenzierbare Funktionen darstellen und (b) den Einfluss mehrerer benachbarter Personen auf den von einem Individuum eingenommen Raum modellieren.

Mithilfe dieser Methode berechnen wir Dichtefelder auf Grundlage von Trajektorien, welche aus Videoaufnahmen eines von uns durchgeführten Fußgängerexperiments gewonnen wurden. Das Resultat vergleichen wir mit zwei in diesem Kontext gebräuchlichen Methoden zur Dichteberechnung: einem auf Voronoi-Diagrammen basierenden Verfahren sowie einem Kerndichteschätzer mit konstanter Bandbreite. Wir kommen zu dem Schluss, dass mit unserem Ansatz Vorteile der beiden alternativen Methoden vereint werden, indem die Berechnung raumzeitlich glatter Dichtefelder ermöglicht wird, welche auf allen Skalen standardmäßig berechnete Dichtewerte approximieren.

1 Introduction

The study of pedestrian dynamics has important applications in crowd management such as devising strategies for the evacuation of buildings or public places. In order to evaluate the predictive power of mathematical models designed to emulate human crowd behaviour, it is a common procedure to compare numerical simulations based on these models with empirical data.

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0138
1432-8364/12/0138 $ 2.25

Furthermore, different modelling approaches demand the extraction of different types of data: For example, the social force model (Helbing & Molnár 1995) and cellular automaton model (Burstedde et al. 2001) aim at predicting pedestrian trajectories, whereas continuum methods adopted from fluid mechanics (Hughes 2002) describe the dynamics via the density and flow of the crowd. In our work, we develop models based on these three approaches in order to simulate intersecting pedestrian flows and compare these simulations with the real world. Here, we describe one important part of this work: the extraction of the trajectories and a dynamic, continuous density field from video recordings of human crowd experiments. The work presented here is an extension of Plaue et al. (2011).

1.1 *Challenges and Contribution*

In section 2, we describe an experiment that was conducted with the purpose of demonstrating the dynamic behaviour of intersecting pedestrian flows (Fig. 1). To this end, we set up an experiment where two unconfined, perpendicularly intersecting pedestrian flows have been recorded by multiple cameras with overlapping fields of view. To the best of our knowledge, no such experiment has previously been conducted.

In that section, we also describe a semi-automatical technique to extract the spatio-temporal positions of pedestrians in a crowd of low density at close range from an arbitrary observation angle. Due to constructional limitations, it was not possible to install the cameras to provide a bird's eye view. This situation is very different from most experimental setups found in the common literature, where a bird's eye view installation of the camera(s) in sufficient height provides an advantageous perspective. In some studies, the heights of the pedestrians are also indicated by visual markers, which also greatly facilitate automated pedestrian tracking (Boltes et al. 2010). Having neither bird's eye views nor markers, we devised a method to extract the floor positions of the pedestrians without knowing their re-

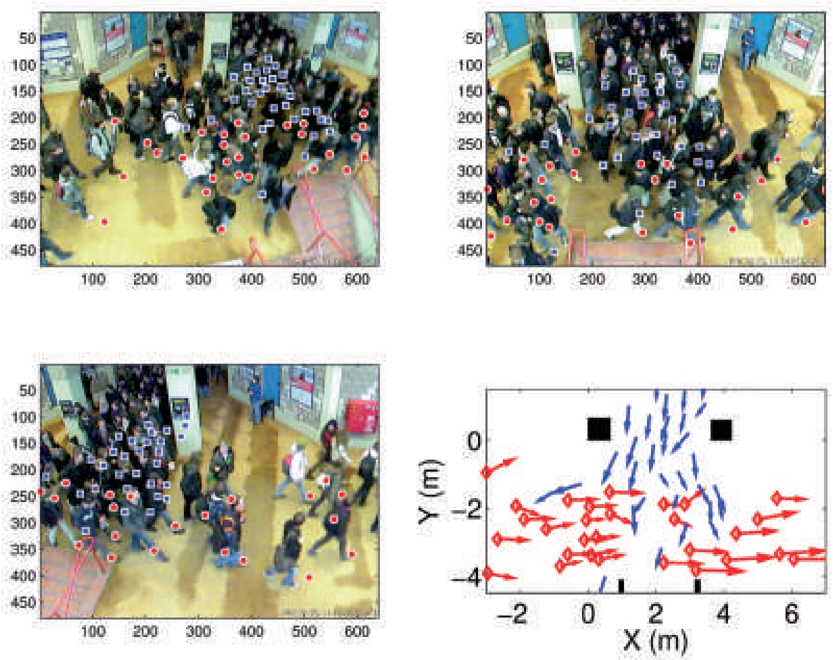


Fig. 1: Human crowd experiment from three different observation points. Bottom right: extracted pedestrian positions at time step *t = 68.2 s* of group *A* (red diamonds) and group *B* (blue). Arrows indicate current velocity; the maximal length corresponds to *1.4 m/s*.

spective body heights beforehand. Furthermore, we perform a fusion of data from multiple video cameras.

The estimation of the density field of a sparse crowd is a challenging task because of the low number of samples. In section 3, we propose a novel method to compute a high-resolution smooth density field from the spatio-temporal positions of the pedestrians based on a kernel density estimator with variable bandwidth. In that section, we also give a comparison of our results with other density estimators. Finally, an overview of future work will be given in section 4.

1.2 *Related Work*

Human crowd experiments

Empirical data for the evaluation of human crowd models are usually extracted from video recordings of either naturally occurring crowds (HELBING et al. 2007) or pedestrian flows that have been produced under controlled conditions (DAAMEN & HOOGENDOORN 2003, GALEA et al. 2011, ZHANG et al. 2011). In general, the latter are devised to demonstrate crowd behaviour in special situations such as evacuation or passage through a bottleneck. In our work, we wish to analyze the dynamics of intersecting pedestrian flows. A very similar experiment with this purpose has been conducted by GUO et al. (2010); however, data from only one camera was processed in that case. Also, in our experiment, the pedestrians did not move along specified, confined corridors.

(Semi-) automatic pedestrian tracking

There exists a large body of literature on methods for the detection and automatic tracking of humans. The range of application of such algorithms varies greatly. For example, SCHMIDT & HINZ (2011) address the problem of tracking the positions of pedestrians from aerial images which provide very low resolution, and CREMERS (2006) proposes a method to track the contour of an individual through noise and occlusion. For an overview on different methodologies, we refer to HU et al. (2004).

In this work, we follow the suggestion of BOLTES et al. (2010), who process video data similar to ours, and use the Lucas-Kanade algorithm (SHI & TOMASI 1994) to facilitate the extraction of the spatio-temporal positions of the pedestrians. Additionally, we employ an algorithm to merge the trajectories from different overlapping camera views via the Hungarian method (KUHN 1955, MUNKRES 1957). Note that we merge the trajectories *after* processing the video data, in contrast to the detection of objects from multiple views during tracking (see for example KAHN et al. 2001).

Pedestrian density estimation

Probably the most basic way to compute a density would be to divide the number of pedestrians in a given region by the area of that region, at a given point in time. However, this "standard" density estimator is not a smooth point-wise density function and yields data with large scatter.

At least two approaches for measuring the (local) density of human crowds have been suggested in the literature as alternatives:

- In HELBING et al. (2007), a local density field is computed via the sum of Gaussians with fixed standard deviation (typically 0.7 m) centred at each pedestrian. Formally, this approach is identical to kernel density estimation with fixed bandwidth, which is a basic tool in statistical data analysis (see SILVERMAN 1986, for example). This method yields a smooth density field defined at every point.
- In STEFFEN & SEYFRIED (2010), estimators are proposed based on the Voronoi diagram defined by the position of each pedestrian as a Voronoi site. The main idea in this approach is to account for the personal space occupied by each pedestrian, and this personal space is represented by the area of the corresponding Voronoi cell. The values for the Voronoi density are very close to standard densities, but with a significantly smaller scatter. However, the Voronoi estimator does not yield a smooth local density defined at every point.

The algorithm that we propose here is conceptually a combination of the Voronoi estimator (accounting for personal space) and the fixed-bandwidth kernel estimator (yielding smooth density fields).

2 Experiments and Trajectory Extraction

In the following, we describe human crowd experiments that we conducted in the lobby of the Department of Mathematics building of Technische Universität Berlin in December 2010, and the extraction of the trajectories of the participants from video streams captured by several cameras.

2.1 *Experimental Setup*

In the experiment which we use to illustrate our method, two pedestrian flows (group *A*, 142 subjects, and group *B*, 83 subjects) intersected at an angle of 90 degrees for one minute in a region of about 25 m^2, reaching a peak density of about five pedestrians per m^2. The scene was recorded from a gallery at a height of about 6 m with five networked and temporally synchronized JVC VN-V25U surveillance video cameras. Here, we will analyze the data provided by the three central cameras which covered the area where the actual intersecting of the pedestrian flows took place (Fig. 1).

2.2 *Extraction of Spatio-Temporal Positions*

For camera calibration, we assumed a pinhole model and estimated the model parameters by measurement of the world and image coordinates of about 30 fixed reference points in the scene. For each camera, this procedure resulted in a camera matrix, and thus enabled us to deduce the parameters of a homography between the camera's image plane and the floor. The video data have been analyzed in a semi-automatical manner for each camera as described in the following (see also Plaue et al. 2011). Our main goal is the supervised extraction of reliable data.

1. The video is played back frame by frame. Image segments corresponding to the heads of pedestrians that newly enter the scene are marked manually. These templates are used to determine the head positions in the next frame via the Lucas-Kanade tracking method (Shi & Tomasi 1994). Based on the difference of spatial positions in consecutive frames as well as the residual i.e., difference between the detected texture and the template, an error score is computed for each person. If this score is too high or the user spots a possible tracking error regardless of the score, the head position/template can be corrected manually.
2. Once the image coordinates of the heads of the pedestrians have been determined, the video is played again from the beginning. For each pedestrian in the frame, the head position is shown together with the corresponding floor position, initially under the assumption that every pedestrian has a standard height of $h = 1.70\ m$. In each frame the user may correct the floor position of a pedestrian by simply clicking into the frame, and the current height and floor position coordinates are updated via the homography determined in camera calibration.
3. Each pedestrian is assigned a final height value equal to either the arithmetic mean of the height values from the corrections in step 2, or equal to the standard height $h = 1.70\ m$ if no user instruction for this pedestrian is available during this step. Based on this final height value and the image coordinates of the head, in each frame we compute the world coordinates $(X, Y, 0)$ of each pedestrian's position on the floor.

Remarks

- In our scenario, the floor position of most of the pedestrians is visible at some point in time, for example before entering or exiting the crowded intersection area. Therefore, manual correction of the floor position is feasible.
- In order to improve the user's corrections of the floor positions it might be reasonable to provide a view of all cameras and the respective positions during step 2.
- One might introduce/implement a pattern recognition module to carry out the func-

tion of an automatic marker. However, even if the whole algorithm provided a fully automatic analysis, manual verification would nevertheless be good experimental practice in order to obtain reliable data.

2.3 *Merging Trajectories from different Camera Views and Smoothing*

The above procedure yields the positions of the pedestrians on the floor covered by each camera. Originally the cameras were positioned so that these floor areas overlapped. However, it was not immediately possible to merge the trajectories since they are not labelled as individual pedestrians. To solve this problem, we implemented the following algorithm:

1. For each pair of pedestrians captured by different cameras, compute their distance in each frame. Compute the mean value across the frames. Due to measurement errors, this value does not vanish even if it is computed for the same pedestrian captured by two different cameras. However, we expect the mean distance to be minimal if the same pedestrian is captured by two different cameras. If two pedestrians do not appear together in at least one frame, a very large distance value is assigned to this pair of pedestrians. If one camera captures fewer pedestrians than the other, pedestrians very far away will be added to this data set to yield a square distance matrix.
2. The problem to find the permutation of labels that yields the minimal distance for each pair of pedestrians is a combinatorial optimization problem that we solve with the Kuhn-Munkres algorithm, also known as the Hungarian method (KUHN 1955, MUNKRES 1957). Data that cannot be assigned automatically can be assigned manually, or be discarded. In our case, data from about 15 pedestrians had to be managed in this way.
3. Due to systematic errors such as lens distortion and due to errors in the measurement process, the positions of the pedestrians from the cameras on the side show a displacement with respect to those obtained from the central view. We use the central view as a reference and shift the positions from the cameras on the side towards the corresponding positions from the central camera for the differences between the locations in different videos to be minimized after merging the data. Finally, for each pedestrian, all available data points are approximated by cubic B-splines to yield smooth trajectories $(t, X(t), Y(t))$. By differentiating these trajectories with respect to the time parameter t, the velocities of the pedestrians can be easily computed (Fig. 1).

A frame of the analyzed video sequence can be seen in Fig. 1. Since we use the central camera as the reference view in step 3 above, measurement errors are particularly visible in the camera views from the side. One can see that in this particular scene, the positions of some pedestrians located near the intersection area are not marked. This is due to the fact that these pedestrians could not be reliably assigned a trajectory over a sufficiently extended time period, and therefore were discarded. Note that with the currently available image size of 640 × 480 pixels, it proves difficult to trace individual pedestrians in a very crowded scene, even for an attentive human observer. In our earlier experiments, we bypassed this problem by equipping the subjects with coloured clothing hoping to establish a better visual contrast.

3 Variable-Bandwidth Kernel Density Estimation

In the following, we describe and investigate a novel method for kernel density estimation. We apply this technique to compute a density field from the trajectories of the pedestrians.

3.1 *Definition*

Consider a Gaussian kernel density estimator with variable bandwidth to compute the density at time t and position $\boldsymbol{x}$:

$$\rho(t,\mathbf{x}) = \frac{1}{2\pi}\sum_{i\in J}\frac{1}{(\lambda d_i(t))^2}\exp\left(-\frac{\left\|\mathbf{x}_i(t)-\mathbf{x}\right\|^2}{2(\lambda d_i(t))^2}\right). \quad (1)$$

Here, J denotes an index set labelling the pedestrians, and λ is an additional dimensionless smoothing parameter. The bandwidth $\lambda d_i(t)$ is estimated from the trajectories $\boldsymbol{x}_i(t)$ of the pedestrians – the formal analogy in statistical data analysis is also known as a sample smoothing estimator (Terrell & Scott 1992). For example, assuming $\lambda = 1$, for the nearest-neighbour kernel estimator (Plaue et al. 2011),

$$d_i(t) = \min_{j\in J, j\neq i} \left\| \mathbf{x}_i(t) - \mathbf{x}_j(t) \right\|. \qquad (2)$$

However, this kernel and therefore the total density are not differentiable with respect to time. Furthermore, it does not account for the fact that the personal space of a pedestrian is affected not only by the nearest pedestrian but also by other pedestrians in the immediate vicinity. Therefore, we propose the following alternative:

$$d_i^{(p)}(t) = \left(\sum_{j\in J, j\neq i} \left\| \mathbf{x}_i(t) - \mathbf{x}_j(t) \right\|^{-p} \right)^{-1/p}. \qquad (3)$$

This is a smooth function and at the same time generalizes the nearest-neighbour kernel as its limiting case of $p \rightarrow \infty$.

3.2 *General Properties and Parameters*

In Fig. 2, a toy-model calculation for a single pedestrian is shown in order to demonstrate how the bandwidth is determined by multiple neighbouring pedestrians for reasonable values of the parameters p and λ. For large values of the parameter p, the bandwidth only depends on the nearest neighbour. For small values of p, the bandwidth is a function of all nearby pedestrians, and it decreases with the number of nearby pedestrians. Therefore, this parameter defines the degree to which other nearby pedestrians influence personal space.

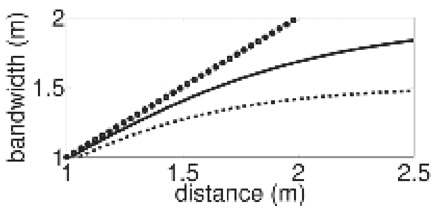


Fig. 2: The bandwidth, defined by (3) with *p = 4*, assigned to a particular pedestrian *A* as a function of the distance to another individual pedestrian *B.* Dotted line: with no other pedestrian present. Solid line (dashed line): with one other pedestrian *C* (three other pedestrians C, D and E) located at a constant distance of two metres to *A*.

Fig. 3 shows the density field computed with this kernel at a particular point in time. By comparison with the fixed-bandwidth estimator, this figure also illustrates how the variable-bandwidth estimator distributes “pedestrian mass” to favour densely crowded regions. We would like to note that this feature is consistent with a model assumption that is frequently found in the description of pedestrian dynamics (“chemotaxis”): interactions between pedestrians are repulsive for short distances and attractive for longer distances (see for example Schadschneider et al. 2002). Also, we expect the proposed density estimator is useful for the visualization of other types of data, in particular if one is interested in highlighting clusters. For large values of the parameter λ, the density field becomes more spatially smoothed and less “fine-grained”, distributing pedestrian mass more broadly.

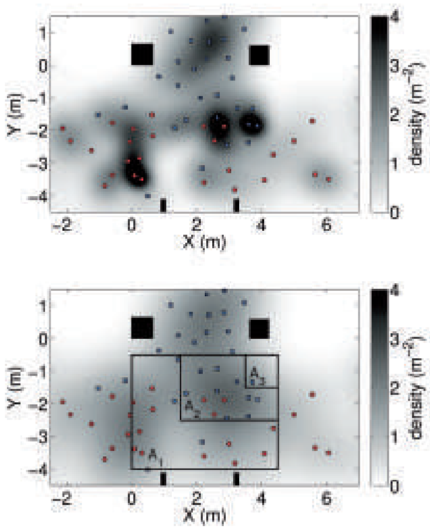


Fig. 3: Top: pedestrian density field, computed with the kernel defined by (3) with *λ = 1*, *p = 4;* bottom: computed with a fixed bandwidth of *d = 1 m*. Black indicates a density *> 4 m⁻²*.

3.3 *Comparison with other Density Estimators*

In Fig. 4, a plot of the density versus time is shown, averaged over the regions marked in Fig. 3, and computed by four methods: the standard method of counting people in the region, a fixed-bandwidth kernel estimator, our variable-bandwidth estimator, and finally the density estimator based on Voronoi diagrams denoted in STEFFEN & SEYFRIED (2010) as "D_V". The regions have the respective areas $A_1 = 15.8\ m^2$, $A_2 = 6\ m^2$, and $A_3 = 1\ m^2$.

Remark

The Voronoi method in its original form is not designed for unconfined crowds; we work around this fact by assuming that the pedestrians stop and cease to move once they exit the area covered by the cameras, thereby limiting the size of the boundary Voronoi cells. More recently, LIDDLE et al. (2011) propose to simply cut off the Voronoi cells beyond a disk of certain radius centred at the respective pedestrian.

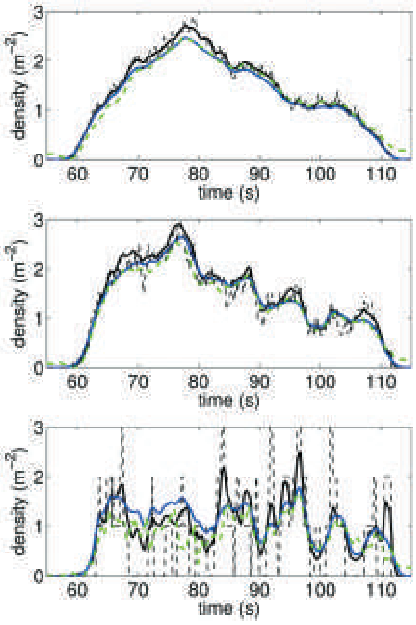


Fig. 4: From top to bottom: pedestrian density, spatially averaged across the areas marked A_1, A_2 and A_3 in Fig. 3. Kernel density with fixed bandwidth $d = 0.7\,m$ (blue line), with variable bandwidth (black line), Voronoi density (dashed green line), standard density (thin black dashed line).

All methods yield results very similar to the standard density when computed for very large regions, with the possible exception of the Voronoi estimator (without cut-off) because of boundary cells of infinite size. However, for smaller regions, the fixed-bandwidth estimator typically yields values that are significantly lower than the standard density since a large portion of the pedestrian mass is located outside the respective region. For very small "microscopic" regions, such as A_3, the densities computed with the fixed-bandwidth estimator can be larger than the standard density since pedestrian mass from outside the region cumulates inside the region regardless of the number of pedestrians already occupying that location. In contrast to this, Fig. 4 demonstrates that the estimator proposed by us yields values that are close to the standard density at all scales. As a result, small temporal variations in density are also described more faithfully by this estimator. Therefore, we may compute pedestrian density data with high spatio-temporal resolution and high precision. We expect that this feature is particularly useful to analyze the fine structure of fundamental diagrams (ZHANG et al. 2011).

4 Conclusion and Future Work

In this work, we present a framework for measuring local density fields from video recordings of human crowds captured by multiple cameras. By utilizing methods from photogrammetric image analysis, we first extract the trajectories of the pedestrians from each camera, and merge these data by matching locations with minimum spatio-temporal distance.

From these trajectories, we compute the pedestrian density field via a modified version of a nearest-neighbour kernel estimator recently proposed by us, with an additional parameter p that serves as a temporal smoothing parameter for the bandwidth. The density obtained in this way is a smooth function of the object coordinates and time, and faithfully represents the standard density when averaged over regions of arbitrary sizes.

In addition, the reader may be aware that the density field is applicable in the estimation of a flow field by requiring that the continuity equation holds. This approach is the subject of present work, the results of which will be reported in Plaue et al. (2012), where we will also describe how obstacles and boundaries can be taken into account.

Note that the automatic, data-driven estimation of values for p and λ is still an open problem in our context. Methods from statistical data analysis for automatic bandwidth selection might prove to be appropriate tools to attack this problem (Comaniciu 2003, Wu 2007).

Acknowledgements

We would like to thank all university staff and students who helped with conducting the experiments, and we especially thank C. Neumann for carrying out the data analysis and implementing the density estimation method based on Voronoi diagrams. Furthermore, we thank the reviewers for their helpful comments and suggestions.

The authors gratefully acknowledge the support of Deutsche Forschungsgemeinschaft (German Research Foundation) for the project SCHW548/5-1 + BA1189/4-1.

The numerical calculations were made with the computing software MATLAB by MathWorks.

Finally, we would like to thank the organizers of the conference Photogrammetric Image Analysis 2011, Technical University of Munich, Germany.

References

Boltes, M., Seyfried, A., Steffen, B. & Schadschneider, A., 2010: Automatic Extraction of Pedestrian Trajectories from Video Recordings. – Pedestrian and Evacuation Dynamics, PED **2008:** 43–54.

Burstedde, C., Klauck, K., Schadschneider, A. & Zittartz, J., 2001: Simulation of Pedestrian Dynamics Using a Two-Dimensional Cellular Automaton. – Physica A **295:** 507–525.

Comaniciu, D., 2003: An Algorithm for Data-Driven Bandwidth Selection. – IEEE Transactions on Pattern Analysis and Machine Intelligence **25** (2): 281–288.

Cremers, D., 2006: Dynamical Statistical Shape Priors for Level Set-Based Tracking. – IEEE Transactions on Pattern Analysis and Machine Intelligence **28** (8): 1262–1273.

Daamen, W. & Hoogendoorn, S.P., 2003: Experimental Research on Pedestrian Walking Behavior. – Transportation Research Board annual meeting: 1–16.

Galea, E.R., Filippidis, L., Wang, Z., Lawrence, P.J. & Ewer, J., 2011: Evacuation Analysis of 1000+ Seat Blended Wing Body Aircraft Configurations: Computer Simulations and Full-Scale Evacuation Experiment. – Pedestrian and Evacuation Dynamics, PED **2010:** 149–151.

Guo, R.-Y., Wong, S.C., Huang, H.-J., Zhang, P. & Lam, W.H.K., 2010: A Microscopic Pedestrian-Simulation Model and Its Application to Intersecting Flows. – Physica A **389:** 515–526.

Helbing, D. & Molnár, P., 1995: Social Force Model for Pedestrian Dynamics. – Physical Review E **51** (5): 4282–4286.

Helbing, D., Johansson, A. & Al-Abideen, H.Z., 2007: Dynamics of Crowd Disasters: An Empirical Study. – Physical Review E **75** (4): 046109.

Hu, W., Tan, T., Wang, L. & Maybank, S., 2004: A Survey on Visual Surveillance of Object Motion and Behaviors. – IEEE Trans. On Systems, Man, and Cybernetics **34** (C3): 334–352.

Hughes, R.L., 2002: A Continuum Theory for the Flow of Pedestrians. – Transportation Research Part B **36**: 507–535.

Khan, S., Javed, O. & Shah, M., 2001: Tracking in Uncalibrated Cameras with Overlapping Field of View. – Performance Evaluation of Tracking and Surveillance (PETS **2001**).

Kuhn, H.W., 1955: The Hungarian Method for the Assignment Problem. – Naval Research Logistic Quaterly **2:** 83–97.

Liddle, J., Seyfried, A., Steffen, B., Klingsch, W., Rupprecht, T., Winkens, A. & Boltes, M., 2011: Microscopic Insights Into Pedestrian Motion Through a Bottleneck, Resolving Spatial and Temporal Variations. – arXiv:1105.1532v1 [physics.soc-ph].

Munkres, J., 1957: Algorithms for the Assignment and Transportation Problems. – Journal of the Society for Industrial and Applied Mathematics **5** (1): 32–38.

Plaue, M., Chen, M., Bärwolff, G. & Schwandt, H., 2011: Trajectory Extraction and Density Analysis of Intersecting Pedestrian Flows from Video Recordings. – ISPRS Conference on Photogrammetric Image Analysis (PIA 2011), LNCS **6952:** 285–296.

Plaue, M., Bärwolff, G. & Schwandt, H., 2012: On Measuring Pedestrian Density and Flow Fields in Dense as well as Sparse Crowds. – to appear in Proceedings Pedestrian and Evacuation Dynamics (PED **2012**).

Schadschneider, A., Kirchner, A. & Nishinari, K., 2002: Cellular Automaton Approach to Collective Phenomena in Pedestrian Dynamics. – 5th International Conference on Cellular Automata for Research and Industry (ACRI 2002), LNCS **2493:** 239–248.

Schmidt, F. & Hinz, S., 2011: A Scheme for the Detection and Tracking of People Tuned for Aerial Image Sequences. – ISPRS Conference on Photogrammetric Image Analysis (PIA 2011), LNCS **6952:** 257–270.

Shi, J. & Tomasi, C., 1994: Good Features to Track. – IEEE Conference on Computer Vision and Pattern Recognition: 593–600.

Silverman, B.W., 1986: Density Estimation for Statistics and Data Analysis. – Chapman and Hall, London, UK.

Steffen, B. & Seyfried, A., 2010: Methods for Measuring Pedestrian Density, Flow, Speed and Direction with Minimal Scatter. – Physica A **389** (9): 1902–1910.

Terrell, G.R. & Scott, D.W., 1992: Variable Kernel Density Estimation. – Annals of Statistics **20** (3): 1236–1265.

Wu, T.-J., Chen, C.-F. & Chen, H.-Y., 2007: A Variable Bandwidth Selector in Multivariate Kernel Density Estimation. – Statistics & Probability Letters **77** (4): 462–467.

Zhang, J., Klingsch, W., Schadschneider, A. & Seyfried, A., 2011: Transitions in Pedestrian Fundamental Diagrams of Straight Corridors and T-Junctions. – Journal of Statistical Mechanics **2011** (6): P06004.

Address of the Authors:

Matthias Plaue, Minjie Chen, Günter Bärwolff & Hartmut Schwandt, Technische Universität Berlin, Institut für Mathematik, D-10623 Berlin, Tel.: +49-30-314-25651, Fax: +49-30-314-21110, web page: http://www.math.tu-berlin.de/projekte/smdpc/, e-mail: {plaue}{minjie.chen}{baerwolf}{schwandt}@math.tu-berlin.de

Manuskript eingereicht: Februar 2012
Angenommen: Juni 2012

PFG 2012 / 5, 0557–0572
Stuttgart, October 2012 **Article**

Testing the Radiometric Performance of Digital Photogrammetric Images: Vicarious vs. Laboratory Calibration on the Leica ADS40, a Study in Spain

David Hernández-López & Beatriz Felipe-García, Albacete, Nilda Sánchez, Diego González-Aguilera & Javier Gomez-Lahoz, Salamanca, Spain

Keywords: ADS40, radiometric calibration, test field, aerial images

Summary: This paper deals with the vicarious calibration of a linear array sensor, ADS40, and the comparison and analysis with the laboratory calibration provided by the manufacturer. For this purpose, a specific test field located in the city of Ávila (Spain) was designed for the two different flights performed during the 8th and 9th of April, 2010 with a spatial resolution of 10 cm and 25 cm, respectively. In addition, a reflectance measurement campaign was performed during the flight using an ASD FieldSpec 3 Hi-Res spectroradiometer, observing a group of 24 targets which included natural and artificial surfaces. Two study cases were considered: with and without atmospheric correction. In order to take into account atmospheric effects, the reflectance method based on the radiative transfer model was applied to establish the relationship between sensor and ground magnitudes. Compared to the calibration results provided by the manufacturer, the results of the vicarious calibration are better. In particular, the root-mean-square grey value error (RMSE) reached through the vicarious calibration in the different reference surfaces is lower than for the laboratory calibration, which considers always the same atmospheric model. Finally, analyzing the stability of the calibration and its relation with the flying height, the calibration of the lower flight (10 cm) was extrapolated to the 25 cm flight, obtaining similar accuracies.

Zusammenfassung: Dieser Artikel beschäftigt sich mit der Feldkalibrierung (vicarious calibration) der Luftbildzeilenkamera ADS40 und vergleicht die Ergebnisse mit denen der vom Hersteller gelieferten Laborkalibrierung. Zu diesem Zweck wurde in Ávila (Spanien) ein Testfeld angelegt, das am 8. und 9. April 2010 in zwei verschiedenen Höhen so überflogen wurde, dass eine Bodenauflösung (Ground Sample Distance, GSD) von 10 cm bzw. 25 cm entstand. Gleichzeitig mit der Überfliegung wurden 24 radiometrische Referenzflächen, die sowohl natürliche als auch künstliche Oberflächen besaßen, vermessen. Dabei kam das Feldspektrometer FieldSpec 3 Hi-Res der Firma ASD zum Einsatz. Die aufgenommenen Kanäle wurden sowohl unter Berücksichtigung eines Atmosphärenmodells als auch ohne ein solches ausgewertet. Als Atmosphärenmodell wurde das 6S (Second simulation of the satellite signal in the solar spectrum), Version 1.1, verwendet. Die Genauigkeit der Feldkalibrierung war etwas höher als die der Laborkalibrierung des Herstellers. Das zeigte sich insbesondere bei unterschiedlichen Oberflächenmaterialien. Ein weiteres Ergebnis war, dass die Ergebnisse der Kalibrierung des niedrigeren Fluges (GSD 10 cm) mit nur einem geringen Genauigkeitsverlust auf die größere Flughöhe (GSD 25 cm) extrapoliert werden konnten.

1 Introduction

Nowadays, the exploitation of physical quantities is still not widespread in thematic applications of aerial images captured by photogrammetric techniques. Nevertheless, as long as the sensor is properly calibrated, it is possible to obtain absolute values of radiance or reflectance, and therefore, these sensors can be used for multispectral classification, vegetation rate calculation and estimation of biophysical variables, providing the advantage over the satellite systems of offering high resolution images and allowing a data acquisition depending

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0139 1432-8364/12/0139 $ 4.00

on user demand (Honkavaara & Markelin 2007).

In 2000, Leica Geosystems introduced the photogrammetric scanning sensor ADS40 (Sandau et al. 2000). This camera is innovative in the field of photogrammetry, not only due to its linear sensor characteristics, but also because of the camera manufacturer provides information about the laboratory radiometric calibration (Beisl et al. 2006), and because the application of some atmospheric correction algorithms and an empirical bidirectional reflectance distribution function (BRDF) correction are part of the photogrammetric workflow provided by the manufacturer's software (Beisl et al. 2008).

There are various radiometric calibration methods in remote sensing (de Vries et al. 2007, Honkavaara et al. 2009, Kaasalainen et al. 2009), but the basic principles are the same (Wagner 2010). While the radiometric processing chain is popular for remote sensing systems, those approaches are not directly applicable in photogrammetric processing due to special features of photogrammetric data acquisition (Read & Graham 2002). In remote sensing, here are three types of radiometric calibration of optical sensors (Dinguirard & Slater 1999, Honkavaara et al. 2009, Wagner 2010): pre-flight (at laboratory with Lambertian or spherical reflectors), on board (taking regular measurements of stable targets) and vicarious (based on object pseudo-invariant reflectance).

The radiometric processing methods of photogrammetric imagery are under development, and several organizations are currently developing radiometric block adjustment software (Honkavaara et al. 2009), specifically for photogrammetric sensors. In a vicarious calibration and validation, the test site performance becomes relevant and must be assessed, although optimal construction of permanent radiometric test sites for high resolution airborne imaging systems is an unresolved issue (Honkavaara et al. 2010). Ideally, the field targets used in radiometric calibration either for satellite or airborne sensors should accomplish a set of characteristics (Teillet et al. 2007, Honkavaara et al. 2010). Specially for the latter ones, the requirements related to stability, location, terrain height and climate/atmospheric properties are the most difficult to realize, i.e., uniformity, sufficiently large footprint, flatness, high signal-to-noise ratio, Lambertian reflectance, as well as minimizing aerosols, water vapour and clouds. Natural and/or anthropogenic targets are used frequently, i.e., rocks or gravel parcels, concrete areas (which can also be painted), sand etc. Specific artificial targets are also common, i.e., transportable tarpaulins or carpets etc.

The authors of this paper conducted some preliminary experiments for radiometric calibration of aerial images with the Intergraph DMC frame camera, analyzing the incorporation of radiometric transfer models into the conventional photogrammetric process (Hernández-López et al. 2011). Recently, very interesting studies on the validation of the radiometric processing chain (Markelin et al. 2010) and the reflectance calibration of the ADS40 were performed (Beisl & Adiguezel 2010). Some of the products processed with the Leica software are evaluated by using ground reflectance measurements as a reference. The results show a reasonably good agreement between the calibration results and the reference.

This article presents an experiment carried out in Ávila (Spain) with the ADS40 sensor. The experiment compared the vicarious calibration of two different flights (10 cm and 25 cm GSD, ground sample distance) with the laboratory calibration provided by the manufacturer: with and without atmospheric correction. A series of check surfaces with different properties have been established for the validation process. During the whole study, the analysis was always based on the Level 0 (L0, raw data) images in nadir direction, without compression and without any geometric or radiometric processing.

This study advances in the research directions raised by the EuroSDR project "Radiometric aspects of digital photogrammetry images" (Honkavaara et al. 2011), which focused on the convenience of performing vicarious calibrations. The main scientific contribution of this work is that remote sensing methodologies for radiometric calibration are adapted, extrapolated and applied to a photogrammetric linear array sensor. Methodologically this implies the application of a rigorous photogrammetric model of the observation geometry to obtain the geometrical parameters required for atmospheric correction. We also carry out an experimental validation of our method, comparing the vicarious calibration to the laboratory calibration and assessing the influence of atmospheric correction (second simulation of the satellite signal in the solar spectrum (6S) model) on the results.

2 Materials and Methods

2.1 *Photogrammetric linear Sensor ADS40-SH52*

The sensor used was the Leica ADS40-SH52, with 62.7 mm focal length, 6.5 µm pixel size, a 12,000 pixels array, and a field of view of 64°. It has four panchromatic CCDs: one oriented forward by 27°, one oriented backwards by 16°, and two nadir views at 2°. It also has four multispectral CCDs: red (R), green (G), blue (B) and near infrared (NIR), providing views that are 2° off nadir. The bandwidths are 465–680 nm for the panchromatic channel, 428–492 nm for the blue band, 533–587 nm for the green band, 608–662 nm for the red band, and 833–887 nm for the infrared band, without overlap between bands (Casella et al. 2008, Markelin et al. 2010).

The camera was calibrated in November 2007 in a laboratory with a uniform white target and an integrating sphere. The calibration parameters are available from Leica on request (Beisl et al. 2006).

2.2 *Photogrammetric Flights*

The city of Ávila, located 100 km from Madrid, was chosen as a test area. In this area, the conditions for photogrammetric flights are favourable due the cleanness of the air in this region. Furthermore, the area comprises an urban area, but also the surroundings including agricultural vegetation, so it has a large size and radiometric range.

The images used in the calibration process were captured on April 8th, 2010 with an excellent illumination and a clean atmosphere. The flight was conducted at 1000 m flying height, resulting in a GSD of 10 cm, distributed in 7 strips in East-West direction and one in North-South over the test field. The area covered was approximately 22 km^2. A flight over the same area with a GSD of 25 cm was performed at the next day, 9th of April, with a flying height of 2500 m, distributed in 10 strips in East-West and another one in North-South direction. The covered area is 75 km^2 (Fig. 1). We worked with images in Level 0 (L0, raw data) of RGB and NIR bands, all of them in the close nadir direction, with 12-bit radiometric resolution.

2.3 *Radiometric Campaign*

The calibration zone was equipped with seven portable reflectance tarpaulins. Six tarpaulins had different spot colours: black, white, blue, red, green and gray, with a size of 5 × 5 m^2. The seventh was a 6-step grey-scale tarpaulin

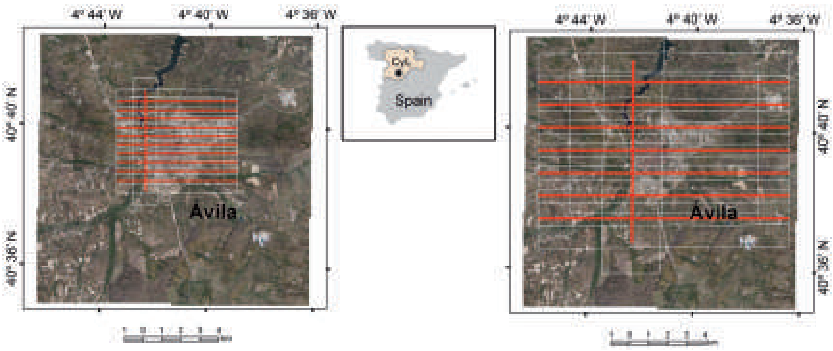


Fig. 1: Flight patterns: GSD of 10 cm on left and of 25 cm on right with control strip in N-S direction in both flights, CyL = Castilla y León.

of 5 × 1 m² (Fig. 2). These tarpaulins (12 targets) were used as control surfaces to perform the calibration process. The number of spectrometer measurements was 5 in each big tarpaulin and 10 in every step of grey-scale tarpaulin.

In addition, 12 check areas were selected for validation (Fig. 3). Three check areas were natural surfaces, covered by sand, grass and a track. Seven check areas corresponded to artificial surfaces of different colours, namely green, white, garnet and the colour of football soil, and to different materials, namely asphalt, concrete soil and cement. The last two check areas corresponded to paved ground and bushes. Their homogeneity was lower than the homogeneity of the other check areas, but their presence is quite common in the study area. All the 12 check areas were used in the validation process.

In order to acquire spectral measurements in the field, an ASD FieldSpec3 spectroradiometer was used as a remote detector of the radiant flux coming from the surface in the visible and near infrared ranges. This model provides a spectral range of 350–2500 nm, a data sampling time of 0.1 seconds per spectrum, and a maximum spectral resolution of 3 nm (±1.0 nm wavelength accuracy) with a maximum noise equivalent radiance of 1.4 10^{-9} $W{\cdot}cm^{-2}{\cdot}sr^{-1}$. The optical fiber was mounted with a 25° field of view registering. In the check areas we defined between one and five check targets, each corresponding to a small patch of approximately 1 × 1 m². In each observation cycle, 120 spectra were measured over each target and then averaged to get the

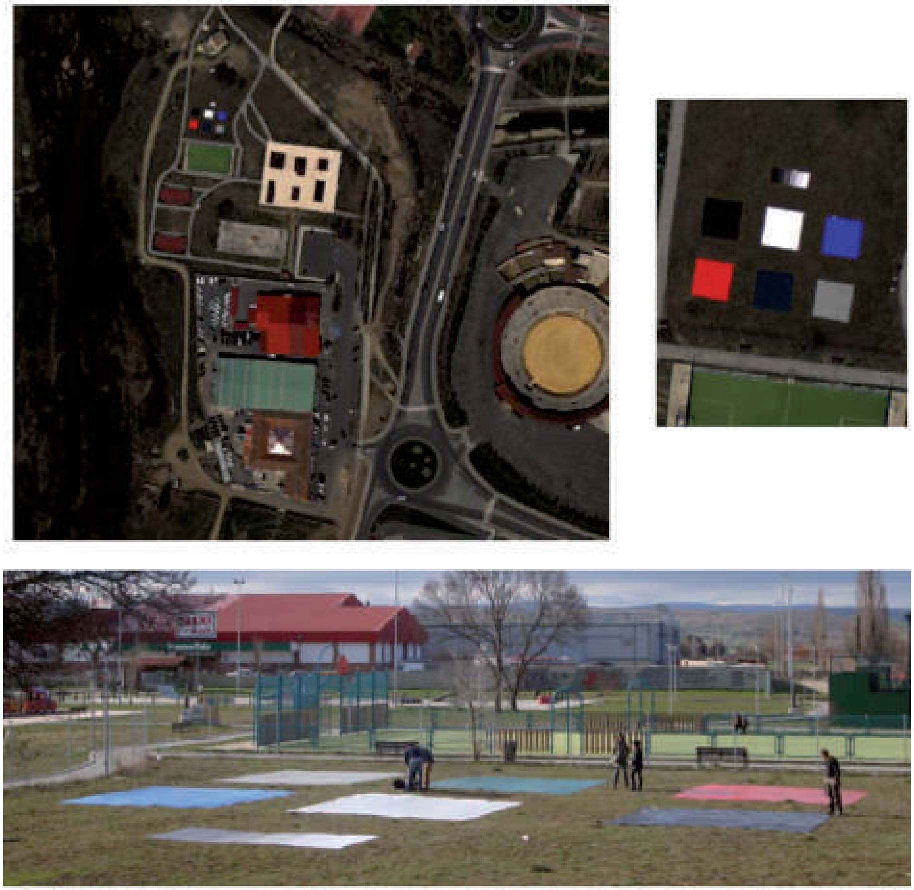


Fig. 2: Test field: Calibration surfaces.

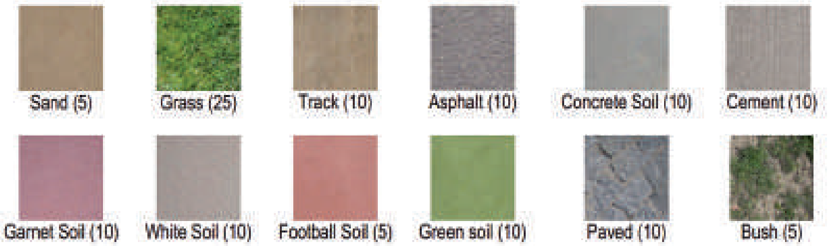


Fig. 3: Examples for targets (1 × 1 m^2) in the check areas and the number of reflectance values derived from spectrometer measurements per check area between brackets. As there are five measurements per target, the number of targets per check area is the number of measurements divided by five. For instance, there was one target over sand, but five targets over grass.

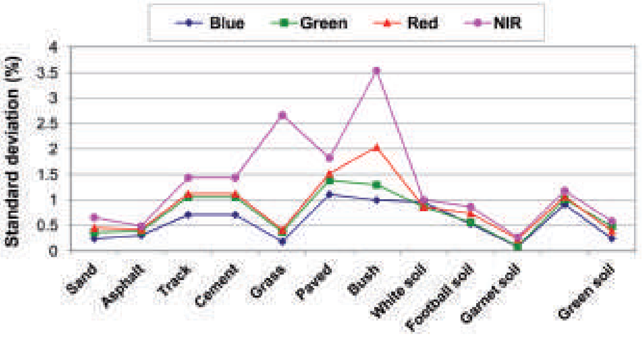


Fig. 4: Standard deviation of measurements (reflectances in %) on the check surfaces.

final spectra. Before each target measurement, a reference measurement was made with a calibrated white reference standard (Spectralon Labsphere ®). The spectra were measured in absolute radiance mode. Afterwards, the radiances were scaled with the reference measurements to produce the reflectance values per target. Finally, these target reflectances were weighted with the ADS40 channel spectral sensitivities to get the reflectances per MS channel. This procedure was repeated five times per target to obtain five representative reflectance values per target. The radiometric samples were georeferenced by determining the centres of the respective targets using a dual-frequency GNSS receiver. The accuracy (RMSE) of the target coordinates was better than 0.1 m. Fig. 3 shows examples for targets (1 × 1 m^2) from each of the check areas and the number of reflectance values per check area used for evaluation. As there were five reflectance values per target, the number of targets per check area is equal to the number of reflectance values divided by five. The overall number of targets in all check areas was 24, and the total number of reflectances measured in the check areas was 120.

Fig. 4 shows the standard deviation of the measurements for each check area. We can observe the highest values in the check area covered by vegetation (grass and bush) in the NIR band. All the spectral measurements per target were used independently for the calibration procedure (in case of the tarpaulins) and for evaluation (targets in check areas). All the targets (calibration or check surfaces) are visible in at least 5 strips. To obtain the corresponding colour vectors in the images, for each image a target appeared in we determined the average digital number per band in a small square window. The size of this window was 7 × 7 pixels in the images with 10 cm GSD (calibration and validation) and 3 × 3 pixels in the images with 25 cm GSD.

2.4 *Radiometric Calibration*

Radiative transfer model

The equation of radiative transfer establishes the relation between reflectances measured on the surface and the apparent reflectance at the sensor, taking into account the processes of atmospheric dispersion and absorption (Vermote et al. 2006).

$$\rho_s(\theta_i,\theta_v,\Phi,z_t,z_s)=\rho_a(\theta_i,\theta_v,\Phi,z_s) + \frac{\rho_t}{1-S(z_t,z_s)\rho_t}\tau_\downarrow(\theta_i,z_t)\tau_\uparrow(\theta_v,z_s)\tau_g(\theta_i,\theta_v,z_t,z_s) \tag{1}$$

In (1), ρ_s is the apparent reflectance at the sensor, ρ_a is the intrinsic reflectance of the atmosphere, S is the atmospheric albedo, τ_g is the global transmittance due to molecular absorption, $\tau_\downarrow$ is the descending transmittance, $\tau^\uparrow$ is the upward transmittance, ρ_t is the target reflectance measured at the surface, θ_i is the zenith angle of the illumination, θ_v is the zenith angle of observation, Φ is the azimuth angle between the plane of observation and the illumination, z_t is the elevation of the ground point and z_s is the altitude of the sensor.

To obtain the atmospheric parameter values necessary to calculate reflectances at the sensor (ρ_a, E_d, $S,\tau_\downarrow$, $\tau^\uparrow$, τ_g) using (1), the model 6S (Vermote et al. 2006) can be used, which requires information about the geometry of illumination and observation (explained in the next section) for each of the target and band image involved in the process. So, θ_i, θ_v, Φ, z_t and z_s will be the input geometric parameters in the model 6S.

If no such atmospheric correction is applied, the sensor and ground reflectance are considered to be equal, in which case (1) is simplified as follows:

$$\rho_s=\rho_t \tag{2}$$

The apparent reflectance of the sensor ρ_s can also be expressed as a function of the observed digital number (DN) recorded by the sensor (Vermote et al. 2006):

$$\rho_s=\frac{\pi L_s}{E_d\cdot\cos\theta_i}=\frac{\pi(c_0+c_1\cdot DN)}{E_d\cdot\cos\theta_i}$$
$$\to c_0+c_1\cdot DN=\frac{E_d\cdot\cos\theta_i}{\pi}\rho_s \tag{3}$$

In (3), L_s is the radiance at the sensor, E_d is the exoatmospheric solar irradiance at the top of the atmosphere (TOA), c_0 (offset) and c_1 (gain) are the calibration coefficients and ρ_s can be obtained either by the 6S model in (1) or by (2).

For the ADS40 sensor, the linear model suggested by the manufacturer is set up in a way that the offset c_0 is corrected by the system during its registration, so it can be set to zero (Beisl 2006). This aspect was checked by the authors. The solution obtained with determining an unknown offset parameter did not result in an improvement, so that a calibration model with only one parameter (the gain c_1) was used in this study. In the calibration process, the reflectance measurements for each target can be transformed to the images via (1) or (2), where they are related to the digital numbers measured in the corresponding image windows and the radiometric calibration coefficient c_1 via (3). Each observed digital number in an image results in one observation equation. This over-determined system is solved by least squares methods (Hernández-López et al. 2011).

In this study, the vicarious calibration is named case '1', referring to the number of coefficients used (only gain) and the calibration of the manufacturer is named case 'M'. If the atmospheric correction is applied the letter 'a' is added. So, case '1' refers to the vicarious calibration without atmospheric correction, case '1a' to the vicarious calibration with atmospheric correction, case 'M' to the calibration provided by the manufacturer without atmospheric correction and case 'Ma' to the calibration provided by the manufacturer with atmospheric correction. In this work, we did not consider the impact of errors in the atmospheric parameters in (1).

Geometry of illumination and observation

To carry out the vicarious calibration with high resolution photogrammetric images it is recommendable to know the three-dimensional geometry of the scene with the highest accuracy possible if atmospheric correction is to be carried out (Hernández-López et al. 2011).

It is not required in the case without atmospheric correction.

The geometry of the scene for each target is defined by the position of the sun, target and sensor.

The GNSS field measurements and the data provided by the flight company are needed to find the three-dimensional geometry of observation, i.e. the relative position between target and sensor. We determined the exterior orientation parameters by aerial triangulation of the L0 images; note that there is an individual set of orientation parameters for each scanline and band: GPS time, position of projection centre (X,Y,Z) and attitude (omega, phi and kappa). In this study, the positions of the (calibration and check) targets were rigorously georeferenced on the ground using a dual-frequency GNSS receiver and back-projected to the L0 images using the orientation parameters. Back-projection requires an iterative process and also a proper handling of structure and image information. To solve this problem, an application based on the ADS Info kit provided by Leica Geosystems was developed (Tempelmann et al. 2005). As a result, we obtain the GPS time for the time each target was imaged and consequently also the projection centre coordinates, which are used to calculate the azimuth and zenith angles of observation between target and sensor required to determine the atmospheric correction parameters in (1). The back-projected points are used to define the $n \times n$ windows to determine the DN values. We do not consider the obliqueness of the terrain since all our targets are in horizontal terrain.

The solar position algorithm (SPA) library (Reda & Andreas 2008) was used to calculate the illumination geometry, relative position between target and sun, expressed as the geodetic azimuth and the zenith angle at the moment of the data acquisition using GPS time and target position (geodetic latitude, geodetic longitude and orthometric height).

Atmospheric characterization

With regard to the atmospheric parameters, the values of aerosol optical depth (AOD) were taken from AERONET (AERONET 2010). Due to the absence of closer stations, the value of AOD was extracted as the average of the values of the two closest stations, Autilla and Cáceres, both at a distance of 100 km from Ávila.

The cleanness of the atmosphere at the day of the flight was remarkable, so that the AOD values were pretty similar (Tab. 1). The "Midlatitude Summer" atmospheric model and the continental aerosol model (Vermote et al. 2006) were used to obtain the atmospheric parameters (ρ_a, S, $\tau_\downarrow$, $\tau^\uparrow$, τ_g) for each of the test field measurements based on the 6S model, using the geometrical parameters derived in the way explained in the previous section. The effects of the quality of the atmospheric parameters on the calibration results were not considered in this study.

Solution

The calculation of the calibration is composed of an overdetermined system based on the calibration targets (Fig. 2), and takes into account that for each target we obtain one corresponding image observation in each image the target is visible in, using the mean digital number (DN) value of window of 7 × 7 pixels. All the reflectances measured at ground level (5–10 per target) are transformed to the images using either (1) (with atmospheric correction) or (2) (no atmospheric correction), where the observation equation is given by (3). In the adjustment, the corrected reflectances are considered as the observations, whereas the digital

Tab. 1: Mean AOD calculated from the closest AERONET stations on April 8th, 2010.

Station	**Long.**	**Lat.**	**AOD**			
			B (at 440 nm)	**G** (at 500 nm)	**R** (at 675 nm)	**NIR** (at 870 nm)
Autilla	−4.603	41.997	0.100	0.084	0.059	0.059
Cáceres	−6.343	39.479	0.104	0.100	0.055	0.045
Mean			**0.102**	**0.092**	**0.057**	**0.052**

numbers were used as constants. Consequently, two solutions were obtained, one without atmospheric correction (case '1') and the other with atmospheric correction (case '1a'). In both cases the overdetermined system has 450 observation equations and one unknown gain parameter c_1 per band. Since some outliers could remain in the measurements, the least squares solution is stabilized with the modified Danish robust method (CASPARY 1987, KRARUP et al. 1980).

2.5 *Validation*

In order to evaluate the vicarious calibration method and compare its effectiveness with the one obtained by the laboratory calibration, we applied all calibration parameters to the image observations in the check targets (Fig. 3) to propagate them into ground reflectances. The ground reflectances thus determined were compared to the ground measurements. The check targets were independent and not used in the calibration process. In this way, for each check measurement and image band, the differences between the ground reflectances $\rho_{calculated}$, obtained by applying the results of the vicarious calibration (with and without atmospheric correction) to the grey levels, and the measured ground reflectances $\rho_{measured}$ were examined. Similarly, the same comparison was done with those reflectances obtained by applying the calibration coefficients provided by the manufacturer. In the cases when atmosphere was considered, the sensor reflectances were transformed into ground reflectances using again the atmospheric characterization provided by the 6S model in (1).

To compare the different cases in a global way, considering all the images and all the targets, the root-mean-square error (RMSE) of the reflectance differences expressed in reflectance units (%) was used:

$$RMSE = \sqrt{\frac{1}{N}\sum\left(\rho_{measured} - \rho_{calculated}\right)^2} \qquad (4)$$

where N is the number of image observations for all check targets.

Furthermore, to analyze each check area (corresponding to several targets of the same land cover type) in an independent way, the following values are determined independently for each check area, considering all the images: the maximal and minimal error, standard deviation, average error, the average absolute error value, and the RMSE.

3 Results

3.1 *Parameters of Flight Calibration, Flight of 10 cm GSD*

In this process all calibration surface tarps were used. Tab. 2 shows the gain coefficient and the RMSE obtained from the vicarious calibration, with atmospheric correction ('1a') and without it ('1'). The calculated values are slightly higher than those provided by the manufacturer in all bands ('M'). The RMSEs were smaller with atmospheric correction, except for the R band.

3.2 *Validation*

Fig. 5 shows the accuracy of these experiments expressed as the RMSE in which all the targets and all the images are considered. Vicarious calibration ('1' and '1a') exhibits errors between 2 and 2.6 % in the RGB bands and a slightly higher one, 3.5 %, in the NIR

Tab. 2: Gain coefficients ($W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1} \cdot DN^{-1}$).

	B		G		R		NIR	
	c_0	RMSE	c_0	RMSE	c_0	RMSE	c_0	RMSE
case '1a'	$5.371 \cdot 10^{-5}$	$9.0 \cdot 10^{-8}$	$3.551 \cdot 10^{-5}$	$9.0 \cdot 10^{-8}$	$3.029 \cdot 10^{-5}$	$9.0 \cdot 10^{-8}$	$3.117 \cdot 10^{-5}$	$7.0 \cdot 10^{-8}$
case '1'	$5.677 \cdot 10^{-5}$	$2.3 \cdot 10^{-7}$	$3.745 \cdot 10^{-5}$	$1.2 \cdot 10^{-7}$	$3.095 \cdot 10^{-5}$	$3.0 \cdot 10^{-8}$	$3.225 \cdot 10^{-5}$	$1.5 \cdot 10^{-7}$
case 'M'	$4.65 \cdot 10^{-5}$	–	$3.39 \cdot 10^{-5}$	–	$2.86 \cdot 10^{-5}$	–	$2.78 \cdot 10^{-5}$	–

Tab. 3: Statistics of the errors of the vicarious calibration (case '1a') using the different check surfaces expressed in reflectances (%), max = maximum error, min = minimum error, std = standard deviation, RMSE = root–mean–square error, ave = average error, ave av = average absolute error value.

Vicarious		Sand	Asphalt	Track	Cement	Grass	Paved	Bush	White Soil	Football Soil	Garnet Soil	Concrete Soil	Green Soil
Blue	max	3.90	3.73	2.73	2.29	3.72	2.99	4.90	2.68	1.20	2.49	1.22	2.06
	min	−6.04	−0.30	−6.19	−8.75	−2.53	−3.54	−1.40	−6.10	−4.35	−0.06	−4.13	−1.05
	std	2.99	1.01	2.44	2.51	1.36	1.59	1.58	2.02	1.60	0.75	1.34	0.83
	RMSE	2.94	2.02	2.45	3.53	1.87	1.61	1.90	2.22	2.08	1.63	2.05	1.32
	ave	0.01	1.75	−0.41	−2.51	1.29	−0.31	1.10	−0.96	−1.37	1.46	−1.56	1.03
	ave av	2.08	1.77	1.76	2.91	1.75	1.31	1.42	1.77	1.60	1.46	1.73	1.22
Green	max	2.21	4.03	4.55	1.75	2.51	3.50	4.25	6.60	1.66	1.07	1.69	2.60
	min	−6.10	−0.26	−7.32	−8.85	−4.64	−2.39	−2.07	−2.14	−4.84	−0.41	−4.31	−1.61
	std	2.37	1.12	2.89	2.71	1.71	1.48	1.64	2.34	1.92	0.39	1.34	0.94
	RMSE	2.43	1.71	2.90	3.31	1.71	1.49	1.72	3.59	2.00	0.62	1.53	1.48
	ave	−0.69	1.30	0.45	−1.93	0.13	0.26	0.63	2.74	−0.68	0.48	−0.77	1.15
	ave av	1.45	1.34	2.46	2.42	1.34	1.23	1.31	3.00	1.31	0.53	1.18	1.35
Red	max	2.68	3.91	5.14	2.24	2.57	11.84	7.36	7.08	1.06	1.29	2.54	1.49
	min	−7.25	−0.71	−7.63	−9.67	−3.84	−2.12	−4.86	−0.98	−5.12	−2.03	−4.01	−1.77
	std	2.78	1.24	3.08	3.12	1.34	3.45	3.11	2.30	1.63	0.96	1.40	0.74
	RMSE	2.95	1.60	3.13	3.80	1.36	3.60	3.07	3.54	2.07	1.02	1.48	0.84
	ave	−1.12	1.01	0.71	−2.21	−0.22	1.15	0.38	2.70	−1.33	−0.35	−0.50	0.41
	ave av	1.90	1.13	2.75	2.79	1.01	2.06	2.61	2.82	1.52	0.87	1.18	0.71
NIR	max	10.70	2.43	4.82	2.21	8.69	2.87	4.24	5.27	1.41	0.62	4.01	1.86
	min	−7.53	−2.04	−9.28	−6.32	−20.08	−3.45	−13.00	−5.90	−5.00	−3.24	−3.01	−2.22
	std	4.92	1.09	3.29	2.14	5.28	1.47	4.81	2.66	1.86	0.93	1.49	1.03
	RMSE	4.94	1.09	3.27	2.38	5.29	1.55	4.96	2.64	1.94	1.41	1.48	1.04
	ave	1.02	0.14	0.03	−1.07	−0.52	−0.54	−1.55	−0.14	−0.67	−1.07	0.15	0.24
	ave av	3.52	0.85	2.71	1.86	3.95	1.27	3.47	2.17	1.52	1.14	1.11	0.88

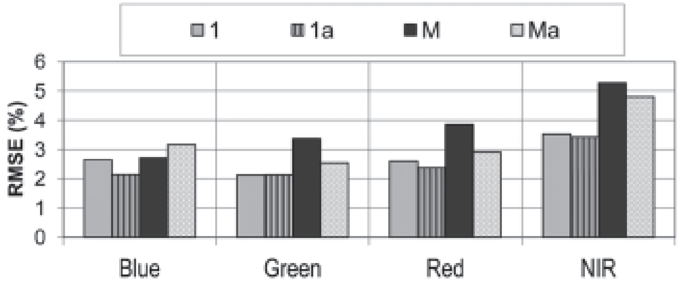


Fig. 5: Comparison of the vicarious calibration derived from the 10 cm flight (cases '1' and '1a') and the laboratory calibration (cases 'M' and 'Ma') applied for the 10 cm flight.

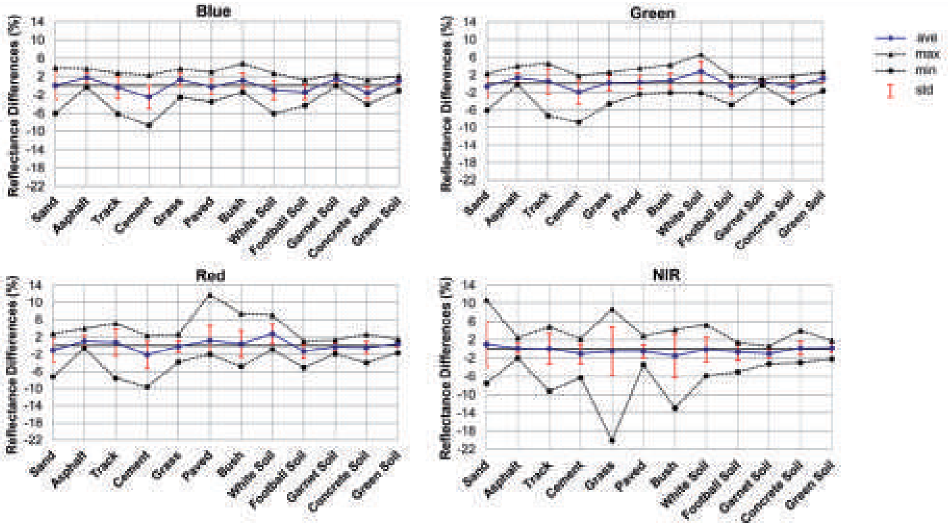


Fig. 6: Maximal and minimal differences between the ground reflectance measurements and those resulting from the vicarious calibration process (case '1a'), with its average and standard deviation (%).

band. The influence of the atmospheric correction in the vicarious calibration was low for all bands, B being the band showing the highest improvement (0.5 %). The RMSEs of the laboratory calibration ('M' and 'Ma') are between 2.6 and 4 % in the RGB band and values close to 5 % in the NIR band. In the case of 'M' the RMSE is higher when the wavelength increases. The results of the calibration coefficients revealed a greater sensitivity to the atmospheric effects ('Ma'), especially in the R and G bands (≈ 1 %). In this case, an increase of the error of 0.5 % for the B band was observed. The RMSEs in reflectances obtained for all the bands after applying the vicarious calibration ('1' and '1a') were smaller than those errors resulting from the laboratory calibration. The NIR band was most significant since it provided an improvement of about 2 %.

Tab. 3 shows the maximum and minimum differences between the ground reflectance measurements at the check surfaces and those resulting from the vicarious calibration with atmospheric correction (case '1a') in absolute terms (expressed in %) and independently for the 12 targets. It shows that the majority of the surface types used as check surfaces may be considered as invariant and suitable for radiometric validation. In general, uniform surfaces that are artificially coloured such as garnet and green provided the best accuracies in the four channels of the spectrum, while more variable surfaces like sand or vegetation had

Tab. 4: Statistics of the errors based on the radiometric manufacturer calibration (case 'Ma') for different check surfaces expressed in reflectances (%), max = maximal error, min = minimal error, std = standard deviation, RMSE = root−mean−square error, ave = average error, ave av = average absolute error value.

ADS40		Sand	Asphalt	Track	Cement	Grass	Paved	Bush	White Soil	Football Soil	Garnet Soil	Concrete Soil	Green Soil
Blue	max	5.52	5.89	5.09	5.85	4.18	6.00	5.93	8.89	3.67	3.32	4.84	3.58
	min	−3.26	1.63	−3.12	−4.48	−1.25	0.37	0.07	0.64	−1.34	1.00	−0.17	0.76
	std	2.62	1.08	2.20	2.35	1.20	1.38	1.47	1.89	1.42	0.68	1.19	0.75
	RMSE	3.31	3.88	3.13	2.77	2.39	3.37	2.79	5.77	1.92	2.45	2.69	2.76
	ave	2.08	3.72	2.25	1.51	2.06	3.09	2.39	5.46	1.32	2.35	2.42	2.66
	ave av	3.08	3.72	2.90	2.30	2.20	3.09	2.39	5.46	1.62	2.35	2.42	2.66
Green	max	3.14	4.76	5.63	3.25	2.76	4.50	4.74	9.10	2.53	1.33	3.29	3.59
	min	−4.86	0.39	−5.95	−7.17	−4.01	−1.01	−1.50	0.49	−3.74	−0.13	−2.82	−0.49
	std	2.28	1.16	2.81	2.67	1.64	1.38	1.62	2.31	1.84	0.39	1.32	0.91
	RMSE	2.26	2.27	3.25	2.66	1.71	1.98	1.97	5.76	1.82	0.84	1.51	2.36
	ave	0.34	1.96	1.66	−0.33	0.50	1.44	1.16	5.29	0.25	0.75	0.75	2.18
	ave av	1.81	1.96	2.91	2.04	1.44	1.62	1.46	5.29	1.50	0.76	1.24	2.21
Red	max	4.03	4.83	6.69	4.22	2.86	12.87	7.84	10.42	2.92	1.76	4.61	2.36
	min	−5.43	0.06	−5.70	−7.43	−3.29	−0.77	−4.03	2.48	−3.03	−1.50	−2.17	−0.77
	std	2.64	1.31	2.98	3.05	1.28	3.30	3.01	2.28	1.55	0.95	1.40	0.71
	RMSE	2.62	2.25	3.83	3.02	1.28	4.15	3.13	6.51	1.64	0.95	1.99	1.49
	ave	0.40	1.84	2.43	−0.08	0.10	2.56	1.06	6.11	0.62	0.15	1.42	1.32
	ave av	2.04	1.84	3.41	2.41	1.02	2.62	2.52	6.11	1.43	0.83	1.67	1.37
NIR	max	13.33	4.22	8.45	6.44	13.57	5.76	6.90	11.56	4.94	1.48	7.68	4.87
	min	−3.18	−0.49	−4.89	−1.85	−13.72	−0.98	−9.05	0.96	−1.01	−2.23	0.01	1.04
	std	4.41	1.18	3.12	2.09	4.94	1.50	4.37	2.54	1.69	0.89	1.71	0.96
	RMSE	6.28	2.08	5.02	3.89	6.44	2.65	4.54	6.87	3.43	0.89	3.85	3.45
	ave	4.55	1.73	3.96	3.29	4.15	2.19	1.51	6.40	3.00	−0.14	3.46	3.32
	ave av	5.34	1.76	4.39	3.51	5.70	2.24	4.04	6.40	3.08	0.72	3.46	3.32

poorer results. These were anisotropic surfaces in which the reflection angles change and the energy that is not gathered by the sensor is lost. That produced a slight increase in its RMSE and standard deviation (std).

The most uniform bands in terms of a similar error were B and G, and the most variable band was NIR (Fig. 6). Here, as expected, the higher error occurred in vegetated samples (bush and grass), because there the NIR band reflectance values and the standard deviation of the measurements were higher.

Next, a comparison with the results obtained using the coefficients of the ADS40 camera and atmospheric correction is done (Tab. 4). In general, for practically all check surfaces the RMSEs of the vicarious calibration are slightly lower than those obtained from the manufacturer calibration coefficients (Fig. 7).

3.3 *Application to 25 cm GSD Flight*

In this section the use of the calibration computed for the 10 cm flight was extrapolated to the 25 cm flight. The main goal was to analyze if it was possible to extrapolate the radiometric results of the vicarious calibration to different ground sample distances and flying heights. The results obtained, considering all the targets and all the images (Fig. 8) show a very similar trend for both flights. In particular, the vicarious calibration exhibited RMSEs lower than 3 % (except for the B band in case ‘1’, where it was 3.5 %). The atmospheric correction in the vicarious calibration improved the results by 1.7 % in the B band, by 0.7 % in the R band and by very small values in the G and NIR bands. The vicarious calibration (‘1’ and ‘1a’) provided higher precisions than those based on the laboratory calibration (‘M’ and ‘Ma’), with a larger variation for the NIR band of about 2 %. Nevertheless, there was an exception for the B band, which provided a different performance. On one hand, without atmospheric correction, the accuracy of the vicarious calibration (‘1’) was slightly lower than the laboratory calibration (‘M’). On the other hand, the application of the atmospheric correction on the laboratory calibration generated a solution slightly worse than the vicarious one (0.7 %).

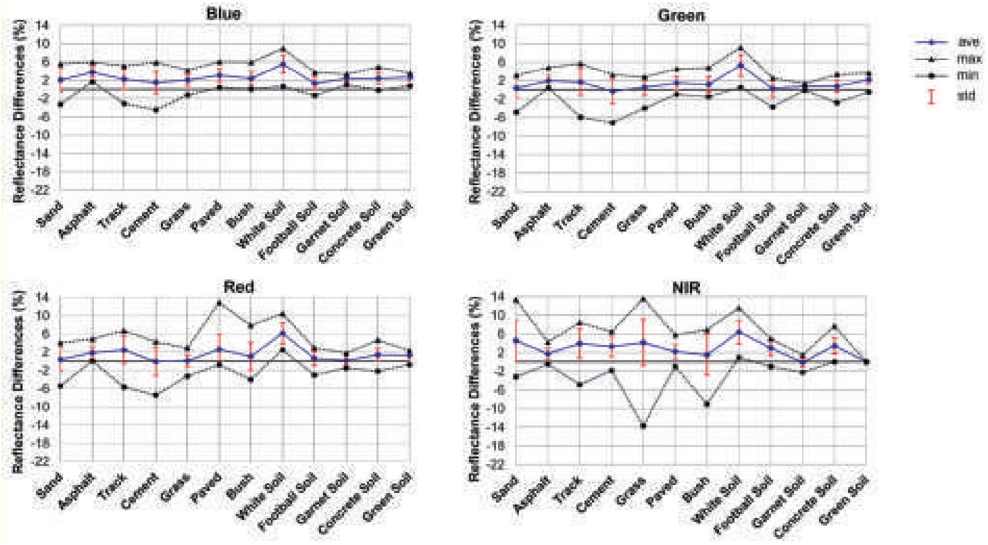


Fig. 7: Maximal and minimal differences between the ground reflectance measurements and the resulting from using the manufacturer calibration coefficients (case ‘Ma’), together with its average and standard deviation (%).

4 Discussion

Other authors such as Honkavaara & Markelin (2007) have empirically analysed the spectral response of different airborne sensors through the use of grey patterns located in the field. In particular, they have examined the spectral response of the ADS40, DMC and UltraCamD systems through the use of a pattern with 8 different levels in grey scale. The conclusion is that these systems present an excellent radiometric quality and can be calibrated. The result of this investigation proved the stability of the radiometric laboratory calibration of the ADS40-SH5.

With regard to the calibration surfaces, we used PVC canvases to cope with the requirements mentioned in other works (Honkavaara et al. 2010) and it has been determined that they are suitable for the correct determination of the radiometric camera calibration, in particular ADS40 sensor (Markelin et al. 2010). However, sensor saturations were detected in the original images in some bands for bright objects, though the images have 12 bit dynamic range.

In our investigation raw sensor data have been used, considering the three-dimensional imaging geometry and a digital elevation model. Processing methods relying on the rectification and the consequent resampling of images could involve alterations to the original radiometric information.

The vicarious calibration is of interest in cases when there is doubt as to the calibration values supplied by the manufacturer or when the calibration needs to be validated. The results are consistent with the recently published papers (Markelin et al. 2010, Beisl 2010), showing that the RMSEs of surfaces reflectances can be better than 5 % in most of the evaluations. The performance varies between the different bands. The performance of the NIR band was lower (3–5 %) than the RGB band in all the cases. The vicarious calibration of the 10 cm flight achieved better results than the laboratory calibration; compared to ground reflectances, RMSEs in the NIR band decreased by about 2 %. In any case, the results achieved for a three-year old laboratory calibration demonstrate the radiometric stability of the sensor. The good weather conditions during the acquisition may have contributed to this result, and thus it would be interesting to repeat the procedure under less favourable conditions. Check surfaces provided large variations in reflectance, depending on the surface and the analyzed band, but they were never above 14 %, reaching even variations lower than 2 % for homogeneous and synthetic surfaces.

The accuracy obtained from the 10 cm calibration flight was preserved in the 25 cm flight. After these experiences, it should be mentioned that the errors increase with the wavelength if the laboratory calibration is used (case 'M'). Likewise, if the atmospheric correction is applied (case 'Ma') the accuracy for the B band gets worse. Therefore, the best solution is the vicarious calibration with atmospheric correction, named here case '1a'. In all the cases, R and G are the most stable bands.

On the other hand, there are some ways for improving the methodology that are worth mentioning: first, concerning the in situ measurements, the atmospheric measurements could be improved. For future work, it would

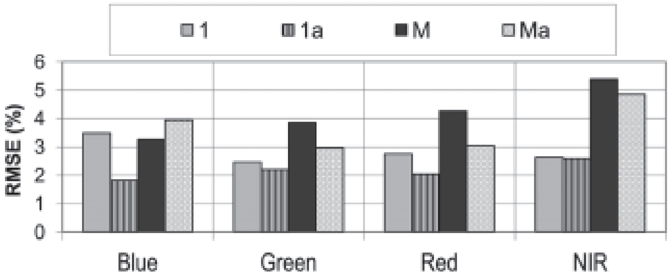


Fig. 8: Comparison of the vicarious calibration derived from the 10 cm flight (cases '1' and '1a') and the laboratory calibration (cases 'M' and 'Ma') applied for the 25 cm flight.

be of interest to analyze the influence of the quality of the meteorological parameters on the final results. Furthermore, it would be interesting to work on the influence of the view direction of the acquisition on the calibration: Although the images were taken approximately in nadir direction (2°) along the flying direction and thus the Lambertian approximation of the surface for the atmospheric correction was applicable, it would be of interest to consider surfaces with an anisotropic behaviour and then analyse bi-directional effects, e.g. applying kernel-driven models (Hernández-López et al. 2011).

5 Conclusions

In the last years there has without doubt been a convergence between photogrammetric and remote sensing methods for radiometric calibration. This study applied a combination of the photogrammetric principles (collinearity constraint, spatial resection, space intersection, …) according to the mathematical model of this linear sensor to establish the geometrical relations required for radiometric calibration (Tempelmann et al. 2005) and the principles of remote sensing, which allow the use of physical models for the image generation, taking into account the problems caused by the atmosphere.

It has been shown that it is possible to perform a radiometric calibration of photogrammetric images through the use of reflectance measurements to obtain and validate the radiometric camera calibration. It has been proven that the solution based on the vicarious calibration with the application of atmospheric correction slightly exceeds the quality of all bands based on the calibration given by the ADS40 sensor manufacturer. However, the ground reflectances obtained from the vicarious calibration and the laboratory calibrations provided by the manufacturer are both acceptable accuracies because the RMSEs achieved by validation for both cases are below 6 %.

It can be concluded that for this investigation, the consideration of atmospheric correction does not affect the results significantly in the imagery acquired at a GSD of 10 cm. Likewise, the use of a vicarious calibration is feasible if it is done shortly before or after the photogrammetric flight, and a different flying height does not significantly affect the results.

The results corroborate the potential of the ADS40 sensor and the need for appropriate laboratory or vicarious calibration to take full advantage of the high radiometric quality of the images.

Acknowledgements

The present contribution has been developed under the Research Project LICAM financed by the "Instituto Tecnológico Agrario de Castilla y León" (ITACYL).

References

AERONET, 2010: Aerosol Robotic Network. – http://aeronet.gsfc.nasa.gov (5.5.2012).

Beisl, U., 2006: Absolute spectroradiometric calibration of the ADS40 sensor. – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XXXVI** (1): on CD-ROM, Paris, France.

Beisl, U., Telaar, J. & Schönermark, M.v., 2008: Atmospheric correction, reflectance calibration and BRDF correction for ADS40 image data. – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XXXVII** (B7): 7–12.

Beisl, U. & Adiguezel, M., 2010: Validation of the reflectance calibration of the ADS40 airborne sensor using ground reflectance measurements. – ISPRS TC VII Symposium – 100 Years ISPRS, IAPRS **XXXVIII** (7B): 80–85, Vienna, Austria.

Casella, V., Franzini, M., Banchini, G. & Gentili, G., 2008: Initial evaluation of the second generation Leica ADS40 Camera. – ISPRS Congress Beijing, Commission I, WG **I/4:** 527–532.

Caspary, W.F., 1987: Concepts of network and deformation analysis. – Monograph **11**, The University of New South Wales, Kensington, N.S.W., Australia.

de Vries, C., Danaher, T., Denham, R., Scarth, P. & Phinn, S., 2007: An operational radiometric calibration procedure for the Landsat sensors based on pseudoinvariant target sites. – Remote Sensing of Environment **107** (3): 414–429.

Dinguirard, M. & Slater, P.N., 1999: Calibration of spatial-multispectral imaging sensors: a review. – Remote Sensing of Environment **68:** 194–205.

Hernández-López, D., Felipe García, B., González Piqueras, J. & Villa Alcázar, G., 2011: An approach to the radiometric aerotriangulation of photogrammetric images. – ISPRS Journal of Photogrammetry and Remote Sensing **66** (2011): 883–893.

Honkavaara, E. & Markelin, L., 2007: Radiometric Performance of Digital Image Data Collection. A Comparison of ADS40/DMC/Ultracam and EmergeDss. – Photogrammetric Week, Stuttgart.

Honkavaara, E., Arbiol, R., Markelin, L., Martinez, L., Cramer, M., Bovet, S., Chandelier, L., Ilves, R., Klonus, S., Marshal, P., Schläpfer, D., Tabor, M., Thom, C. & Veje, N., 2009: Digital Airborne Photogrammetry – A new tool for quantitative remote sensing? A State-of-the-art review on radiometric aspects of digital photogrammetric images. – Remote Sensing **1** (3): 577–605.

Honkavaara, E., Hakala, T., Peltoniemi, J., Suomalainen, J., Ahokas, E. & Markelin, L., 2010: Analysis of properties of reflectance reference targets for permanent radiometric test sites of high resolution airborne imaging systems. – Remote Sensing **2:** 1892–1917.

Honkavaara, E., Arbiol, R., Markelin, L., Martínez, L., Cramer, M., Bovet, S., Chandelier, L., Ilves, R., Klonus, S., Marshal, P., Schläpfer, D., Tabor, M., Thom, C. & Veje, N., 2011: The EuroSDR project Radiometric aspects of digital photogrammetric images. – Results of the empirical phase International Archives of the Photogrammetry. – Remote Sensing and Spatial Information Sciences **38** (Part 4/W19).

Kaasalainen, S., Krooks, A., Kukko, A. & Kaartinen, H., 2009: Radiometric calibration of terrestrial laser scanners with external reference targets. – Remote Sensing **1** (3): 144–158.

Krarup, T., Kubik, K. & Juhl, J., 1980: Gotterdammerung over least squares adjustment. – 14th Congress of the International Society of Photogrammetry **B3:** 369–378, Hamburg.

Markelin, L., Honkavaara, E., Beisl, U. & Korpela, I., 2010: Validation of the radiometric processing chain of the Leica ADS40 airborne photogrammetric sensor. – ISPRS TC VII Symposium – 100 Years ISPRS, IAPRS **XXXVIII** (7A): 145–150, Vienna, Austria.

Read, R.E. & Graham, R.W., 2002: Manual of Air Survey: Primary Data Acquisition. – Caithness, Scotland, UK, Whittles Publishing.

Reda, I. & Andreas, A., 2008: Solar Position Algorithm for Solar Radiation Applications. – Technical report, National Renewable Energy Laboratory.

Sandau, R., Braunecker, B., Driescher, H., Eckardt, A., Hilbert, S., Hutton, J., Kirchhofer, W., Lithopoulos, E., Reulke, R. & Wicki, S., 2000: Design principles of the LH systems ADS40 airborne digital sensor. – International Archives of Photogrammetry and Remote Sensing **33** (1): 258–265.

Teillet, P.M., Barsi, J.A., Chander, G. & Thome, K.J., 2007: Prime candidate earth targets for the post-launch radiometric calibration of satellite sensors. – SPIE International Symposium **6677**, San Diego, CA, USA.

Tempelmann, U., Borner, A., Chaplin, B., Hinsken, L., Mykhalevych, B., Miller, S., Recke, U., Reulke, R. & Uebbing, R., 2005: Photogrammetric Software for the LH Systems ADS40 Airbone Digital Sensor. – Technical report, Leica Geosystems Geospatial Imaging.

Vermote, E.F., Tanré, D., Deuzé, J.L., Herman, M. & Morcrette, J.J., 2006: Second Simulation of the Satellite Signal in the Solar Spectrum – Vector (6SV). – 55, Users Guide Version 3.0, University of Maryland, Greenbelt (MD) and Laboratoire d'Optique Atmosphérique, Lille (France).

Wagner, W., 2010: Radiometric calibration of small-footprint full-waveform airborne laser scanner measurements: Basic physical concepts. – ISPRS Journal of Photogrammetry & Remote Sensing **65:** 505–513.

Addresses of the Authors:

David Hernández-López & Beatriz Felipe-García, Regional Development Institute, University of Castilla-La Mancha, Albacete, Spain, e-mail: {david.hernandez}{beatriz.felipe}@uclm.es

Nilda Sánchez , CIALE, Spanish-Portuguese Centre of Agricultural Research, University of Salamanca, Spain, e-mail: nilda@usal.es

Diego González-Aguilera & Javier Gomez-Lahoz, Department of Cartographic and Land Engineering, University of Salamanca, Spain, e-mail: {daguilera}{fotod}@usal.es

Manuskript eingereicht: November 2011
Angenommen: Juni 2012

PFG 2012 / 5, 0557–0572
Stuttgart, October 2012

Article

Testing the Radiometric Performance of Digital Photogrammetric Images: Vicarious vs. Laboratory Calibration on the Leica ADS40, a Study in Spain

David Hernández-López & Beatriz Felipe-García, Albacete, Nilda Sánchez, Diego González-Aguilera & Javier Gomez-Lahoz, Salamanca, Spain

Keywords: ADS40, radiometric calibration, test field, aerial images

Summary: This paper deals with the vicarious calibration of a linear array sensor, ADS40, and the comparison and analysis with the laboratory calibration provided by the manufacturer. For this purpose, a specific test field located in the city of Ávila (Spain) was designed for the two different flights performed during the 8th and 9th of April, 2010 with a spatial resolution of 10 cm and 25 cm, respectively. In addition, a reflectance measurement campaign was performed during the flight using an ASD FieldSpec 3 Hi-Res spectroradiometer, observing a group of 24 targets which included natural and artificial surfaces. Two study cases were considered: with and without atmospheric correction. In order to take into account atmospheric effects, the reflectance method based on the radiative transfer model was applied to establish the relationship between sensor and ground magnitudes. Compared to the calibration results provided by the manufacturer, the results of the vicarious calibration are better. In particular, the root-mean-square grey value error (RMSE) reached through the vicarious calibration in the different reference surfaces is lower than for the laboratory calibration, which considers always the same atmospheric model. Finally, analyzing the stability of the calibration and its relation with the flying height, the calibration of the lower flight (10 cm) was extrapolated to the 25 cm flight, obtaining similar accuracies.

Zusammenfassung: Dieser Artikel beschäftigt sich mit der Feldkalibrierung (vicarious calibration) der Luftbildzeilenkamera ADS40 und vergleicht die Ergebnisse mit denen der vom Hersteller gelieferten Laborkalibrierung. Zu diesem Zweck wurde in Ávila (Spanien) ein Testfeld angelegt, das am 8. und 9. April 2010 in zwei verschiedenen Höhen so überflogen wurde, dass eine Bodenauflösung (Ground Sample Distance, GSD) von 10 cm bzw. 25 cm entstand. Gleichzeitig mit der Überfliegung wurden 24 radiometrische Referenzflächen, die sowohl natürliche als auch künstliche Oberflächen besaßen, vermessen. Dabei kam das Feldspektrometer FieldSpec 3 Hi-Res der Firma ASD zum Einsatz. Die aufgenommenen Kanäle wurden sowohl unter Berücksichtigung eines Atmosphärenmodells als auch ohne ein solches ausgewertet. Als Atmosphärenmodell wurde das 6S (Second simulation of the satellite signal in the solar spectrum), Version 1.1, verwendet. Die Genauigkeit der Feldkalibrierung war etwas höher als die der Laborkalibrierung des Herstellers. Das zeigte sich insbesondere bei unterschiedlichen Oberflächenmaterialien. Ein weiteres Ergebnis war, dass die Ergebnisse der Kalibrierung des niedrigeren Fluges (GSD 10 cm) mit nur einem geringen Genauigkeitsverlust auf die größere Flughöhe (GSD 25 cm) extrapoliert werden konnten.

1 Introduction

Nowadays, the exploitation of physical quantities is still not widespread in thematic applications of aerial images captured by photogrammetric techniques. Nevertheless, as long as the sensor is properly calibrated, it is possible to obtain absolute values of radiance or reflectance, and therefore, these sensors can be used for multispectral classification, vegetation rate calculation and estimation of biophysical variables, providing the advantage over the satellite systems of offering high resolution images and allowing a data acquisition depending

DOI: 10.1127/1432-8364/2012/0139 1432-8364/12/0139 $ 4.00

on user demand (Honkavaara & Markelin 2007).

In 2000, Leica Geosystems introduced the photogrammetric scanning sensor ADS40 (Sandau et al. 2000). This camera is innovative in the field of photogrammetry, not only due to its linear sensor characteristics, but also because of the camera manufacturer provides information about the laboratory radiometric calibration (Beisl et al. 2006), and because the application of some atmospheric correction algorithms and an empirical bidirectional reflectance distribution function (BRDF) correction are part of the photogrammetric workflow provided by the manufacturer's software (Beisl et al. 2008).

There are various radiometric calibration methods in remote sensing (de Vries et al. 2007, Honkavaara et al. 2009, Kaasalainen et al. 2009), but the basic principles are the same (Wagner 2010). While the radiometric processing chain is popular for remote sensing systems, those approaches are not directly applicable in photogrammetric processing due to special features of photogrammetric data acquisition (Read & Graham 2002). In remote sensing, here are three types of radiometric calibration of optical sensors (Dinguirard & Slater 1999, Honkavaara et al. 2009, Wagner 2010): pre-flight (at laboratory with Lambertian or spherical reflectors), on board (taking regular measurements of stable targets) and vicarious (based on object pseudo-invariant reflectance).

The radiometric processing methods of photogrammetric imagery are under development, and several organizations are currently developing radiometric block adjustment software (Honkavaara et al. 2009), specifically for photogrammetric sensors. In a vicarious calibration and validation, the test site performance becomes relevant and must be assessed, although optimal construction of permanent radiometric test sites for high resolution airborne imaging systems is an unresolved issue (Honkavaara et al. 2010). Ideally, the field targets used in radiometric calibration either for satellite or airborne sensors should accomplish a set of characteristics (Teillet et al. 2007, Honkavaara et al. 2010). Specially for the latter ones, the requirements related to stability, location, terrain height and climate/atmospheric properties are the most difficult to realize, i.e., uniformity, sufficiently large footprint, flatness, high signal-to-noise ratio, Lambertian reflectance, as well as minimizing aerosols, water vapour and clouds. Natural and/or anthropogenic targets are used frequently, i.e., rocks or gravel parcels, concrete areas (which can also be painted), sand etc. Specific artificial targets are also common, i.e., transportable tarpaulins or carpets etc.

The authors of this paper conducted some preliminary experiments for radiometric calibration of aerial images with the Intergraph DMC frame camera, analyzing the incorporation of radiometric transfer models into the conventional photogrammetric process (Hernández-López et al. 2011). Recently, very interesting studies on the validation of the radiometric processing chain (Markelin et al. 2010) and the reflectance calibration of the ADS40 were performed (Beisl & Adiguezel 2010). Some of the products processed with the Leica software are evaluated by using ground reflectance measurements as a reference. The results show a reasonably good agreement between the calibration results and the reference.

This article presents an experiment carried out in Ávila (Spain) with the ADS40 sensor. The experiment compared the vicarious calibration of two different flights (10 cm and 25 cm GSD, ground sample distance) with the laboratory calibration provided by the manufacturer: with and without atmospheric correction. A series of check surfaces with different properties have been established for the validation process. During the whole study, the analysis was always based on the Level 0 (L0, raw data) images in nadir direction, without compression and without any geometric or radiometric processing.

This study advances in the research directions raised by the EuroSDR project "Radiometric aspects of digital photogrammetry images" (Honkavaara et al. 2011), which focused on the convenience of performing vicarious calibrations. The main scientific contribution of this work is that remote sensing methodologies for radiometric calibration are adapted, extrapolated and applied to a photogrammetric linear array sensor. Methodologically this implies the application of a rigorous photogrammetric model of the observation geometry to obtain the geometrical parameters required for atmospheric correction. We also carry out an experimental validation of our method, comparing the vicarious calibration to the laboratory calibration and assessing the influence of atmospheric correction (second simulation of the satellite signal in the solar spectrum (6S) model) on the results.

2 Materials and Methods

2.1 *Photogrammetric linear Sensor ADS40-SH52*

The sensor used was the Leica ADS40-SH52, with 62.7 mm focal length, 6.5 µm pixel size, a 12,000 pixels array, and a field of view of 64°. It has four panchromatic CCDs: one oriented forward by 27°, one oriented backwards by 16°, and two nadir views at 2°. It also has four multispectral CCDs: red (R), green (G), blue (B) and near infrared (NIR), providing views that are 2° off nadir. The bandwidths are 465–680 nm for the panchromatic channel, 428–492 nm for the blue band, 533–587 nm for the green band, 608–662 nm for the red band, and 833–887 nm for the infrared band, without overlap between bands (Casella et al. 2008, Markelin et al. 2010).

The camera was calibrated in November 2007 in a laboratory with a uniform white target and an integrating sphere. The calibration parameters are available from Leica on request (Beisl et al. 2006).

2.2 *Photogrammetric Flights*

The city of Ávila, located 100 km from Madrid, was chosen as a test area. In this area, the conditions for photogrammetric flights are favourable due the cleanness of the air in this region. Furthermore, the area comprises an urban area, but also the surroundings including agricultural vegetation, so it has a large size and radiometric range.

The images used in the calibration process were captured on April 8th, 2010 with an excellent illumination and a clean atmosphere. The flight was conducted at 1000 m flying height, resulting in a GSD of 10 cm, distributed in 7 strips in East-West direction and one in North-South over the test field. The area covered was approximately 22 km². A flight over the same area with a GSD of 25 cm was performed at the next day, 9th of April, with a flying height of 2500 m, distributed in 10 strips in East-West and another one in North-South direction. The covered area is 75 km² (Fig. 1). We worked with images in Level 0 (L0, raw data) of RGB and NIR bands, all of them in the close nadir direction, with 12-bit radiometric resolution.

2.3 *Radiometric Campaign*

The calibration zone was equipped with seven portable reflectance tarpaulins. Six tarpaulins had different spot colours: black, white, blue, red, green and gray, with a size of 5 × 5 m². The seventh was a 6-step grey-scale tarpaulin

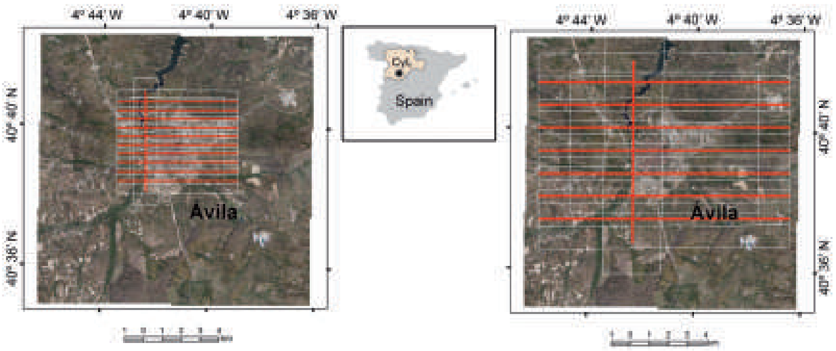


Fig. 1: Flight patterns: GSD of 10 cm on left and of 25 cm on right with control strip in N-S direction in both flights, CyL = Castilla y León.

of 5 × 1 m^2 (Fig. 2). These tarpaulins (12 targets) were used as control surfaces to perform the calibration process. The number of spectrometer measurements was 5 in each big tarpaulin and 10 in every step of grey-scale tarpaulin.

In addition, 12 check areas were selected for validation (Fig. 3). Three check areas were natural surfaces, covered by sand, grass and a track. Seven check areas corresponded to artificial surfaces of different colours, namely green, white, garnet and the colour of football soil, and to different materials, namely asphalt, concrete soil and cement. The last two check areas corresponded to paved ground and bushes. Their homogeneity was lower than the homogeneity of the other check areas, but their presence is quite common in the study area. All the 12 check areas were used in the validation process.

In order to acquire spectral measurements in the field, an ASD FieldSpec3 spectroradiometer was used as a remote detector of the radiant flux coming from the surface in the visible and near infrared ranges. This model provides a spectral range of 350–2500 nm, a data sampling time of 0.1 seconds per spectrum, and a maximum spectral resolution of 3 nm (±1.0 nm wavelength accuracy) with a maximum noise equivalent radiance of 1.4 10^{-9} W·cm^{-2}·sr^{-1}. The optical fiber was mounted with a 25° field of view registering. In the check areas we defined between one and five check targets, each corresponding to a small patch of approximately 1 × 1 m^2. In each observation cycle, 120 spectra were measured over each target and then averaged to get the

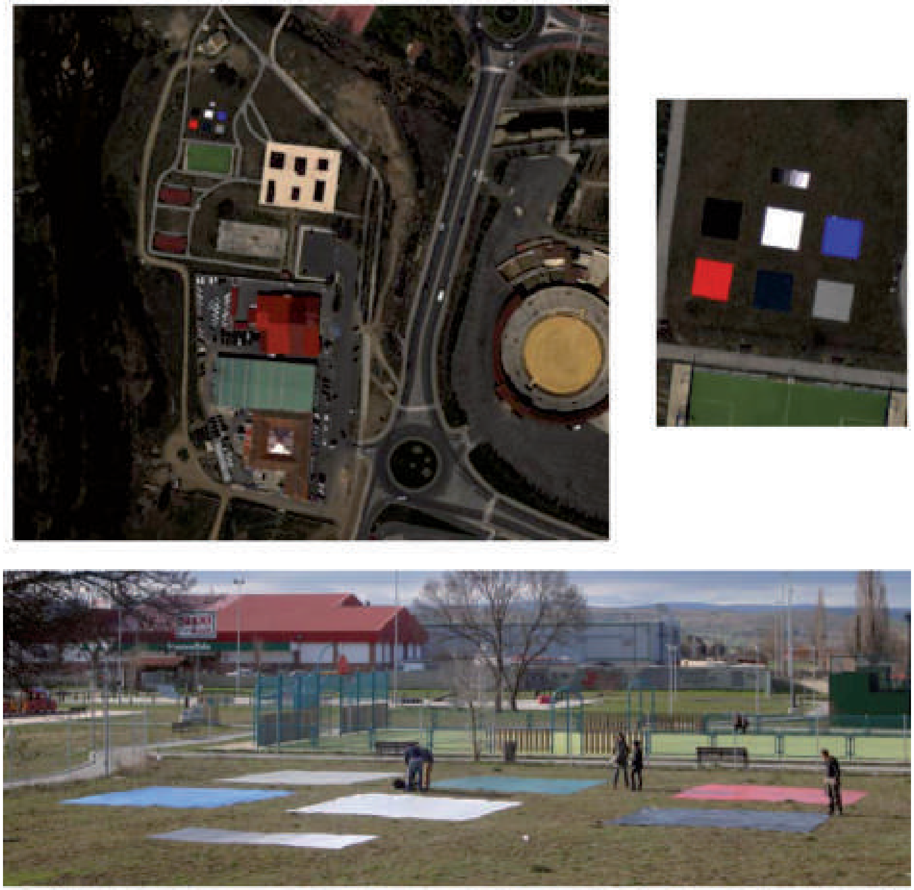


Fig. 2: Test field: Calibration surfaces.

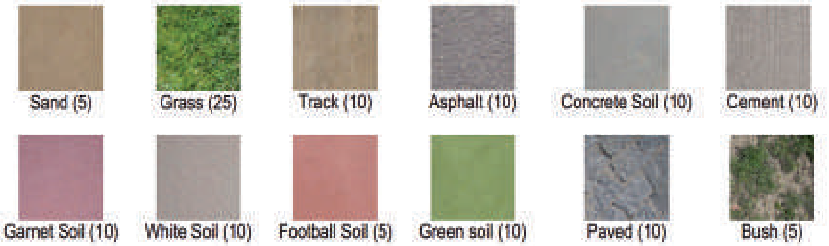


Fig. 3: Examples for targets (1 × 1 m²) in the check areas and the number of reflectance values derived from spectrometer measurements per check area between brackets. As there are five measurements per target, the number of targets per check area is the number of measurements divided by five. For instance, there was one target over sand, but five targets over grass.

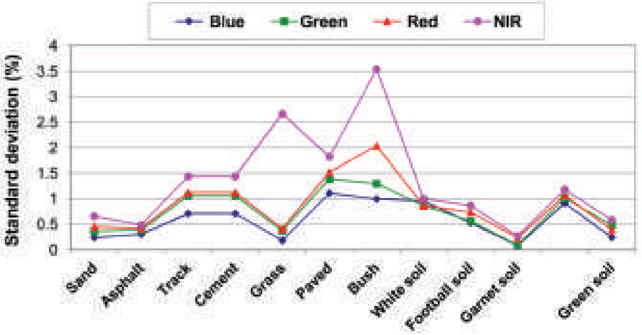


Fig. 4: Standard deviation of measurements (reflectances in %) on the check surfaces.

final spectra. Before each target measurement, a reference measurement was made with a calibrated white reference standard (Spectralon Labsphere ®). The spectra were measured in absolute radiance mode. Afterwards, the radiances were scaled with the reference measurements to produce the reflectance values per target. Finally, these target reflectances were weighted with the ADS40 channel spectral sensitivities to get the reflectances per MS channel. This procedure was repeated five times per target to obtain five representative reflectance values per target. The radiometric samples were georeferenced by determining the centres of the respective targets using a dual-frequency GNSS receiver. The accuracy (RMSE) of the target coordinates was better than 0.1 m. Fig. 3 shows examples for targets (1 × 1 m²) from each of the check areas and the number of reflectance values per check area used for evaluation. As there were five reflectance values per target, the number of targets per check area is equal to the number of reflectance values divided by five. The overall number of targets in all check areas was 24, and the total number of reflectances measured in the check areas was 120.

Fig. 4 shows the standard deviation of the measurements for each check area. We can observe the highest values in the check area covered by vegetation (grass and bush) in the NIR band. All the spectral measurements per target were used independently for the calibration procedure (in case of the tarpaulins) and for evaluation (targets in check areas). All the targets (calibration or check surfaces) are visible in at least 5 strips. To obtain the corresponding colour vectors in the images, for each image a target appeared in we determined the average digital number per band in a small square window. The size of this window was 7 × 7 pixels in the images with 10 cm GSD (calibration and validation) and 3 × 3 pixels in the images with 25 cm GSD.

2.4 *Radiometric Calibration*

Radiative transfer model

The equation of radiative transfer establishes the relation between reflectances measured on the surface and the apparent reflectance at the sensor, taking into account the processes of atmospheric dispersion and absorption (Vermote et al. 2006).

$$\rho_s(\theta_i,\theta_v,\Phi,z_t,z_s)=\rho_a(\theta_i,\theta_v,\Phi,z_s)$$
$$+\frac{\rho_t}{1-S(z_t,z_s)\rho_t}\tau_{\downarrow}(\theta_i,z_t)\tau_{\uparrow}(\theta_v,z_s)\tau_g(\theta_i,\theta_v,z_t,z_s) \tag{1}$$

In (1), ρ_s is the apparent reflectance at the sensor, ρ_a is the intrinsic reflectance of the atmosphere, S is the atmospheric albedo, τ_g is the global transmittance due to molecular absorption, $\tau_{\downarrow}$ is the descending transmittance, $\tau^{\uparrow}$ is the upward transmittance, ρ_t is the target reflectance measured at the surface, θ_i is the zenith angle of the illumination, θ_v is the zenith angle of observation, Φ is the azimuth angle between the plane of observation and the illumination, z_t is the elevation of the ground point and z_s is the altitude of the sensor.

To obtain the atmospheric parameter values necessary to calculate reflectances at the sensor (ρ_a, E_d, S,$\tau_{\downarrow}$, $\tau^{\uparrow}$, τ_g) using (1), the model 6S (Vermote et al. 2006) can be used, which requires information about the geometry of illumination and observation (explained in the next section) for each of the target and band image involved in the process. So, θ_i, θ_v, Φ, z_t and z_s will be the input geometric parameters in the model 6S.

If no such atmospheric correction is applied, the sensor and ground reflectance are considered to be equal, in which case (1) is simplified as follows:

$$\rho_s=\rho_t \tag{2}$$

The apparent reflectance of the sensor ρ_s can also be expressed as a function of the observed digital number (*DN*) recorded by the sensor (Vermote et al. 2006):

$$\rho_s=\frac{\pi L_s}{E_d\cdot\cos\theta_i}=\frac{\pi(c_0+c_1\cdot DN)}{E_d\cdot\cos\theta_i}$$
$$\rightarrow c_0+c_1\cdot DN=\frac{E_d\cdot\cos\theta_i}{\pi}\rho_s \tag{3}$$

In (3), L_s is the radiance at the sensor, E_d is the exoatmospheric solar irradiance at the top of the atmosphere (TOA), c_0 (offset) and c_1 (gain) are the calibration coefficients and ρ_s can be obtained either by the 6S model in (1) or by (2).

For the ADS40 sensor, the linear model suggested by the manufacturer is set up in a way that the offset c_0 is corrected by the system during its registration, so it can be set to zero (Beisl 2006). This aspect was checked by the authors. The solution obtained with determining an unknown offset parameter did not result in an improvement, so that a calibration model with only one parameter (the gain c_1) was used in this study. In the calibration process, the reflectance measurements for each target can be transformed to the images via (1) or (2), where they are related to the digital numbers measured in the corresponding image windows and the radiometric calibration coefficient c_1 via (3). Each observed digital number in an image results in one observation equation. This over-determined system is solved by least squares methods (Hernández-López et al. 2011).

In this study, the vicarious calibration is named case '1', referring to the number of coefficients used (only gain) and the calibration of the manufacturer is named case 'M'. If the atmospheric correction is applied the letter 'a' is added. So, case '1' refers to the vicarious calibration without atmospheric correction, case '1a' to the vicarious calibration with atmospheric correction, case 'M' to the calibration provided by the manufacturer without atmospheric correction and case 'Ma' to the calibration provided by the manufacturer with atmospheric correction. In this work, we did not consider the impact of errors in the atmospheric parameters in (1).

Geometry of illumination and observation

To carry out the vicarious calibration with high resolution photogrammetric images it is recommendable to know the three-dimensional geometry of the scene with the highest accuracy possible if atmospheric correction is to be carried out (Hernández-López et al. 2011).

It is not required in the case without atmospheric correction.

The geometry of the scene for each target is defined by the position of the sun, target and sensor.

The GNSS field measurements and the data provided by the flight company are needed to find the three-dimensional geometry of observation, i.e. the relative position between target and sensor. We determined the exterior orientation parameters by aerial triangulation of the L0 images; note that there is an individual set of orientation parameters for each scanline and band: GPS time, position of projection centre (X,Y,Z) and attitude (omega, phi and kappa). In this study, the positions of the (calibration and check) targets were rigorously georeferenced on the ground using a dual-frequency GNSS receiver and back-projected to the L0 images using the orientation parameters. Back-projection requires an iterative process and also a proper handling of structure and image information. To solve this problem, an application based on the ADS Info kit provided by Leica Geosystems was developed (Tempelmann et al. 2005). As a result, we obtain the GPS time for the time each target was imaged and consequently also the projection centre coordinates, which are used to calculate the azimuth and zenith angles of observation between target and sensor required to determine the atmospheric correction parameters in (1). The back-projected points are used to define the $n \times n$ windows to determine the DN values. We do not consider the obliqueness of the terrain since all our targets are in horizontal terrain.

The solar position algorithm (SPA) library (Reda & Andreas 2008) was used to calculate the illumination geometry, relative position between target and sun, expressed as the geodetic azimuth and the zenith angle at the moment of the data acquisition using GPS time and target position (geodetic latitude, geodetic longitude and orthometric height).

Atmospheric characterization

With regard to the atmospheric parameters, the values of aerosol optical depth (AOD) were taken from AERONET (AERONET 2010). Due to the absence of closer stations, the value of AOD was extracted as the average of the values of the two closest stations, Autilla and Cáceres, both at a distance of 100 km from Ávila.

The cleanness of the atmosphere at the day of the flight was remarkable, so that the AOD values were pretty similar (Tab. 1). The "Midlatitude Summer" atmospheric model and the continental aerosol model (Vermote et al. 2006) were used to obtain the atmospheric parameters (ρ_a, S, $\tau_\downarrow$, $\tau^\uparrow$, τ_g) for each of the test field measurements based on the 6S model, using the geometrical parameters derived in the way explained in the previous section. The effects of the quality of the atmospheric parameters on the calibration results were not considered in this study.

Solution

The calculation of the calibration is composed of an overdetermined system based on the calibration targets (Fig. 2), and takes into account that for each target we obtain one corresponding image observation in each image the target is visible in, using the mean digital number (DN) value of window of 7 × 7 pixels. All the reflectances measured at ground level (5–10 per target) are transformed to the images using either (1) (with atmospheric correction) or (2) (no atmospheric correction), where the observation equation is given by (3). In the adjustment, the corrected reflectances are considered as the observations, whereas the digital

Tab. 1: Mean AOD calculated from the closest AERONET stations on April 8th, 2010.

Station	Long.	Lat.	AOD			
			B (at 440 nm)	**G** (at 500 nm)	**R** (at 675 nm)	**NIR** (at 870 nm)
Autilla	−4.603	41.997	0.100	0.084	0.059	0.059
Cáceres	−6.343	39.479	0.104	0.100	0.055	0.045
Mean			**0.102**	**0.092**	**0.057**	**0.052**

numbers were used as constants. Consequently, two solutions were obtained, one without atmospheric correction (case '1') and the other with atmospheric correction (case '1a'). In both cases the overdetermined system has 450 observation equations and one unknown gain parameter c_1 per band. Since some outliers could remain in the measurements, the least squares solution is stabilized with the modified Danish robust method (Caspary 1987, Krarup et al. 1980).

2.5 *Validation*

In order to evaluate the vicarious calibration method and compare its effectiveness with the one obtained by the laboratory calibration, we applied all calibration parameters to the image observations in the check targets (Fig. 3) to propagate them into ground reflectances. The ground reflectances thus determined were compared to the ground measurements. The check targets were independent and not used in the calibration process. In this way, for each check measurement and image band, the differences between the ground reflectances $\rho_{calculated}$, obtained by applying the results of the vicarious calibration (with and without atmospheric correction) to the grey levels, and the measured ground reflectances $\rho_{measured}$ were examined. Similarly, the same comparison was done with those reflectances obtained by applying the calibration coefficients provided by the manufacturer. In the cases when atmosphere was considered, the sensor reflectances were transformed into ground reflectances using again the atmospheric characterization provided by the 6S model in (1).

To compare the different cases in a global way, considering all the images and all the targets, the root-mean-square error (RMSE) of the reflectance differences expressed in reflectance units (%) was used:

$$RMSE = \sqrt{\frac{1}{N}\sum\left(\rho_{measured} - \rho_{calculated}\right)^2} \qquad (4)$$

where N is the number of image observations for all check targets.

Furthermore, to analyze each check area (corresponding to several targets of the same land cover type) in an independent way, the following values are determined independently for each check area, considering all the images: the maximal and minimal error, standard deviation, average error, the average absolute error value, and the RMSE.

3 Results

3.1 *Parameters of Flight Calibration, Flight of 10 cm GSD*

In this process all calibration surface tarps were used. Tab. 2 shows the gain coefficient and the RMSE obtained from the vicarious calibration, with atmospheric correction ('1a') and without it ('1'). The calculated values are slightly higher than those provided by the manufacturer in all bands ('M'). The RMSEs were smaller with atmospheric correction, except for the R band.

3.2 *Validation*

Fig. 5 shows the accuracy of these experiments expressed as the RMSE in which all the targets and all the images are considered. Vicarious calibration ('1' and '1a') exhibits errors between 2 and 2.6 % in the RGB bands and a slightly higher one, 3.5 %, in the NIR

Tab. 2: Gain coefficients ($W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1} \cdot DN^{-1}$).

	B		G		R		NIR	
	c_0	RMSE	c_0	RMSE	c_0	RMSE	c_0	RMSE
case '1a'	$5.371 \cdot 10^{-5}$	$9.0 \cdot 10^{-8}$	$3.551 \cdot 10^{-5}$	$9.0 \cdot 10^{-8}$	$3.029 \cdot 10^{-5}$	$9.0 \cdot 10^{-8}$	$3.117 \cdot 10^{-5}$	$7.0 \cdot 10^{-8}$
case '1'	$5.677 \cdot 10^{-5}$	$2.3 \cdot 10^{-7}$	$3.745 \cdot 10^{-5}$	$1.2 \cdot 10^{-7}$	$3.095 \cdot 10^{-5}$	$3.0 \cdot 10^{-8}$	$3.225 \cdot 10^{-5}$	$1.5 \cdot 10^{-7}$
case 'M'	$4.65 \cdot 10^{-5}$	–	$3.39 \cdot 10^{-5}$	–	$2.86 \cdot 10^{-5}$	–	$2.78 \cdot 10^{-5}$	–

Tab. 3: Statistics of the errors of the vicarious calibration (case '1a') using the different check surfaces expressed in reflectances (%), max = maximum error, min = minimum error, std = standard deviation, RMSE = root–mean–square error, ave = average error, ave av = average absolute error value.

Vicarious		Sand	Asphalt	Track	Cement	Grass	Paved	Bush	White Soil	Football Soil	Garnet Soil	Concrete Soil	Green Soil
Blue	max	3.90	3.73	2.73	2.29	3.72	2.99	4.90	2.68	1.20	2.49	1.22	2.06
	min	−6.04	−0.30	−6.19	−8.75	−2.53	−3.54	−1.40	−6.10	−4.35	−0.06	−4.13	−1.05
	std	2.99	1.01	2.44	2.51	1.36	1.59	1.58	2.02	1.60	0.75	1.34	0.83
	RMSE	2.94	2.02	2.45	3.53	1.87	1.61	1.90	2.22	2.08	1.63	2.05	1.32
	ave	0.01	1.75	−0.41	−2.51	1.29	−0.31	1.10	−0.96	−1.37	1.46	−1.56	1.03
	ave av	2.08	1.77	1.76	2.91	1.75	1.31	1.42	1.77	1.60	1.46	1.73	1.22
Green	max	2.21	4.03	4.55	1.75	2.51	3.50	4.25	6.60	1.66	1.07	1.69	2.60
	min	−6.10	−0.26	−7.32	−8.85	−4.64	−2.39	−2.07	−2.14	−4.84	−0.41	−4.31	−1.61
	std	2.37	1.12	2.89	2.71	1.71	1.48	1.64	2.34	1.92	0.39	1.34	0.94
	RMSE	2.43	1.71	2.90	3.31	1.71	1.49	1.72	3.59	2.00	0.62	1.53	1.48
	ave	−0.69	1.30	0.45	−1.93	0.13	0.26	0.63	2.74	−0.68	0.48	−0.77	1.15
	ave av	1.45	1.34	2.46	2.42	1.34	1.23	1.31	3.00	1.31	0.53	1.18	1.35
Red	max	2.68	3.91	5.14	2.24	2.57	11.84	7.36	7.08	1.06	1.29	2.54	1.49
	min	−7.25	−0.71	−7.63	−9.67	−3.84	−2.12	−4.86	−0.98	−5.12	−2.03	−4.01	−1.77
	std	2.78	1.24	3.08	3.12	1.34	3.45	3.11	2.30	1.63	0.96	1.40	0.74
	RMSE	2.95	1.60	3.13	3.80	1.36	3.60	3.07	3.54	2.07	1.02	1.48	0.84
	ave	−1.12	1.01	0.71	−2.21	−0.22	1.15	0.38	2.70	−1.33	−0.35	−0.50	0.41
	ave av	1.90	1.13	2.75	2.79	1.01	2.06	2.61	2.82	1.52	0.87	1.18	0.71
NIR	max	10.70	2.43	4.82	2.21	8.69	2.87	4.24	5.27	1.41	0.62	4.01	1.86
	min	−7.53	−2.04	−9.28	−6.32	−20.08	−3.45	−13.00	−5.90	−5.00	−3.24	−3.01	−2.22
	std	4.92	1.09	3.29	2.14	5.28	1.47	4.81	2.66	1.86	0.93	1.49	1.03
	RMSE	4.94	1.09	3.27	2.38	5.29	1.55	4.96	2.64	1.94	1.41	1.48	1.04
	ave	1.02	0.14	0.03	−1.07	−0.52	−0.54	−1.55	−0.14	−0.67	−1.07	0.15	0.24
	ave av	3.52	0.85	2.71	1.86	3.95	1.27	3.47	2.17	1.52	1.14	1.11	0.88

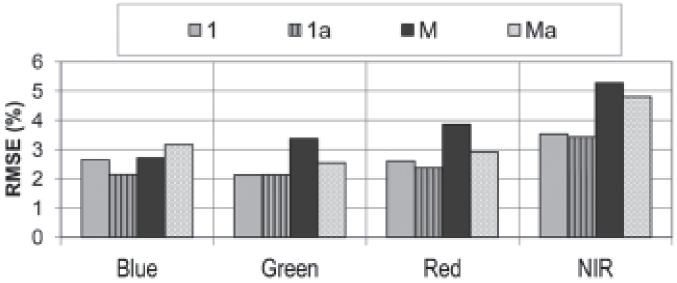


Fig. 5: Comparison of the vicarious calibration derived from the 10 cm flight (cases '1' and '1a') and the laboratory calibration (cases 'M' and 'Ma') applied for the 10 cm flight.

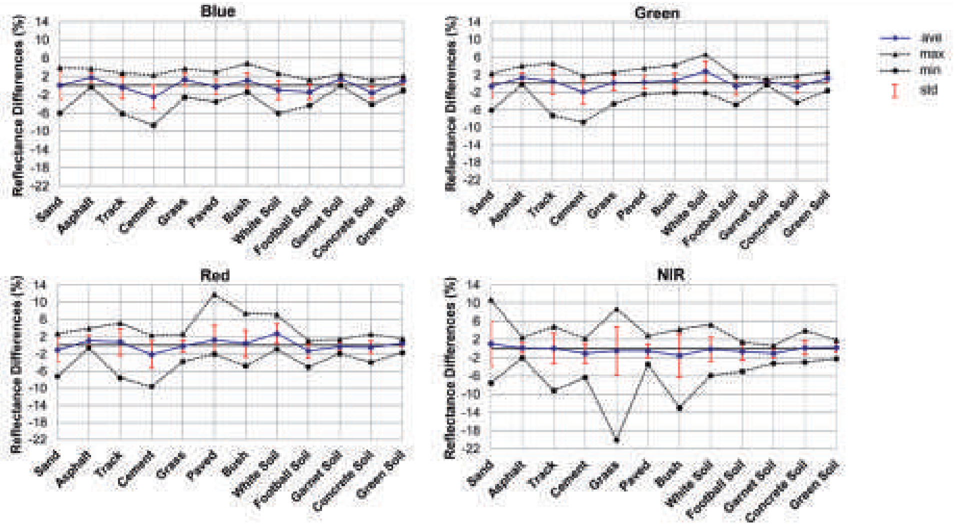


Fig. 6: Maximal and minimal differences between the ground reflectance measurements and those resulting from the vicarious calibration process (case '1a'), with its average and standard deviation (%).

band. The influence of the atmospheric correction in the vicarious calibration was low for all bands, B being the band showing the highest improvement (0.5 %). The RMSEs of the laboratory calibration ('M' and 'Ma') are between 2.6 and 4 % in the RGB band and values close to 5 % in the NIR band. In the case of 'M' the RMSE is higher when the wavelength increases. The results of the calibration coefficients revealed a greater sensitivity to the atmospheric effects ('Ma'), especially in the R and G bands (≈ 1 %). In this case, an increase of the error of 0.5 % for the B band was observed. The RMSEs in reflectances obtained for all the bands after applying the vicarious calibration ('1' and '1a') were smaller than those errors resulting from the laboratory calibration. The NIR band was most significant since it provided an improvement of about 2 %.

Tab. 3 shows the maximum and minimum differences between the ground reflectance measurements at the check surfaces and those resulting from the vicarious calibration with atmospheric correction (case '1a') in absolute terms (expressed in %) and independently for the 12 targets. It shows that the majority of the surface types used as check surfaces may be considered as invariant and suitable for radiometric validation. In general, uniform surfaces that are artificially coloured such as garnet and green provided the best accuracies in the four channels of the spectrum, while more variable surfaces like sand or vegetation had

Tab. 4: Statistics of the errors based on the radiometric manufacturer calibration (case ‘Ma’) for different check surfaces expressed in reflectances (%), max = maximal error, min = minimal error, std = standard deviation, RMSE = root−mean−square error, ave = average error, ave av = average absolute error value.

ADS40		Sand	Asphalt	Track	Cement	Grass	Paved	Bush	White Soil	Football Soil	Garnet Soil	Concrete Soil	Green Soil
Blue	max	5.52	5.89	5.09	5.85	4.18	6.00	5.93	8.89	3.67	3.32	4.84	3.58
	min	−3.26	1.63	−3.12	−4.48	−1.25	0.37	0.07	0.64	−1.34	1.00	−0.17	0.76
	std	2.62	1.08	2.20	2.35	1.20	1.38	1.47	1.89	1.42	0.68	1.19	0.75
	RMSE	3.31	3.88	3.13	2.77	2.39	3.37	2.79	5.77	1.92	2.45	2.69	2.76
	ave	2.08	3.72	2.25	1.51	2.06	3.09	2.39	5.46	1.32	2.35	2.42	2.66
	ave av	3.08	3.72	2.90	2.30	2.20	3.09	2.39	5.46	1.62	2.35	2.42	2.66
Green	max	3.14	4.76	5.63	3.25	2.76	4.50	4.74	9.10	2.53	1.33	3.29	3.59
	min	−4.86	0.39	−5.95	−7.17	−4.01	−1.01	−1.50	0.49	−3.74	−0.13	−2.82	−0.49
	std	2.28	1.16	2.81	2.67	1.64	1.38	1.62	2.31	1.84	0.39	1.32	0.91
	RMSE	2.26	2.27	3.25	2.66	1.71	1.98	1.97	5.76	1.82	0.84	1.51	2.36
	ave	0.34	1.96	1.66	−0.33	0.50	1.44	1.16	5.29	0.25	0.75	0.75	2.18
	ave av	1.81	1.96	2.91	2.04	1.44	1.62	1.46	5.29	1.50	0.76	1.24	2.21
Red	max	4.03	4.83	6.69	4.22	2.86	12.87	7.84	10.42	2.92	1.76	4.61	2.36
	min	−5.43	0.06	−5.70	−7.43	−3.29	−0.77	−4.03	2.48	−3.03	−1.50	−2.17	−0.77
	std	2.64	1.31	2.98	3.05	1.28	3.30	3.01	2.28	1.55	0.95	1.40	0.71
	RMSE	2.62	2.25	3.83	3.02	1.28	4.15	3.13	6.51	1.64	0.95	1.99	1.49
	ave	0.40	1.84	2.43	−0.08	0.10	2.56	1.06	6.11	0.62	0.15	1.42	1.32
	ave av	2.04	1.84	3.41	2.41	1.02	2.62	2.52	6.11	1.43	0.83	1.67	1.37
NIR	max	13.33	4.22	8.45	6.44	13.57	5.76	6.90	11.56	4.94	1.48	7.68	4.87
	min	−3.18	−0.49	−4.89	−1.85	−13.72	−0.98	−9.05	0.96	−1.01	−2.23	0.01	1.04
	std	4.41	1.18	3.12	2.09	4.94	1.50	4.37	2.54	1.69	0.89	1.71	0.96
	RMSE	6.28	2.08	5.02	3.89	6.44	2.65	4.54	6.87	3.43	0.89	3.85	3.45
	ave	4.55	1.73	3.96	3.29	4.15	2.19	1.51	6.40	3.00	−0.14	3.46	3.32
	ave av	5.34	1.76	4.39	3.51	5.70	2.24	4.04	6.40	3.08	0.72	3.46	3.32

poorer results. These were anisotropic surfaces in which the reflection angles change and the energy that is not gathered by the sensor is lost. That produced a slight increase in its RMSE and standard deviation (std).

The most uniform bands in terms of a similar error were B and G, and the most variable band was NIR (Fig. 6). Here, as expected, the higher error occurred in vegetated samples (bush and grass), because there the NIR band reflectance values and the standard deviation of the measurements were higher.

Next, a comparison with the results obtained using the coefficients of the ADS40 camera and atmospheric correction is done (Tab. 4). In general, for practically all check surfaces the RMSEs of the vicarious calibration are slightly lower than those obtained from the manufacturer calibration coefficients (Fig. 7).

3.3 *Application to 25 cm GSD Flight*

In this section the use of the calibration computed for the 10 cm flight was extrapolated to the 25 cm flight. The main goal was to analyze if it was possible to extrapolate the radiometric results of the vicarious calibration to different ground sample distances and flying heights. The results obtained, considering all the targets and all the images (Fig. 8) show a very similar trend for both flights. In particular, the vicarious calibration exhibited RMSEs lower than 3 % (except for the B band in case '1', where it was 3.5 %). The atmospheric correction in the vicarious calibration improved the results by 1.7 % in the B band, by 0.7 % in the R band and by very small values in the G and NIR bands. The vicarious calibration ('1' and '1a') provided higher precisions than those based on the laboratory calibration ('M' and 'Ma'), with a larger variation for the NIR band of about 2 %. Nevertheless, there was an exception for the B band, which provided a different performance. On one hand, without atmospheric correction, the accuracy of the vicarious calibration ('1') was slightly lower than the laboratory calibration ('M'). On the other hand, the application of the atmospheric correction on the laboratory calibration generated a solution slightly worse than the vicarious one (0.7 %).

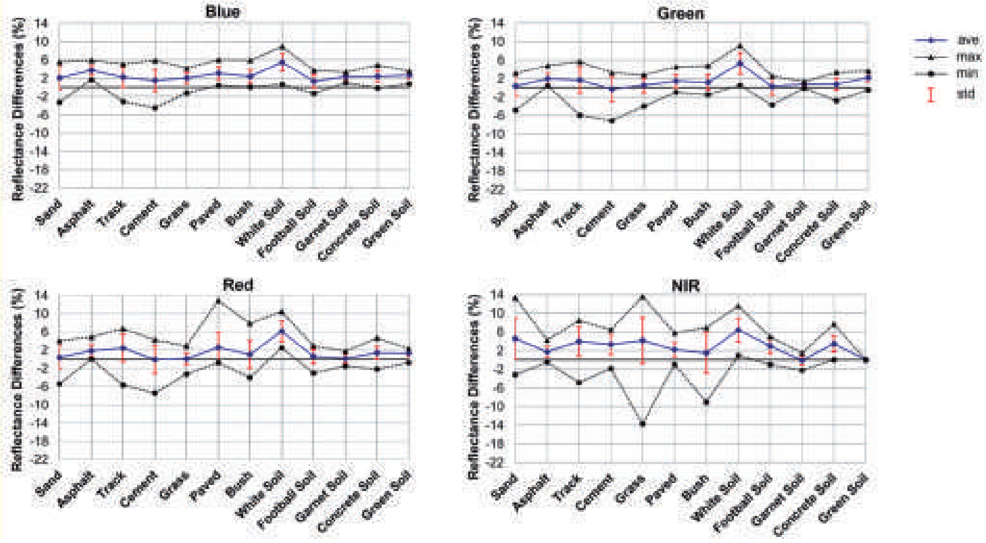


Fig. 7: Maximal and minimal differences between the ground reflectance measurements and the resulting from using the manufacturer calibration coefficients (case 'Ma'), together with its average and standard deviation (%).

4 Discussion

Other authors such as Honkavaara & Markelin (2007) have empirically analysed the spectral response of different airborne sensors through the use of grey patterns located in the field. In particular, they have examined the spectral response of the ADS40, DMC and UltraCamD systems through the use of a pattern with 8 different levels in grey scale. The conclusion is that these systems present an excellent radiometric quality and can be calibrated. The result of this investigation proved the stability of the radiometric laboratory calibration of the ADS40-SH5.

With regard to the calibration surfaces, we used PVC canvases to cope with the requirements mentioned in other works (Honkavaara et al. 2010) and it has been determined that they are suitable for the correct determination of the radiometric camera calibration, in particular ADS40 sensor (Markelin et al. 2010). However, sensor saturations were detected in the original images in some bands for bright objects, though the images have 12 bit dynamic range.

In our investigation raw sensor data have been used, considering the three-dimensional imaging geometry and a digital elevation model. Processing methods relying on the rectification and the consequent resampling of images could involve alterations to the original radiometric information.

The vicarious calibration is of interest in cases when there is doubt as to the calibration values supplied by the manufacturer or when the calibration needs to be validated. The results are consistent with the recently published papers (Markelin et al. 2010, Beisl 2010), showing that the RMSEs of surfaces reflectances can be better than 5 % in most of the evaluations. The performance varies between the different bands. The performance of the NIR band was lower (3–5 %) than the RGB band in all the cases. The vicarious calibration of the 10 cm flight achieved better results than the laboratory calibration; compared to ground reflectances, RMSEs in the NIR band decreased by about 2 %. In any case, the results achieved for a three-year old laboratory calibration demonstrate the radiometric stability of the sensor. The good weather conditions during the acquisition may have contributed to this result, and thus it would be interesting to repeat the procedure under less favourable conditions. Check surfaces provided large variations in reflectance, depending on the surface and the analyzed band, but they were never above 14 %, reaching even variations lower than 2 % for homogeneous and synthetic surfaces.

The accuracy obtained from the 10 cm calibration flight was preserved in the 25 cm flight. After these experiences, it should be mentioned that the errors increase with the wavelength if the laboratory calibration is used (case 'M'). Likewise, if the atmospheric correction is applied (case 'Ma') the accuracy for the B band gets worse. Therefore, the best solution is the vicarious calibration with atmospheric correction, named here case '1a'. In all the cases, R and G are the most stable bands.

On the other hand, there are some ways for improving the methodology that are worth mentioning: first, concerning the in situ measurements, the atmospheric measurements could be improved. For future work, it would

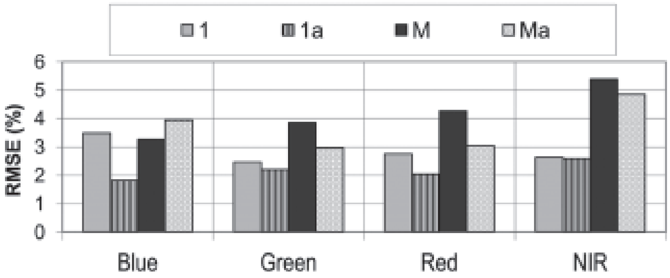


Fig. 8: Comparison of the vicarious calibration derived from the 10 cm flight (cases '1' and '1a') and the laboratory calibration (cases 'M' and 'Ma') applied for the 25 cm flight.

be of interest to analyze the influence of the quality of the meteorological parameters on the final results. Furthermore, it would be interesting to work on the influence of the view direction of the acquisition on the calibration: Although the images were taken approximately in nadir direction (2°) along the flying direction and thus the Lambertian approximation of the surface for the atmospheric correction was applicable, it would be of interest to consider surfaces with an anisotropic behaviour and then analyse bi-directional effects, e.g. applying kernel-driven models (Hernández-López et al. 2011).

5 Conclusions

In the last years there has without doubt been a convergence between photogrammetric and remote sensing methods for radiometric calibration. This study applied a combination of the photogrammetric principles (collinearity constraint, spatial resection, space intersection, …) according to the mathematical model of this linear sensor to establish the geometrical relations required for radiometric calibration (Tempelmann et al. 2005) and the principles of remote sensing, which allow the use of physical models for the image generation, taking into account the problems caused by the atmosphere.

It has been shown that it is possible to perform a radiometric calibration of photogrammetric images through the use of reflectance measurements to obtain and validate the radiometric camera calibration. It has been proven that the solution based on the vicarious calibration with the application of atmospheric correction slightly exceeds the quality of all bands based on the calibration given by the ADS40 sensor manufacturer. However, the ground reflectances obtained from the vicarious calibration and the laboratory calibrations provided by the manufacturer are both acceptable accuracies because the RMSEs achieved by validation for both cases are below 6 %.

It can be concluded that for this investigation, the consideration of atmospheric correction does not affect the results significantly in the imagery acquired at a GSD of 10 cm. Likewise, the use of a vicarious calibration is feasible if it is done shortly before or after the photogrammetric flight, and a different flying height does not significantly affect the results.

The results corroborate the potential of the ADS40 sensor and the need for appropriate laboratory or vicarious calibration to take full advantage of the high radiometric quality of the images.

Acknowledgements

The present contribution has been developed under the Research Project LICAM financed by the "Instituto Tecnológico Agrario de Castilla y León" (ITACYL).

References

AERONET, 2010: Aerosol Robotic Network. – http://aeronet.gsfc.nasa.gov (5.5.2012).

Beisl, U., 2006: Absolute spectroradiometric calibration of the ADS40 sensor. – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XXXVI** (1): on CD-ROM, Paris, France.

Beisl, U., Telaar, J. & Schönermark, M.v., 2008: Atmospheric correction, reflectance calibration and BRDF correction for ADS40 image data. – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XXXVII** (B7): 7–12.

Beisl, U. & Adiguezel, M., 2010: Validation of the reflectance calibration of the ADS40 airborne sensor using ground reflectance measurements. – ISPRS TC VII Symposium – 100 Years ISPRS, IAPRS **XXXVIII** (7B): 80–85, Vienna, Austria.

Casella, V., Franzini, M., Banchini, G. & Gentili, G., 2008: Initial evaluation of the second generation Leica ADS40 Camera. – ISPRS Congress Beijing, Commission I, WG **I/4:** 527–532.

Caspary, W.F., 1987: Concepts of network and deformation analysis. – Monograph **11**, The University of New South Wales, Kensington, N.S.W., Australia.

de Vries, C., Danaher, T., Denham, R., Scarth, P. & Phinn, S., 2007: An operational radiometric calibration procedure for the Landsat sensors based on pseudoinvariant target sites. – Remote Sensing of Environment **107** (3): 414–429.

Dinguirard, M. & Slater, P.N., 1999: Calibration of spatial-multispectral imaging sensors: a review. – Remote Sensing of Environment **68:** 194–205.

Hernández-López, D., Felipe García, B., González Piqueras, J. & Villa Alcázar, G., 2011: An approach to the radiometric aerotriangulation of photogrammetric images. – ISPRS Journal of Photogrammetry and Remote Sensing **66** (2011): 883–893.

Honkavaara, E. & Markelin, L., 2007: Radiometric Performance of Digital Image Data Collection. A Comparison of ADS40/DMC/Ultracam and EmergeDss. – Photogrammetric Week, Stuttgart.

Honkavaara, E., Arbiol, R., Markelin, L., Martinez, L., Cramer, M., Bovet, S., Chandelier, L., Ilves, R., Klonus, S., Marshal, P., Schläpfer, D., Tabor, M., Thom, C. & Veje, N., 2009: Digital Airborne Photogrammetry – A new tool for quantitative remote sensing? A State-of-the-art review on radiometric aspects of digital photogrammetric images. – Remote Sensing **1** (3): 577–605.

Honkavaara, E., Hakala, T., Peltoniemi, J., Suomalainen, J., Ahokas, E. & Markelin, L., 2010: Analysis of properties of reflectance reference targets for permanent radiometric test sites of high resolution airborne imaging systems. – Remote Sensing **2:** 1892–1917.

Honkavaara, E., Arbiol, R., Markelin, L., Martínez, L., Cramer, M., Bovet, S., Chandelier, L., Ilves, R., Klonus, S., Marshal, P., Schläpfer, D., Tabor, M., Thom, C. & Veje, N., 2011: The EuroSDR project Radiometric aspects of digital photogrammetric images. – Results of the empirical phase International Archives of the Photogrammetry. – Remote Sensing and Spatial Information Sciences **38** (Part 4/W19).

Kaasalainen, S., Krooks, A., Kukko, A. & Kaartinen, H., 2009: Radiometric calibration of terrestrial laser scanners with external reference targets. – Remote Sensing **1** (3): 144–158.

Krarup, T., Kubik, K. & Juhl, J., 1980: Gotterdammerung over least squares adjustment. – 14th Congress of the International Society of Photogrammetry **B3:** 369–378, Hamburg.

Markelin, L., Honkavaara, E., Beisl, U. & Korpela, I., 2010: Validation of the radiometric processing chain of the Leica ADS40 airborne photogrammetric sensor. – ISPRS TC VII Symposium – 100 Years ISPRS, IAPRS **XXXVIII** (7A): 145–150, Vienna, Austria.

Read, R.E. & Graham, R.W., 2002: Manual of Air Survey: Primary Data Acquisition. – Caithness, Scotland, UK, Whittles Publishing.

Reda, I. & Andreas, A., 2008: Solar Position Algorithm for Solar Radiation Applications. – Technical report, National Renewable Energy Laboratory.

Sandau, R., Braunecker, B., Driescher, H., Eckardt, A., Hilbert, S., Hutton, J., Kirchhofer, W., Lithopoulos, E., Reulke, R. & Wicki, S., 2000: Design principles of the LH systems ADS40 airborne digital sensor. – International Archives of Photogrammetry and Remote Sensing **33** (1): 258–265.

Teillet, P.M., Barsi, J.A., Chander, G. & Thome, K.J., 2007: Prime candidate earth targets for the post-launch radiometric calibration of satellite sensors. – SPIE International Symposium **6677**, San Diego, CA, USA.

Tempelmann, U., Borner, A., Chaplin, B., Hinsken, L., Mykhalevych, B., Miller, S., Recke, U., Reulke, R. & Uebbing, R., 2005: Photogrammetric Software for the LH Systems ADS40 Airbone Digital Sensor. – Technical report, Leica Geosystems Geospatial Imaging.

Vermote, E.F., Tanré, D., Deuzé, J.L., Herman, M. & Morcrette, J.J., 2006: Second Simulation of the Satellite Signal in the Solar Spectrum – Vector (6SV). – 55, Users Guide Version 3.0, University of Maryland, Greenbelt (MD) and Laboratoire d'Optique Atmosphérique, Lille (France).

Wagner, W., 2010: Radiometric calibration of small-footprint full-waveform airborne laser scanner measurements: Basic physical concepts. – ISPRS Journal of Photogrammetry & Remote Sensing **65:** 505–513.

Addresses of the Authors:

David Hernández-López & Beatriz Felipe-García, Regional Development Institute, University of Castilla-La Mancha, Albacete, Spain, e-mail: {david.hernandez}{beatriz.felipe}@uclm.es

Nilda Sánchez , CIALE, Spanish-Portuguese Centre of Agricultural Research, University of Salamanca, Spain, e-mail: nilda@usal.es

Diego González-Aguilera & Javier Gomez-Lahoz, Department of Cartographic and Land Engineering, University of Salamanca, Spain, e-mail: {daguilera}{fotod}@usal.es

Manuskript eingereicht: November 2011
Angenommen: Juni 2012

PFG 2012 / 5, 0573 – 0588
Stuttgart, October 2012

Article

Eignung von WorldView-2 Satellitenbildern für die Baumartenklassifizierung unter besonderer Berücksichtigung der vier neuen Spektralkanäle

Markus Immitzer, Clement Atzberger & Tatjana Koukal, Wien, Österreich

Keywords: tree species, classification, temperate forest, WorldView-2, linear discriminant analysis

Summary: *Suitability of WorldView-2 data for tree species classification with special emphasis on the four new spectral bands*. There is an increasing demand for information on tree species composition and spatial distribution. Detailed tree species maps are essential for addressing different ecological problems. Moreover, they are important in modern forest management as close to nature forest management is becoming more and more common. This study examines the potential of 8-band WorldView-2 data with 2 m ground sample distance (GSD) for identifying 10 tree species in a mid-European forest. We delineated well-illuminated tree crowns manually and did a supervised classification using linear discriminant analysis (LDA) implemented in a bootstrapping environment (1465 observations, 500 bootstrap samples). The overall accuracy for the 10 tree species was around 84 % (8 bands) with class-specific producer's accuracies ranging between 54 % (*Carpinus betulus*) and 96 % (*Fagus sylvatica*). In general, the accuracy was higher for the 5 deciduous tree species (89 %) than for the 5 coniferous tree species (79 %). The classification accuracies decreased significantly, if only the 4 conventional bands Blue, Green, Red, and Near Infrared 1 were used (78 % overall accuracy). However, when the classification focused only on the 4 main tree species in the investigation area, i.e. *Picea abies*, *Pinus sylvestris*, *Fagus sylvatica*, *Quercus robur*, the 4 conventional bands were sufficient to achieve high classification accuracies (95 % overall accuracy). Adding the 4 new bands (Coastal, Yellow, Red Edge, Near Infrared 2) to the input feature set did not further improve the overall classification accuracy for these 4 main tree species. Hence, the positive impact of the additional 4 new bands resulted from strongly increased classification accuracies of the 6 secondary tree species. Green and Near Infrared 1 (conventional bands) as well as the bands Red Edge and Near Infrared 2 (new bands) contributed most to class separability (according to Wilks' Lambda).

Zusammenfassung: Informationen über die Baumartenzusammensetzung und -verteilung werden vermehrt nachgefragt. Detaillierte Baumartenkarten sind beispielsweise bei vielen ökologischen Fragestellungen unverzichtbar, stellen aber auch in der forstlichen Praxis aufgrund der mehr und mehr praktizierten naturnahen Waldbewirtschaftung eine wichtige Datengrundlage dar. In dieser Studie wird das Potenzial der 8-Band WorldView-2 Daten mit einer Bodenauflösung von 2 m für die Unterscheidung von 10 Baumarten in einem mitteleuropäischen Testgebiet untersucht. Dazu wurden gut beleuchtete Teile von Baumkronen manuell abgegrenzt und mit Hilfe der Linearen Diskriminanzanalyse (LDA) klassifiziert. Durch Bootstrapping (1465 Referenzflächen, 500 Wiederholungen) wurde für die 10 Baumarten eine Gesamtgenauigkeit von 84 % ermittelt (8 Kanäle). Die Ergebnisse variierten von Baumart zu Baumart zum Teil beträchtlich (Produzentengenauigkeit bei Hainbuche 54 % und bei Buche 96 %). Generell lag die Klassifikationsgenauigkeit bei den 5 Laubbaumarten höher (89 %) als bei den 5 Nadelbaumarten (79 %). Die Klassifikationsgenauigkeiten nahmen deutlich ab, wenn nur die 4 konventionellen Spektralkanäle Blue, Green, Red und Near Infrared 1 verwendet wurden (78 % Gesamtgenauigkeit). Wurden jedoch lediglich die 4 Hauptbaumarten kartiert (Fichte, Kiefer, Buche und Eiche), konnte bereits mit den 4 konventionellen Kanälen eine hohe Gesamtgenauigkeit erzielt werden (95 %). Für die 4 Hauptbaumarten konnte durch die zusätzliche Verwendung der 4 neuen Kanäle (Coastal, Yellow, Red Edge, Near Infrared 2) keine Steigerung der Klassifikationsgenauigkeit erzielt werden. Der positive Einfluss der 4 zusätzlichen Kanäle ist dagegen bei den 6 Nebenbaumarten stark ausgeprägt. Die Untersuchung der Trennkraft der einzelnen Kanäle ergab, dass von den 4 konventionellen Kanälen Green und Near Infrared 1 am meisten zur Trennung der Klassen beitragen und von den 4 neuen Kanälen Red Edge und Near Infrared 2 (Wilks' Lambda).

DOI: 10.1127/1432-8364/2012/0140 1432-8364/12/0140 $ 4.00

1 Einleitung

Die Identifikation und Kartierung von Baumarten stellt eine Aufgabe dar, die sowohl forstwirtschaftlich als auch ökologisch von großer Relevanz ist. Die Baumartenzusammensetzung sowie die räumliche Verteilung verschiedener Baumarten sind beispielsweise wesentliche Eingangsparameter bei Habitatmodellierungen zur Beschreibung der Artenbiodiversität und für Monitoringzwecke. Aber auch in der forstlichen Planung sind baumartenbezogene Aussagen unverzichtbar (Wulder et al. 2004, McDermid et al. 2009). Gerade die zunehmende naturnahe Bewirtschaftung erhöht den Aufwand der Baumartenerkennung, da großflächige Monokulturen in strukturierte, gemischte Bestände mit kleinflächigen Nutzungen übergeführt werden. Die traditionellen, rein terrestrischen Methoden der Datenerfassung wie Stichprobeninventur und Taxation liefern lediglich punktuelle Informationen. Diese können zwar auf die Fläche hochgerechnet werden, ermöglichen aber keine detaillierte flächenhafte Darstellung. Dies ist besonders bei baumartenreichen Wäldern oftmals unzureichend.

Der Einsatz von Fernerkundungsmethoden kann einen wesentlichen Beitrag zur Beseitigung von forstlichen Informationsdefiziten leisten (Hildebrandt 1996, Olsson et al. 2008, Falkowski et al. 2009). So können bei der visuellen Stereointerpretation von Farbinfrarot-Luftbildern, u.a. durch die charakteristische Farbe, Kronenform und -struktur, Baumarten generell sehr gut erfasst werden (Albertz 2009). Eine großflächige Anwendung scheitert dabei aber meist an der zeitaufwändigen Bearbeitung und den damit verbundenen hohen Kosten. Die automatisierte Baumartenkartierung stellt daher einen aktuellen Forschungsschwerpunkt der forstlichen Fernerkundung dar (Brosinger 2010, Buck et al. 2010). Dabei weisen die zur Verfügung stehenden Sensoren systemspezifische Vor- und Nachteile auf: Digitale Orthophotos bieten zwar eine hohe räumliche Auflösung bei geringen Erstellungskosten, sind jedoch auf die Spektralkanäle Blau, Grün, Rot und Nahes Infrarot beschränkt. Die über das Bild gesehen stark variierende Aufnahmerichtung und die daraus ableitbare Information über das richtungsabhängige Reflexionsverhalten kann zwar prinzipiell in einer Klassifizierung genutzt werden (Heikkinen et al. 2011, Koukal & Atzberger 2012). Da diese direktionalen Effekte im Zuge der Bildauswertung üblicherweise jedoch nicht berücksichtigt werden, stellen sie in den meisten Anwendungen eher eine Fehlerquelle dar. Satellitengestützt stehen sowohl multi- als auch hyperspektrale Sensoren zur Verfügung, jedoch bei meist geringerer räumlicher Auflösung als bei Luftbilddaten. Die geringe räumliche Auflösung verhindert eine einzelbaumweise Erfassung und die resultierenden Mischpixel erschweren eine korrekte Identifizierung der Baumarten (Schlerf et al. 2003). Dem hohen spektralen Informationsgehalt von Hyperspektraldaten steht die meist eher geringe Verfügbarkeit bei hohen Kosten gegenüber.

Mit WorldView-2 steht seit 2010 ein satellitengestützter Sensor zur Verfügung, der durch die hohe räumliche Auflösung sowie durch 4 neue Spektralkanäle ein hohes Potenzial für die Untersuchung von Vegetation aufweist (siehe Kapitel 2). Erste Ergebnisse über die Verwendung dieser Daten zur Ermittlung forstlicher Strukturparameter wurden bereits veröffentlicht (Huang & Cao 2011, Ozdemir & Karnieli 2011). In einer Überblicksstudie bewerteten Marchisio et al. (2010) die Steigerung der Klassifikationsgenauigkeit durch die erstmals verfügbaren Kanäle (Coastal, Yellow, Red Edge und Near Infrared 2) gegenüber traditionellen 4-Kanal-Bildern von Sensoren mit vergleichbarer räumlicher Auflösung. Die Klassifikationsergebnisse konnten um 5 bis 20 % erhöht werden. Auch Collin & Planes (2011) erzielten durch die 4 zusätzlichen Kanäle eine Steigerung der Klassifikationsgenauigkeiten. Weitere Arbeiten mit WorldView-2 Daten wurden im Zuge der von DigitalGlobe ausgeschriebenen „8-Band Challenge“ publiziert. Darunter finden sich auch einige Studien, die sich mit dem Themenkomplex der „Baumartenerkennung“ beschäftigten: Chen (2011) konnte bei der Baumartenidentifizierung auf Hawaii mit WorldView-2 Bildern deutlich höhere Genauigkeiten erzielen als mit IKONOS Daten. Hamdan (2010) erreichte bei der Klassifikation von 10 tropischen Baumarten in Malaysia sehr gute Ergebnisse. Sridharan (2011) klassifizierte urbane Wälder in Texas auf mehreren Detaillierungsstufen bis hinunter aufs Artenniveau und erreichte nur geringfügig geringere Klassifikationsgenauigkeiten als mit AISA Hyperspektraldaten.

Der Einsatz von WorldView-2 für forstliche Fragestellungen ist auf Grund der hohen räumlichen Auflösung von 0,5 m (panchromatisch) bzw. 2 m (multispektral), der zusätzlichen 4 Kanäle und der raschen Verfügbarkeit der Daten eine interessante Alternative zu Orthophotos. Für mitteleuropäische Wälder sind derzeit jedoch noch kaum Untersuchungen publiziert. In dieser Studie werden daher die forstlichen Einsatzmöglichkeiten von WorldView-2 Daten im Bereich der Baumartenerkennung in einem mitteleuropäischen Waldgebiet untersucht. Folgende Fragen stehen im Mittelpunkt: (1) Welche Baumarten können auf Grund der spektralen Information aus WorldView-2 Daten unterschieden werden, (2) welchen Mehrwert haben die 4 neuen Kanäle im Vergleich zur alleinigen Verwendung der 4 konventionellen Kanäle, und (3) welche der insgesamt 8 Kanäle tragen am meisten zur Baumartenunterscheidung bei?

Für die Beantwortung dieser Fragen werden die Reflexionswerte in den multispektralen Kanälen mit 2 m Pixelgröße von manuell abgegrenzten, beleuchteten Kronenteilen verwendet. Zur Baumartenklassifikation und Bestimmung der spektralen Trennkraft wurde die Lineare Diskriminanzanalyse (LDA) verwendet.

2 Daten und Methodik

2.1 *Untersuchungsgebiet*

Das Untersuchungsgebiet befindet sich im Osten Österreichs (Burgenland, Bezirk Oberpullendorf) und ist durch hügelige Landschaft geprägt. Mit Seehöhen von 290 bis 670 m über NN liegt es in der submontanen bis tiefmontanen Höhenstufe des forstlichen Wuchsgebiets 5.2 „Bucklige Welt". Der Jahresniederschlag liegt, mit einem sommerlichen Maximum, zwischen 700 und 1100 mm. Die Grundgesteine sind vor allem basenarme Silikate, Gneis und Quarzphyllit. Die potentielle natürliche Waldgesellschaft reicht vom Eichen - Hainbuchenwald über sauren Weiß-Kiefern - Eichenwald bis hin zum Tannen - Buchenwald mit Beimischungen von Eichen, Edelkastanie und Weiß-Kiefer (Kilian et al. 1994).

Das Untersuchungsgebiet umfasst eine Fläche von knapp 3000 ha, auf welchen die beiden Nadelhölzer Gemeine Fichte (*Picea abies*, (L) Karst.) und Weiß-Kiefer (*Pinus sylvestris*, L.) gemeinsam mit den beiden Laubhölzern Rot-Buche (*Fagus sylvatica*, L.) und Stiel-Eiche (*Quercus robur*, L.) die Hauptbaumarten darstellen. Weiters kommen die Baumarten Weiß-Tanne (*Abis alba*, Mill.), Europäische Lärche (*Larix decidua*, Mill.), Douglasie (*Pseudotsuga menziesii*, (Mirb.) Franco), Scheinzypresse (*Chamaecyparis lawsoniana* (A.Murr.) Parl.), Riesen-Tanne (*Abies grandis*, Lindl), Gelb-Kiefer (*Pinus ponderosa*, Douglas ex P. et C.Laws), sowie Gemeine Esche (*Fraxinus excelsior*, L.), Hainbuche (*Carpinus betulus*, L.), Zerreiche (*Quercus cerris*, L.), Weiß-Birke (*Betula pendula*, Roth), Schwarz-Erle (*Alnus glutinosa*, L.), Wal-Nuss (*Juglans regia*, L.), Robinie (*Robinia pseudoacacia*, L.), Edelkastanie (*Castanea sativa*, Mill.), Vogel-Kirsche (*Prunus avium*, L.), Linden- (*Tilia* sp.), Ahorn- (*Acer* sp.), Ulmen- (*Ulmus* sp.) und Pappel-Arten (*Populus* sp.) vor.

2.2 *WorldView-2 Daten*

Der WorldView-2 Satellit liefert seit Anfang 2010 räumlich hochauflösende Daten

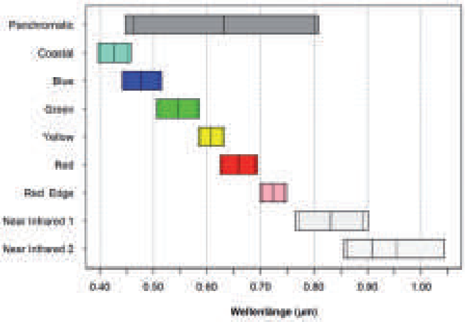


Abb. 1: Lage und Bandbreite der 8 Farbkanäle und des panchromatischen Kanals des WorldView-2 Sensors sowie Bandbreite (gestrichelt) und mittlere Wellenlänge (durchgezogen) bei 50 % der maximalen Sensitivität (Updike & Comp 2010).

in 8 Spektralkanälen (Abb. 1). Bei senkrechter Aufnahmerichtung beträgt die Bodenauflösung des panchromatischen Kanals 50 cm und die der multispektralen Kanäle 200 cm. Zusätzlich zu den 4 üblichen Kanälen Blue, Green, Red und Near Infrared 1 stehen 4 weitere Kanäle, Coastal, Yellow, Red Edge und Near Infrared 2, zur Verfügung, für welche vom Anbieter ein großes Potenzial für vegetationskundliche Fragestellungen postuliert wird (DigitalGlobe 2009).

Die verwendete WorldView-2 Szene wurde am 10. Juli 2010 aufgenommen und deckt eine Fläche von 7500 ha ab (Scandirection: forward, Mean Satellite Elevation: 77,8°, Mean Satellite Azimuth: 76,6°, Mean Off Nadir View Angle: 11°).

Zur radiometrischen und atmosphärischen Korrektur der Aufnahme wurden zunächst die Pixelgrauwerte in „at-sensor“ Strahldichten umgerechnet (Updike & Comp 2010). Anschließend wurde das Bild mit Hilfe des ENVI-Moduls (ENVI 4.8) FLAASH atmosphärisch korrigiert. Die Einstellungen wurden dabei iterativ durch Plausibilitätsprüfung der resultierenden Reflexionswerte eruiert. Für das Pansharpening wurde der speziell für WorldView-2 Daten entwickelte und in ERDAS Imagine 2010 verfügbare HCS-Algorithmus (Hyperspherical Colour Space Algorithmus) angewendet, welcher die Einbeziehung aller 8 Kanäle ermöglicht (Padwick et al. 2010). Die Georektifizierung wurde mittels ERDAS (Control Points) unter Verwendung eines Digitalen Geländemodells (5 m × 5 m Raster) vorgenommen. Die Passpunktkoordinaten stammten aus einem Farbinfrarot-Orthophoto (Aufnahmejahr 2007, Pixelgröße 0,5 m). Die erreichte mittlere Lagegenauigkeit (RMSE) lag im 2 m Bild bei 0,70 Pixel (x: 0,50, y: 0,47).

2.3 *Auswahl von Referenzflächen*

Zur Auswahl von Referenzflächen wurde eine digitale Bestandeskarte herangezogen. Diese enthält verschiedenste Bestandesattribute, die im Zuge der Forsteinrichtung durch Taxation und Stichprobeninventur erhoben wurden. Die Taxation dient der bestandesweisen Zustandserhebung. Dabei werden in den einzelnen Beständen durch Winkelzählproben die Baumartenanteile und Vorräte ermittelt. Auf Grund der üblichen Mindestgröße von 0,5 ha können aber durchaus auch, bezüglich Aufbau, Struktur und Baumarten, uneinheitliche Waldteile in Beständen zusammengefasst sein. Dadurch ist eine direkte Umlegung der Taxationsdaten auf die gesamte Bestandesfläche nicht ohne Weiteres möglich.

Ausgehend von den angegebenen Baumartenanteilen wurden möglichst reine Bestände sowie Bestände mit eindeutig identifizierbaren Baumarten ausgewählt. Durch visuelle Interpretation wurden die Taxationsangaben auf Plausibilität überprüft und anschließend Beispielbäume ausgewählt. Dabei wurde auf eine gute Verteilung der ausgewählten Bestände über das Untersuchungsgebiet und die Erfassung verschiedener Altersklassen geachtet. Um Beschattungs- und Beleuchtungseinflüsse zu minimieren, wurden nur die gut beleuchteten Baumkronen(-teile) verwendet (Clark et al. 2005, Leckie et al. 2005b, Greenberg et al. 2006).

Die Auswahl der Beispielkronen erfolgte auf Basis des Bildes mit 0,5 m Pixelgröße (pansharpened) auf Grund der Taxationsdaten ohne visuelle Vorinterpretation (Abb. 2). Gut beleuchtete Kronenbereiche, die möglichst sicher einer Baumart zuzuordnen waren, wurden anschließend im Bild mit 2 m Pixelgröße

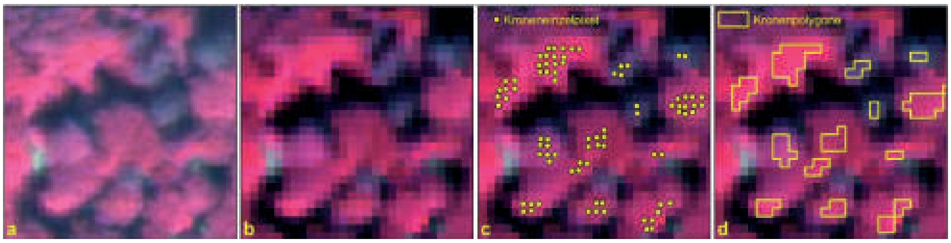


Abb. 2: Auswahl beleuchteter Baumkronen(-teile) als Referenzflächen: Baumkronen a) im Bild mit 0,5 m Pixelgröße, b) im Bild mit 2 m Pixelgröße, c) Auswahl der Kronenpixel durch Punktsetzung, d) Referenzpolygone entsprechend den pixelweise ausgewählten Kronenteilen.

aufgesucht und durch pixelweises Setzen von Punkten (Shapefile) markiert. Den Punkten wurden die wesentlichen Attribute wie Baumart und Alter aus den Forsteinrichtungsdaten zugeordnet. Punkte in benachbarten Rasterzellen wurden anschließend zu Polygonen zusammengefügt und in eine Vektordatei (Polygon-Shapefile) konvertiert. Dieser Datensatz wurde schließlich für das Auslesen der Spektralinformationen aus dem WorldView-2 Bild herangezogen. Für jede Referenzfläche wurden die Mittelwerte pro Kanal ausgelesen.

Neben den 4 Hauptbaumarten Gemeine Fichte, Weiß-Kiefer, Rot-Buche und Stiel-Eiche konnten auch Bestände mit den Baumarten Europäische Lärche, Douglasie, Scheinzypresse, Hainbuche, Weiß-Birke und Schwarz-Erle eindeutig identifiziert werden. Diese 6 Nebenbaumarten wurden ebenfalls in die Analysen aufgenommen. Weitere Baumarten konnten nicht berücksichtigt werden, da nicht genügend Individuen eindeutig identifiziert werden konnten.

Für die 10 untersuchten Baumarten sind in Tab. 1 Angaben über die Anzahl der Referenzflächen und Flächenanteile aufgeführt. Insgesamt wurden 1465 Referenzflächen generiert. Die unterschiedlichen Anteile spiegeln ungefähr die relativen Flächenverhältnisse vor Ort wider.

2.4 *Spektrale Trennbarkeit und Klassifikation der Baumarten*

Für die Auswertung wurden die spektralen Mittelwerte der einzelnen Referenzpolygone verwendet. Die Analyse und Visualisierung der Daten erfolgte mit dem Statistikprogramm R 2.14.1 (R Development Core Team 2011) mit den Zusatzpaketen MASS (Venables & Ripley 2002) und scatterplot3d (Ligges & Mächler 2003). Die Baumartenklassifizierung erfolgte mit Hilfe der Linearen Diskriminanzanalyse (LDA). Die LDA ist ein strukturprüfendes Verfahren, welches die Abhängigkeit der Gruppenzugehörigkeit (nominal skaliert) von den metrisch skalierten Merkmalsvariablen prüft. Dabei werden durch Linearkombinationen der Merkmalsvariablen sogenannte Diskriminanzfunktionen erstellt. Deren Eigenwerte beschreiben den Anteil der durch diese Funktionen erklärten Varianz. Die Trennkraft der einzelnen Kanäle wurde sowohl univariat (ANOVA, Wilks' Lambda) als auch multivariat (standardisierte mittlere Diskriminanzkoeffizienten) beurteilt (Backhaus et al. 2008).

Die Klassifikationsgenauigkeiten wurden durch Bootstrapping (Efron & Tibshirani 1993) ermittelt. Dabei wird durch das zufällige Ziehen mit Zurücklegen ein Trainingsdatensatz der Größe der Stichprobe gewonnen. Jene Beispieldaten, die nicht im Trainingsdatensatz enthalten sind (ca. 36,8 % der Daten), verbleiben als unabhängiger Testdatensatz.

Tab. 1: Verteilung der Referenzflächen auf die 10 untersuchten Baumarten (getrennt in Nadel- und Laubholz und gereiht nach der forstlichen Bedeutung).

Baumart	Wissenschaftliche Bezeichnung	Abkürzung	Anzahl Pixel	Anzahl Referenzflächen	Anteil (%)
Gemeine Fichte	*Picea abies*	Fi	1084	226	15,4
Weiß-Kiefer	*Pinus sylvestris*	Ki	807	235	16,0
Europäische Lärche	*Larix decidua*	Lä	472	122	8,3
Douglasie	*Pseudotsuga menziesii*	Dgl	677	178	12,1
Scheinzypresse	*Chamaecypris lawsoniana*	SZy	166	42	2,9
Rot-Buche	*Fagus sylvatica*	Bu	1519	247	16,9
Stiel-Eiche	*Quercus robur*	Ei	1770	152	10,4
Hainbuche	*Carpinus betulus*	HBu	445	81	5,5
Weiß-Birke	*Betula pendula*	Bi	397	86	5,9
Schwarz-Erle	*Alnus glutinosa*	SEr	387	96	6,6
		Summe	**7724**	**1465**	**100,0**

Die Anzahl der Bootstrap-Stichproben betrug in unserer Studie 500, wobei die zufällige Referenzflächenauswahl entsprechend der jeweiligen Anzahl an Referenzflächen pro Klasse stratifiziert erfolgte.

Die Analyse der Klassifikationsergebnisse der einzelnen Baumarten erfolgte durch die Erstellung von Klassifikationsmatrizen (Congalton & Green 1999). Im Zuge der 500 Bootstrap-Stichproben wurde jede Referenzfläche mehrmals klassifiziert. Für die Klassifikationsmatrizen wurde für jede Referenzfläche die am häufigsten klassifizierte Klasse (Modalwert) verwendet. Der Vergleich von Ergebnissen unterschiedlicher Methoden bzw. Datengrundlagen erfolgte über den Mittelwert der Gesamtgenauigkeiten aus den einzelnen Bootstrap-Wiederholungen.

2.5 *Flächige Anwendung auf Testfläche*

Für die flächige Anwendung der Klassifizierung wurde im Untersuchungsgebiet eine Testfläche mit großer Baumartenvielfalt ausgewählt. Innerhalb dieser Testfläche wurden alle Baumkronen abgegrenzt, wobei die Abgrenzung wie bei den Referenzflächen erfolgte (Abb. 2). Zur Plausibilitätsprüfung der Klassifizierungsergebnisse standen Taxationsdaten zur Verfügung. Obwohl die Taxation keine absolut zuverlässige Referenzinformation für die gesamte Fläche im Sinne einer Vollaufnahme darstellt, kann diese Information dennoch für eine näherungsweise Beurteilung der Klassifikationsergebnisse verwendet werden.

3 Ergebnisse

3.1 *Spektrale Signaturen*

In Abb. 3 sind die gemittelten Spektralsignaturen über alle Referenzpolygone einer Baumart dargestellt. Der Vergleich der spektralen Signaturen zeigt wie erwartet bei den Laubhölzern deutlich höhere Reflexionswerte im nahen Infrarot als bei den Nadelhölzern, wobei die Rot-Buche, gefolgt von der Hainbuche, die höchsten Werte aufweist. Zwischen Stiel-Eiche, Weiß-Birke und Schwarz-Erle sind im nahen Infrarot nur geringe Unterschiede erkennbar. Bei den Nadelhölzern zeigt die Scheinzypresse die höchsten Reflexionswerte im nahen Infrarot. Mit abnehmenden Reflexionsgraden folgen Douglasie, Europäische Lärche, Weiß-Kiefer und Gemeine Fichte. Der Detailausschnitt für den Wellenlängenbereich des sichtbaren Lichts zeigt, dass die Unterschiede zwischen den Baumarten in den Kanälen Green und Yellow am größten sind. Eine eindeutige Differenzierung zwischen Laub-

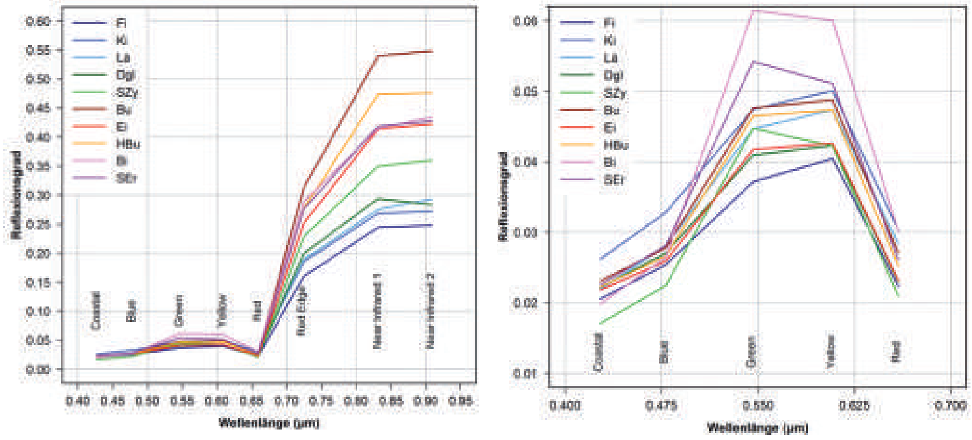


Abb. 3: Mittlere spektrale Signaturen der 10 untersuchten Baumarten über alle Referenzpolygone, links: gesamter Wellenlängenbereich, rechts: Detailansicht im Bereich des sichtbaren Lichts.

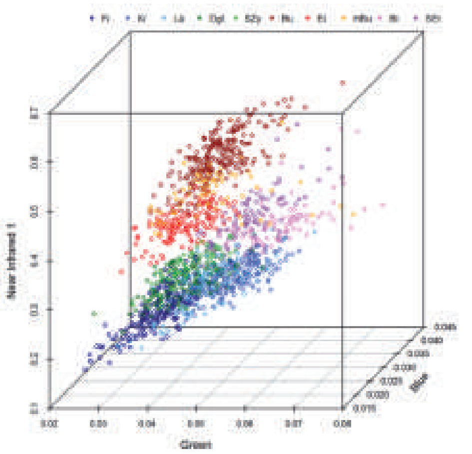


Abb. 4: Verteilung der Referenzflächen der untersuchten Baumarten in den Kanälen Blue, Green und Near Infrared 1.

und Nadelbäumen ist hier jedoch nicht möglich. Die Ergebnisse entsprechen typischen Signaturen für Laub- und Nadelbäume, wie sie auch in der Literatur beschrieben werden (Kadro 1981, Hosgood et al. 1994, Jones et al. 2010).

In Abb. 4 sind die Reflexionswerte aller Referenzflächen in den Kanälen Blue, Green und Near Infrared 1 dargestellt. Dies erlaubt einerseits Rückschlüsse auf die Trennbarkeit der Baumarten und zeigt anderseits auch die Streuungen innerhalb der Baumarten sowie die baumartenspezifischen Korrelationen zwischen den 3 Spektralkanälen. Während die Trennung in Nadel- und Laubbäume anhand dieses Streudiagramms noch relativ gut möglich ist, überlagern sich die Punktwolken der einzelnen Baumarten innerhalb dieser zwei Gruppen teilweise sehr stark.

3.2 *Lineare Diskriminanzanalyse (LDA)*

Für die LDA wurden alle 8 Kanäle verwendet, da laut univariater ANOVA alle Spektralkanäle die 10 Baumarten höchst signifikant (F-Test, P<0,01) trennen. Die erstellten 8 Diskriminanzfunktionen wurden ebenfalls auf ihre Trennkraft untersucht. Dabei wurde festgestellt, dass alle Funktionen einen signifikanten Beitrag zur Trennung liefern (Chi^2-Test, P<0,05). Auch wenn die höchste Trennkraft bei der Verwendung aller Funktionen erzielt wurde, nahmen die Eigenwerte der einzelnen Funktionen von der ersten zur achten Diskriminanzfunktion stark ab. So trug die erste Funktion bereits 75,6 % der Varianz; die zweite trug 14,0 %, die dritte 6,6 % und die vierte 2,7 % zur Erklärung der Varianz bei. Die Anteile der anderen 4 Funktionen lagen unter 1 %.

Zur Veranschaulichung der Trennkraft der ersten 4 Diskriminanzfunktionen sind in Abb. 5 Streudiagramme von jeweils 2 Funktionen sowie die dazugehörigen univariaten Dichteverteilungen dargestellt, welche für jede Baumart die Streuung in der jeweiligen Funktion zeigen. Die Abbildung macht ersichtlich, dass durch die Kombination der ersten beiden Funktionen bereits eine deutliche Trennung der Nadel- von den Laubbäumen erzielt wird. Aber auch innerhalb der Laubbäume sind die einzelnen Baumarten als gut trennbare Gruppen erkennbar, wobei sich Schwarz-Erle und Weiß-Birke von den anderen 3 Baumarten deutlicher absetzen. Die Kombinationen der ersten mit der dritten und in geringerem Maße auch mit der vierten Diskriminanzfunktion separieren dann auch die Arten innerhalb der Gruppe der Nadelbäume besser.

3.3 *Klassifikationsergebnisse*

Das durch Bootstrapping ermittelte Klassifikationsergebnis bei Verwendung aller 8 Kanäle ist in Form einer Klassifikationsmatrix in Tab. 2 dargestellt. Über alle 10 Baumarten konnten 83,7 % der 1465 Referenzflächen richtig klassifiziert werden (Kappa: 0,814). Die Trennung zwischen Nadel- und Laubbäumen erfolgte zu 99,1 % richtig. Innerhalb der Nadelbäume sind 79,3 % der Referenzflächen richtig klassifiziert. Die häufigsten Fehlklassifikationen finden sich zwischen Douglasien und Fichten (16,3 %) und zwischen Lärchen und Kiefern (16,4 %) bzw. Fichten (15,8 %). Mit 64,8 % weist die Lärche die geringste Produzentengenauigkeit der Nadelbäume auf, während die Kiefer mit 87,2 % die höchste Produzentengenauigkeit zeigt.

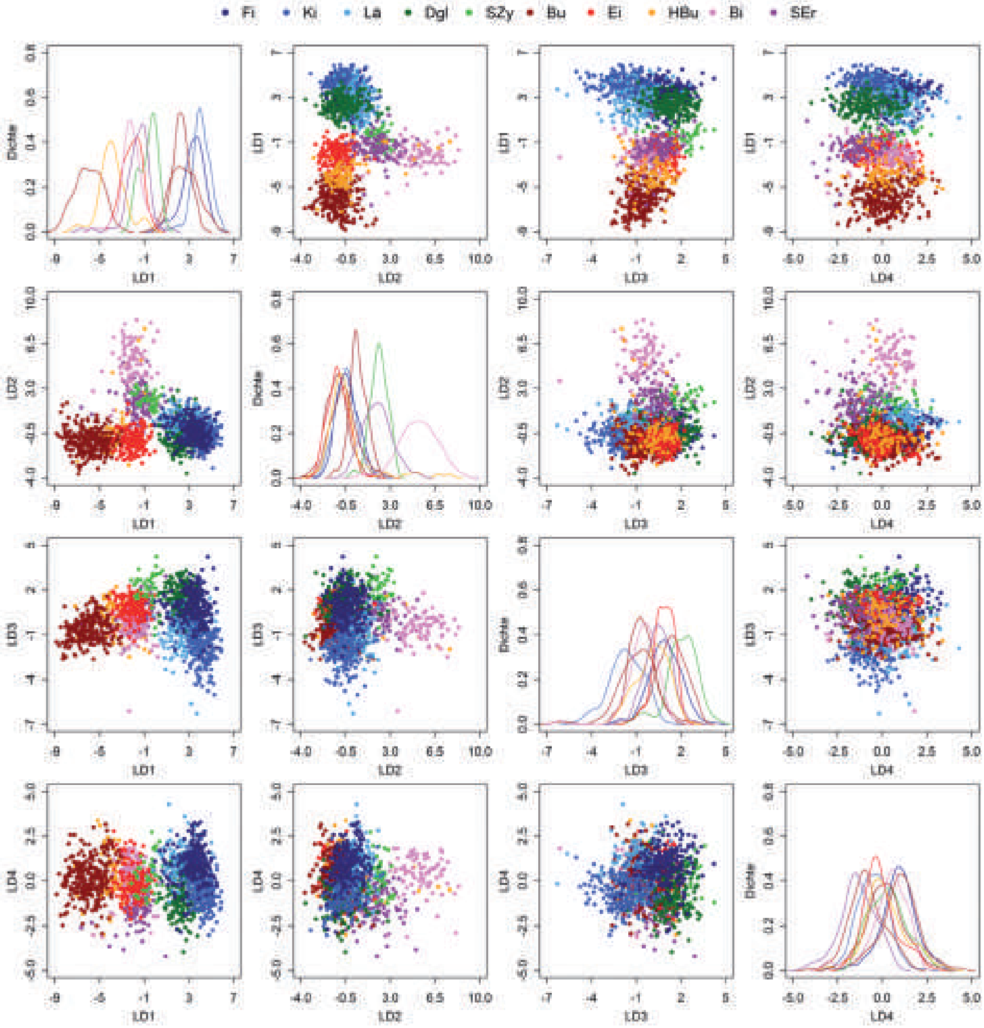


Abb. 5: Streudiagramme der ersten 4 Diskriminanzfunktionen (LD1 bis LD4) zur Visualisierung der Trennkraft der einzelnen Funktionen sowie die zugehörigen Dichtefunktionen für die 10 Baumarten.

Die 662 Referenzflächen der Laubbaumarten werden zu 89,0 % richtig klassifiziert. Die häufigsten Fehlklassifikationen bei Laubbäumen sind bei Hainbuche zu beobachten. Das beste Klassifikationsergebnis zeigt die Buche, die mit 96,0 % die höchste Produzenten- und mit 91,9 % gemeinsam mit Birke (92,0 %) auch die höchste Nutzergenauigkeit aller Baumarten aufweist.

Werden für die Analysen nur die 4 konventionellen Kanäle verwendet, zeigt sich über beinahe alle Baumarten eine deutliche Verschlechterung der Produzenten- und Nutzergenauigkeiten (Tab. 3). Unverändert bleiben dagegen die sehr geringen Fehlklassifikationen zwischen Nadel- und Laubbäumen (<1 %). Der Anstieg der Fehlerraten beschränkt sich somit auf die bereits erläuterten Verwechslungen innerhalb der Nadel- bzw. Laubhölzer. Am deutlichsten zeigt sich die Verschlechterung bei Hainbuche: konnten mit 8 Kanälen noch 44 von den 81 Referenzflächen richtig

Tab. 2: Klassifikationsmatrix für die Klassifikation mit 8 Kanälen (Modalwerte aus 500 Bootstrap-Wiederholungen, PG = Produzentengenauigkeit, NG = Nutzergenauigkeit).

		Referenz											
		Fi	**Ki**	**Lä**	**Dgl**	**SZy**	**Bu**	**Ei**	**HBu**	**Bi**	**SEr**	**Σ**	**NG**
Klassifiziert als	**Fi**	183	17	19	29	0	0	0	0	0	0	248	**0,738**
	Ki	12	205	20	8	0	0	0	0	0	0	245	**0,837**
	Lä	13	11	79	5	2	0	0	0	0	0	110	**0,718**
	Dgl	17	2	3	136	0	0	0	0	0	1	159	**0,855**
	SZy	0	0	0	0	34	0	0	0	1	3	38	**0,895**
	Bu	0	0	0	0	0	237	0	18	1	2	258	**0,919**
	Ei	1	0	1	0	0	1	144	10	0	4	161	**0,894**
	HBu	0	0	0	0	0	9	4	44	1	0	58	**0,759**
	Bi	0	0	0	0	0	0	0	5	80	2	87	**0,920**
	SEr	0	0	0	0	6	0	4	4	3	84	101	**0,832**
	Σ	**226**	**235**	**122**	**178**	**42**	**247**	**152**	**81**	**86**	**96**	**1465**	
	PG	**0,810**	**0,872**	**0,648**	**0,764**	**0,810**	**0,960**	**0,947**	**0,543**	**0,930**	**0,875**		**0,837**

Tab. 3: Klassifikationsmatrix für die Klassifikation mit den 4 konventionellen Kanälen Blue, Green, Red,NearInfrared1(Modalwerteaus500Bootstrap-Wiederholungen,PG=Produzentengenauigkeit, NG=Nutzergenauigkeit).

		Referenz											
		Fi	**Ki**	**Lä**	**Dgl**	**SZy**	**Bu**	**Ei**	**HBu**	**Bi**	**SEr**	**Σ**	**NG**
Klassifiziert als	**Fi**	167	12	19	33	0	0	0	0	0	0	**231**	**0,723**
	Ki	12	205	22	10	0	0	0	0	0	0	**249**	**0,823**
	Lä	10	14	72	11	1	0	0	1	0	0	**109**	**0,661**
	Dgl	36	4	8	120	0	0	0	0	0	0	**168**	**0,714**
	SZy	1	0	0	3	37	0	0	0	1	4	**46**	**0,804**
	Bu	0	0	0	0	0	228	0	32	0	2	**262**	**0,870**
	Ei	0	0	1	0	0	1	143	18	0	3	**166**	**0,861**
	HBu	0	0	0	0	0	18	6	21	1	0	**46**	**0,457**
	Bi	0	0	0	0	0	0	0	5	75	6	**86**	**0,872**
	SEr	0	0	0	1	4	0	3	4	9	81	**102**	**0,794**
	Σ	**226**	**235**	**122**	**178**	**42**	**247**	**152**	**81**	**86**	**96**	**1465**	
	PG	**0,739**	**0,872**	**0,590**	**0,674**	**0,881**	**0,923**	**0,941**	**0,259**	**0,872**	**0,844**		**0,784**

klassifiziert werden, waren es mit 4 Kanälen nur noch 21. Auch bei Fichte ist ein deutlicher Anstieg der fälschlich als Douglasien klassifizierten Referenzflächen zu beobachten (36 Fälle bei 4 Kanälen gegenüber 17 bei 8 Kanälen).

Eine Zusammenfassung aller Ergebnisse ist in Tab. 4 dargestellt. Die Genauigkeitsmaße aus den oben angeführten Klassifikationsmatrizen werden ergänzt durch die Klassifikationsergebnisse für die Trennung der 4 Hauptbaumarten Gemeine Fichte, Weiß-Kiefer, Rot-Buche und Stiel-Eiche (im Vergleich zu den insgesamt 10 Baumarten). Außerdem werden die Mittelwerte (MW) der Gesamtgenauigkeiten und deren Standardabweichungen (SD) aus den 500 Bootstrap-Wiederholungen aufgelistet.

Werden statt der 8 verfügbaren nur die 4 konventionellen Kanäle für die Klassifikation verwendet, verschlechtert sich das Klassifikationsergebnis deutlich. Bei Betrachtung der mittleren Gesamtgenauigkeiten der 500 Einzelklassifikationen ist eine signifikante Abnahme von 83,0 % auf 78,2 % (T-Test, t=-54,05, df=998, P<0,01) feststellbar. Beschränkt sich die Klassifikation hingegen auf die 4 Hauptbaumarten, zeigen die 4 zusätzlichen Kanäle keine Verbesserung. So liegt im Untersuchungsgebiet die mittlere Gesamtgenauigkeit bei der Unterscheidung der 4 Hauptbaumarten Fichte, Kiefer, Buche und Eiche bei Verwendung von 4 Kanälen mit 94,8 % sogar geringfügig über den 94,5 % bei Verwendung von 8 Kanälen (T-Test, t=4,91, df=998, P<0,01).

Tab. 4: Übersicht der Klassifikationsergebnisse für die 10 Baumarten (Fi, Ki, Lä, Dgl, SZy, Bu, Ei, HBu, Bi, SEr) bzw. 4 Baumarten (Fi, Ki, Bu, Ei) mit 8 bzw. 4 Kanälen (beste Klassifikationsergebnisse jeweils in Fettschrift; Mittelwerte, Standardabweichungen und Modalwerte aus 500 Bootstrap-Wiederholungen).

Bewertungskriterium	10 Baumarten		4 Baumarten	
	8 Kanäle	4 Kanäle	8 Kanäle	4 Kanäle
Mittelwert der Gesamtgenauigkeiten	**0,830**	0,782	0,945	**0,948**
Standardabweichung der Gesamtgenauigkeiten	0,013	0,015	0,011	0,011
Gesamtgenauigkeit aus Modalwerten (Klassifikationsmatrix)	**0,837**	0,784	0,945	**0,950**
Kappa aus Modalwerten (Klassifikationsmatrix)	0,814	0,754	0,926	0,933

Die in Tab. 4 aufgeführten Mittelwerte der Gesamtgenauigkeiten und deren Standardabweichungen aus den 500 Bootstrap-Wiederholungen berücksichtigen die Streuung (unterschiedliche Klassifikationen) innerhalb der Referenzflächen und liegen erwartungsgemäß etwas unter den Werten aus den Klassifikationsmatrizen, die aus den Modalwerten pro Referenzfläche erstellt wurden. Die Unterschiede zwischen den beiden Gesamtgenauigkeitsmaßen fallen aber über alle Modelle nur sehr gering aus und bescheinigen den Modellen damit eine hohe Robustheit gegenüber den verwendeten Trainingsdaten.

3.4 *Klassifikation einer Testfläche*

In dem Bestandesteil, der als Testfläche ausgewählt wurde, kommen laut Taxation im nördlichen Teil die Baumarten Stiel-Eiche, Schwarz-Erle, Hainbuche, Rot-Buche und einzelne Europäische Lärchen vor, während im Süden Weiß-Kiefer und Gemeine Fichte dominieren. Das Klassifizierungsergebnis (10 Baumarten, 8 Kanäle) stimmt mit dieser Beschreibung weitgehend überein. Die Vergesellschaftung von Stiel-Eiche mit Hainbuche ist plausibel (Eichen-Hainbuchenwald), jedoch erscheint der Anteil der Hainbuche etwas höher als es die Taxationsdaten erwarten lassen. Die Baumart Douglasie wird in der Bestandesbeschreibung nicht erwähnt. Bei den klassifizierten Douglasien handelt es sich daher vermutlich um Fehlklassifikationen.

3.5 *Trennkraft der einzelnen Spektralkanäle*

Im Zuge der LDA wurden zwei Kennwerte berechnet, die Aufschluss über die Trennkraft

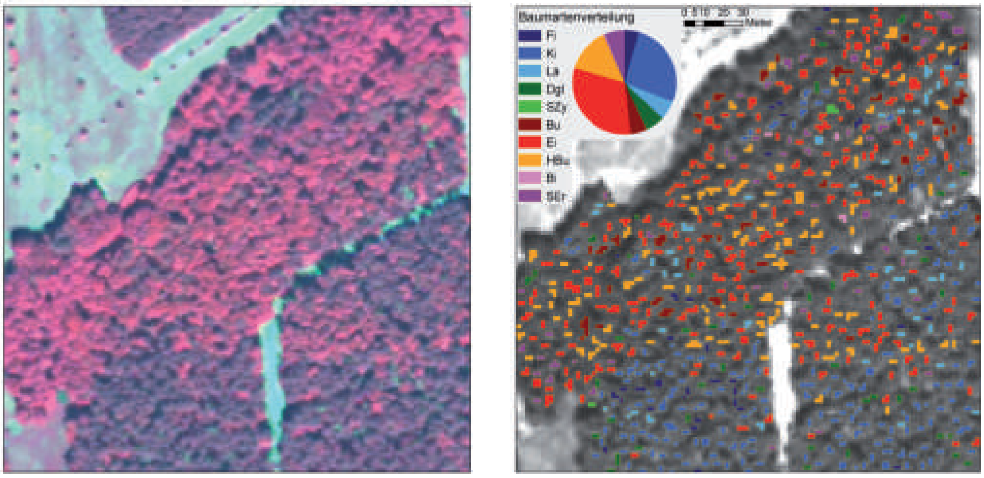


Abb. 6: Klassifizierung der Kronenpolygone der Testfläche: links: Bestand in CIR-Darstellung (Pixelgröße 0,5 m), rechts: mittels LDA klassifizierte Kronenpolygone.

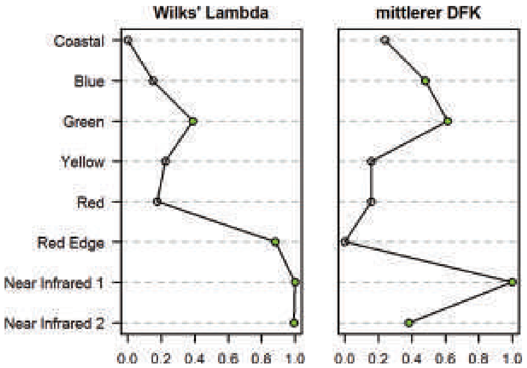


Abb. 7: Normalisierte Wichtigkeiten (0 = geringste, 1 = höchste Wichtigkeit) der einzelnen Kanäle für die Klassifizierung der 10 Baumarten. Die jeweils 4 wichtigsten sind grün hervorgehoben.

der einzelnen Spektralkanäle liefern (Abb. 7). Die Wilks'-Lambda-Werte der univariaten ANOVA geben an, wie gut die einzelnen Kanäle die 10 Baumarten trennen können. Mit den geringsten Wilks'-Lambda-Werten trennen die Kanäle Near Infrared 1 und 2, bzw. auch der Red Edge Kanal univariat deutlich besser als die Kanäle im Wellenlängenbereich des sichtbaren Lichts. Bewertet man die Trennkraft auf Basis der standardisierten und gemittelten Diskriminanzkoeffizienten aus der LDA, ergibt sich folgende Reihung: Near Infrared 1, Green, Blue und Near Infrared 2. Interessanterweise befinden sich damit 2 Kanäle aus dem Wellenlängenbereich des sichtbaren Lichts unter den 3 wichtigsten Kanälen. Für den direkten Vergleich der einzelnen Maße wurden die Kenngrößen auf den Wertebereich von 0 geringste Wichtigkeit) bis 1 (höchste Wichtigkeit) normalisiert. Die Ergebnisse in Abb. 7 zeigen eine gute Übereinstimmung zwischen den Reihungen anhand der beiden Wichtigkeitsmaße.

Die Kanäle Near Infrared 1, Near Infrared 2 und Green sind bei beiden Wichtigkeitsmaßen unter den 4 wichtigsten Kanälen. Bei den Mittleren Diskriminanzkoeffizienten fallen Red Edge und Near Infrared 2 jedoch verglichen mit der Reihung nach den Wilks' Lambda Werten stark gegenüber dem Kanal Near Infrared 1 ab. Dies ist vermutlich durch hohe Korrelationen zwischen den genannten Kanälen zu erklären. Von den 4 neuen Kanälen (Coastal, Yellow, Red Edge und Near Infrared 2) sind laut Wilks' Lambda nur 2 unter den 4 wichtigsten Kanälen (Red Edge und Near Infrared 2); unter Verwendung des Mittleren Diskriminanzkoeffizienten ist es nur der Near Infrared 2 Kanal.

4 Diskussion

Die verwendeten WorldView-2 Satellitendaten zeigten eine gute Eignung für die Trennung der untersuchten Baumarten im Testgebiet. Bei der Klassifikation der 4 Hauptbaumarten (Fi, Ki, Bu, Ei) zeigten bereits die 4 konventionellen Kanäle (Blau, Grün, Rot und Near Infrared 1) sehr gute Ergebnisse, welche durch die Mitverwendung der 4 neuen Kanäle (Coastal, Yellow, Red Edge und Near Infrared 2) nicht mehr verbessert wurden. Bezüglich der 6 Nebenbaumarten konnten die zusätzlichen Kanäle jedoch eine deutliche Verbesserung des Klassifikationsergebnisses bewirken. So konnte die Gesamtgenauigkeit der Klassifikation der 10 untersuchten Baumarten von 78 % (4 Kanäle) auf 84 % (8 Kanäle) gesteigert werden. Die geringen Genauigkeitsunterschiede bei Verwendung von 4 gegenüber 8 Kanälen bei der Analyse weniger (Haupt-) Baumarten decken sich mit den Ergebnissen der Studie von CHEN (2011). Dieser erzielte bei der Klassifikation zweier Baumarten mit verschiedenen Klassifikationsalgorithmen (u.a. LDA) mit 4 Kanälen immer annähernd die gleichen Ergebnisse wie mit 8 Kanälen.

Die Detailergebnisse zeigten, dass die Unterscheidung zwischen Nadel- und Laubbäumen zu 99 % richtig erfolgte. Die erzielten Nutzergenauigkeiten der einzelnen Baumarten lagen zwischen 71 % und 92 %, die Produzentengenauigkeiten zwischen 54 % und 96 %, wobei die geringsten Werte bei Hainbuche bzw. Lärche festgestellt wurden. Gerade bei diesen Baumarten ist vermutlich ein Teil der Fehlklassifikationen auf Fehler in den Referenzdaten (falsche Baumartenzuordnung) zurückzuführen: Beide Baumarten bilden im Untersuchungsgebiet keine Reinbestände, vielmehr handelt es sich um begleitende (Hainbuche) oder einzeln eingesprengte (Lärche) Baumarten. Dadurch kann es bei der Referenzdatengewinnung zu fehlerhaften Zuordnungen gekommen sein. Die Tatsache, dass einzelne Baumarten bei sonst sehr ausgeglichenem Klassifikationsergebnis deutlich geringere Genauigkeiten erzielen, ist bei nahezu allen Studien, die eine größere Anzahl an Arten untersuchen, zu beobachten (CARLEER & WOLFF 2004, VOSS & SUGUMARAN 2008, WASER et al. 2010, 2011, JONES et al. 2010, HAMDAN 2010).

Hervorzuheben sind auch die generell geringeren Klassifikationsgenauigkeiten bei den Nadelbaumarten im Vergleich zu den Laubbaumarten. Neben den geringeren spektralen Unterschieden zwischen den Nadelbäumen könnte die höhere Fehlerrate teilweise auch durch die Kronenform begründet sein: Durch die für Nadelbäume typische, schmale Kronenform sind bei Bilddaten mit 2 m Pixelgröße nur wenige Pixel gut beleuchtet und baumartenrein. Dadurch fielen bei der Referenzdatengewinnung deutlich weniger Pixel pro Beispielkrone an, wodurch einzelne Mischpixel einen größeren Einfluss auf den Mittelwert des Referenzpolygons ausüben. Durch die geringe Nadel- und Kronentransmissivität sind Schatteneffekte zudem stärker ausgeprägt als bei Laubbäumen (SCHLERF & ATZBERGER 2012).

Ähnlichen Studien in temperierten Wäldern zeigten bei Verwendung von Sensoren unterschiedlicher spektraler und räumlicher Auflösung Gesamtgenauigkeiten zwischen 45 % und 96 %. Die besten Werte wurden in der Regel erzielt, wenn nur wenige Baumarten, wie z.B. 3 Arten bei HEIKKINEN et al. (2011), untersucht oder zusätzliche Eingangsdaten, wie z.B. LiDAR bei HOLMGREN et al. (2008), verwendet wurden. In der vorliegenden Studie konnten die 4 Hauptbaumarten Gemeine Fichte, Weiß-Kiefer, Rot-Buche und Stiel-Eiche mit einer Gesamtgenauigkeit von 95 % klassifiziert werden, unabhängig davon ob nur die 4 konventionellen oder alle 8 Kanäle zum Einsatz kamen. Dieses Ergebnis liegt über vergleichbaren Studien, in denen ebenfalls 4 bis 5 Baumarten rein spektral klassifiziert wurden. Auch bei der Klassifikation der 10 Baumarten liegen die erzielten Gesamtgenauigkeiten mit 84 % im Spitzenfeld der Ergebnisse vergleichbarer Studien (CARLEER & WOLFF 2004, WASER et al. 2010, JONES et al. 2010). Die Gründe dafür liegen vermutlich einerseits in den spektralen und geometrischen Eigenschaften der verwendeten WorldView-2 Daten und andrerseits in der Klassifizierung ausschließlich gut beleuchteter Kronenteile. Durch diese Vorgangsweise kann der Einfluss von Beleuchtungsunterschieden auf das Klassifikationsergebnis deutlich reduziert werden.

Die Abgrenzung der beleuchteten Kronenteile erfolgte manuell. Für die großflächige Umsetzung der Baumartenkartierung ist es notwendig, diesen Arbeitsschritt zu automatisieren. Dies wurde bereits in mehreren Studien z.B. durch die Kombination von Spektraldaten mit 3D-Informationen erfolgreich bewerkstelligt (z.B. VOSS & SUGUMARAN 2008, HOLMGREN et al. 2008, DALPONTE et al. 2008, STRAUB et al. 2010, JONES et al. 2010, WASER et al. 2011), wobei die Spektraldaten in erster Linie der Artunterscheidung dienten. Die 3D-Informationen z.B. in Form von Digitalen Oberflächenmodellen wurden hingegen vor allem für die Abgrenzung einzelner Baumindividuen verwendet. Für diese Zwecke sind LiDAR-Daten gut nutzbar (z.B. HEINZEL et al. 2008, HOLMGREN et al. 2008, STRAUB et al. 2010, WASER et al. 2011). Selbstverständlich können geeignete Oberflächenmodelle auch aus dem Bildmaterial selbst erstellt werden (z.B. HIRSCHMUGL et al. 2007). Bei der Verwendung von LiDAR Daten könnten diese neben den Spektralkanälen auch als Eingangsdaten für die Klassifikation dienen und zu einer weiteren Verbesserung des Ergebnisses führen (VOSS & SUGUMARAN 2008, HOLMGREN et al. 2008, JONES et al. 2010).

Andere Ansätze beschränken sich auf nur eine Datenquelle und verwenden verschiedene Segmentierungsalgorithmen zur Einzelbaumabgrenzung. Diese reichen von vorgefertigten Verfahren z.B. mit Definiens (SRIDHARAN

2011), über weiterentwickelte Watershed Segmentation Ansätze (z.B. Kanda et al. 2004, Wang et al. 2004) bis hin zu speziell entwickelten Verfahren (z.B. Brandtberg 1999, Atzberger & Schlerf 2002, Culvenor 2002, Erikson 2003, Leckie et al. 2005a). Ein ausführlicher Überblick über die unterschiedlichen Verfahren findet sich unter anderem in Culvenor (2003) und Wolf & Heipke (2007). Alle Studien verwenden hochauflösende Bilddaten, vorzugsweise Luftbilder. Mit dem panchromatischen Kanal mit einer Auflösung von 0,5 m sollten auch WorldView-2 Daten für derartige Verfahren geeignet sein.

5 Schlussfolgerungen und Ausblick

Im Vergleich zu Luftbildern bieten WorldView-2 Daten aufgrund der höheren spektralen Auflösung bei gleichzeitig geringeren Blickrichtungseffekten deutliche Vorteile. Nachteilig sind die geringere Flexibilität bezüglich des Aufnahmezeitpunkts und die geringere räumliche Auflösung. Gegenüber anderen satellitengestützten Sensoren, wie z.B. QuickBird, kann WorldView-2 vor allem durch die höhere spektrale und räumliche Auflösung punkten.

Da der Erwerb aller 8 anstelle der 4 konventionellen Kanäle zu einer Verdoppelung der Datenkosten führt, ist für jede konkrete Anwendung eine Kosten-/Nutzenanalyse sinnvoll. Wie in dieser Studie gezeigt werden konnte, reichen bei Klassifikationsaufgaben mit wenigen Zielklassen die konventionellen Kanäle aus, während sich die 4 neuen Kanäle (Coastal, Yellow, Red Edge und Near Infrared 2) bei komplexeren Fragestellungen als vorteilhaft erwiesen haben.

Der vorgestellte Ansatz beschränkt sich auf die Analyse der beleuchteten Kronenteile, welche manuell abgegrenzt wurden. Ziel unserer weiteren Forschungsarbeit ist die automatisierte, großflächige Umsetzung der Methode. Soweit möglich soll dabei die Einzelbaumabgrenzung rein auf Basis des WorldView-2 Bildes ohne Zuhilfenahme zusätzlicher Daten, wie z.B. LiDAR Daten, erfolgen, um im Sinne einer praxistauglichen Anwendung den Datenaufwand möglichst gering zu halten.

Neben der großflächigen Anwendung ist eine Ausweitung der Analyse in Bezug auf weitere Baumarten und die Berücksichtigung des Baumalters geplant. Im Zuge dessen werden für alle relevanten Baumarten Referenzdaten vor Ort erhoben. Damit können Unsicherheiten, wie sie im Zusammenhang mit Taxationsdaten speziell bei sehr heterogenen, gemischten Beständen auftreten, in den weiteren Untersuchungen vermieden werden. Auch ein direkter Vergleich der Klassifikation für das gleiche Untersuchungsgebiet mit einem weiteren Sensor, vorzugsweise mit Luftbildern, ist geplant.

In der vorliegenden Studie wurden lediglich LDA-basierte Klassifikationsergebnisse vorgestellt. Nicht gezeigt wurden parallel dazu durchgeführte Untersuchungen mit Hilfe des Algorithmus Random Forests (Breiman 2001), welche in jeder Hinsicht sehr ähnliche Ergebnisse ergaben. Dies betrifft sowohl den möglichen Informationsgewinn durch die Verwendung der zusätzlichen WorldView-2 Kanäle als auch die Analyse der spektralen Trennkraft der einzelnen Kanäle.

Im spektralen Merkmalsraum abgeleitete Texturmerkmale (Atzberger 2003) wurden in der vorliegenden Studie nicht untersucht, könnten aber Genauigkeitssteigerungen bewirken. Ebenso wäre die Optimierung der Klassifikation durch die Verwendung eines späteren Aufnahmezeitpunktes innerhalb der Vegetationsperiode oder eines multi-saisonalen Datensatzes desselben Sensors denkbar (Hill et al. 2010).

Danksagung

Die Arbeiten wurden zum Teil durch den FFG Innovationscheck finanziert. Die Autoren bedanken sich bei Günther Bronner und Bernhard Pfandl von der Fa. Umweltdata GmbH für die Bereitstellung der WorldView-2 Szene und der Referenzdaten. Matteo Mattiuzzi danken wir für die Unterstützung bei der Umsetzung der Analysen mit dem Softwarepaket R. Schließlich bedanken wir uns bei den anonymen Gutachtern für ihre wertvollen Anregungen zum Manuskript.

Referenzen

Albertz, J., 2009: Einführung in die Fernerkundung: Grundlagen der Interpretation von Luft- und Satellitenbildern. – **4.** aktualisierte Auflage, 254 S., Wissenschaftliche Buchgesellschaft, Darmstadt.

Atzberger, C. & Schlerf, M., 2002: Automatisierte Bestimmung der Bestockungsdichte in Nadelwäldern aus räumlich hochauflösenden Ortholuftbildern. – PFG **2002** (3): 171–180.

Atzberger, C., 2003: Möglichkeiten und Grenzen der fernerkundlichen Bestimmung biophysikalischer Vegetationsparameter mittels physikalisch basierter Reflexionsmodelle. – PFG **2003** (3): 51–61.

Backhaus, K., Erichson, B., Plinke, W. & Weiber, R., 2008: Multivariate Analysemethoden - Eine anwendungsorientierte Einführung. – **12th** ed., 575 S., Springer, Berlin.

Brandtberg, T., 1999: Automatic individual tree based analysis of high spatial resolution aerial images on naturally regenerated boreal forests. – Canadian Journal of Forest Research **29** (10): 1464–1478.

Breiman, L., 2001: Random forests. – Machine learning **45** (1): 5–32.

Brosinger, F., 2010: Erwartungen der Forstverwaltung an die Fernerkundung. Der gepixelte Wald – forstliche Fernerkundung vor dem Hintergrund aktueller Entwicklungen in Umwelt und Technik. – Fachtagung am Zentrum Wald-Forst-Holz in Freising-Weihenstephan 2008, Frank, München.

Buck, G., Seitz, R. & Troycke, A., 2010: Fernerkundung an der Bayerischen Landesanstalt für Wald und Forstwirtschaft (LWF) – Umsetzung von Forschungsergebnissen in die forstliche Praxis. – PFG **2010** (4): 295–303.

Carleer, A. & Wolff, E., 2004: Exploitation of Very High Resolution Satellite Data for Tree Species Identification. – Photogrammetric Engineering and Remote Sensing **70** (1): 135–140.

Chen, Q., 2011: Comparison of Worldview-2 and IKONOS-2 imagery for identifying tree species in the habitat of an endangered bird species in Hawaii. 8-Band Research Challenge, DigitalGlobe, http://dgl.us.neolane.net/res/img/f3e1241921965c60aeae869dd949bb48.pdf (8.3.2011).

Clark, M.L., Roberts, D.A. & Clark, D.B., 2005: Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. – Remote Sensing of Environment **96** (3–4): 375–398.

Collin, A. & Planes, S., 2011: What is the value added of 4 bands within the submetric remote sensing of tropical coastscape? QuickBird-2 vs WorldView-2. – IEEE International Geoscience and Remote Sensing Symposium (IGARSS): 2165–2168, Vancouver, BC, Canada.

Congalton, R. & Green, K., 1999: Assessing the accuracy of remotely sensed data principles and practices. – 137 S., Lewis, Boca Raton.

Culvenor, D.S., 2002: TIDA: An algorithm for the delineation of tree crowns in high spatial resolution remotely sensed imagery. – Computers and Geosciences **28** (1): 33–44.

Culvenor, D.S., 2003: Extracting individual tree information: a survey of techniques for high spatial resolution imagery. Remote Sensing of Forest Environments: Concepts and Case Studies. – **1st** ed., 255–277, Kluwer Academic Publishers, Boston, Dordrecht, London.

Dalponte, M., Bruzzone, L. & Gianelle, D., 2008: Fusion of Hyperspectral and LIDAR Remote Sensing Data for Classification of Complex Forest Areas. – IEEE Transactions on Geoscience and Remote Sensing **46:** 1416–1427.

DigitalGlobe, 2009: White Paper – The Benefits of the 8 Spectral Bands of WorldView-2, http://worldview2.digitalglobe.com/docs/WorldView-2_8-Band_Applications_Whitepaper.pdf (12.5.2011).

Efron, B. & Tibshirani, R., 1993: An Introduction to the Bootstrap. – Chapman and Hall, New York.

Erikson, M., 2003: Segmentation of individual tree crowns in colour aerial photographs using region growing supported by fuzzy rules. – Canadian Journal of Forest Research **33** (8): 1557–1563.

Falkowski, M.J., Wulder, M.A., White, J.C. & Gillis, M.D., 2009: Supporting large-area, sample-based forest inventories with very high spatial resolution satellite imagery. – Progress in Physical Geography **33** (3): 403–423.

Greenberg, J.A., Dobrowski, S.Z., Ramirez, C.M., Tull, J.L. & Ustin, S.L., 2006: A bottom-up approach to vegetation mapping of the Lake Tahoe Basin using hyperspatial image analysis. – Photogrammetric Engineering and Remote Sensing **72** (5): 581–589.

Hamdan, O., 2010: Commercial Timber Tree Species Identification Using Multispectral WorldView-2 Data. 8-Band Research Challenge, DigitalGlobe, http://dgl.us.neolane.net/res/img/e809c430f196bad7ffd19e3ed53bdac2.pdf (8.3.2011).

Heikkinen, V., Korpela, I., Tokola, T., Honkavaara, E. & Parkkinen, J., 2011: An SVM Classification of Tree Species Radiometric Signatures Based on the Leica ADS40 Sensor. – IEEE Transactions on Geoscience and Remote Sensing **49** (11): 4539–4551.

Heinzel, J.N., Weinacker, H. & Koch, B., 2008: Full automatic detection of tree species based on delineated single tree crowns – a data fusion approach for airborne laser scanning data and aerial photographs. – SilviLaser **8th** international conference on LiDAR applicatons in forest assessment and inventory: 76–85, Edinburgh UK.

Hildebrandt, G., 1996: Fernerkundung und Luftbildmessung für Forstwirtschaft, Vegetationskartierung und Landschaftsökologie. – 1. Aufl., 676 S., Herbert Wichmann Verlag.

Hill, R.A., Wilson, A.K., George, M. & Hinsley, S.A., 2010: Mapping tree species in temperate deciduous woodland using time-series multispectral data. – Applied Vegetation Science **13:** 86–99.

Hirschmugl, M., Ofner, M., Raggam, J. & Schardt, M., 2007: Single tree detection in very high resolution remote sensing data. – Remote Sensing of Environment **110:** 533–544.

Holmgren, J., Persson, Å. & Söderman, U., 2008: Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. – International Journal of Remote Sensing **29** (5): 1537–1552.

Hosgood, B., Jacquemoud, S., Andreoli, G., Verdebout, J., Pedrini, G. & Schmuck, G., 1994: Leaf Optical Properties EXperiment 93 (LOPEX93). – European Commission, Joint Research Centre, Institute for Remote Sensing Applications, Report EUR **16095 EN:** 11.

Huang, H. & Cao, B., 2011: Experiment on extracting forest canopy height from Worldview-2. – IEEE Fuzzy Systems and Knowledge Discovery (FSKD): 2614–2617, Shanghai.

Jones, T.G., Coops, N.C. & Sharma, T., 2010: Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. – Remote Sensing of Environment **114** (12): 2841–2852.

Kadro, A., 1981: Untersuchung der spektralen Reflexionseigenschaften verschiedener Vegetationsbestände. – Dissertation, Universität Freiburg.

Kanda, F., Kubo, M. & Muramoto, K., 2004: Watershed segmentation and classification of tree species using high resolution forest imagery. – IEEE International Geoscience and Remote Sensing Symposium (IGARSS): 3822–3825.

Kilian, W., Müller, F. & Starlinger, F., 1994: Die forstlichen Wuchsgebiete Österreichs – Eine Naturraumgliederung nach waldökologischen Gesichtspunkten. – BFW, Wien.

Koukal, T. & Atzberger, C., 2012: Potential of Multi-Angular Data Derived From a Digital Aerial Frame Camera for Forest Classification. – IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **5** (1): 30–43.

Leckie, D.G., Gougeon, F.A., Tinis, S., Nelson, T., Burnett, C.N. & Paradine, D., 2005a: Automated tree recognition in old growth conifer stands with high resolution digital imagery. – Remote Sensing of Environment **94** (3): 311–326.

Leckie, D.G., Tinis, S., Nelson, T., Burnett, C., Gougeon, F.A., Cloney, E. & Paradine, D., 2005b: Issues in species classification of trees in old growth conifer stands. – Canadian Journal of Remote Sensing **31** (2): 175–190.

Ligges, U. & Mächler, M., 2003: Scatterplot3d – an R package for Visualizing Multivariate Data. – Journal of Statistical Software **8** (11): 1–20.

Marchisio, G., Pacifici, F. & Padwick, C., 2010: On the relative predictive value of the new spectral bands in the WorldView-2 sensor. – IEEE International Geoscience and Remote Sensing Symposium (IGARSS): 2723–2726, Honolulu, Hawaii, USA.

McDermid, G.J., Hall, R.J., Sanchez-Azofeifa, G.A., Franklin, S.E., Stenhouse, G.B., Kobliuk, T. & LeDrew, E.F., 2009: Remote sensing and forest inventory for wildlife habitat assessment. – Forest Ecology and Management **257** (11): 2262–2269.

Olsson, H., Schneider, W. & Koukal, T., 2008: Preface: 3D Remote Sensing in Forestry. – International Journal of Remote Sensing **29** (5): 1239–1242.

Ozdemir, I. & Karnieli, A., 2011: Predicting forest structural parameters using the image texture derived from WorldView-2 multispectral imagery in a dryland forest, Israel. – International Journal of Applied Earth Observation and Geoinformation **13** (5): 701–710.

Padwick, C., Deskevich, M., Pacifici, F. & Smallwood, S., 2010: WorldView-2 Pan-Sharpening. – ASPRS 2010.

R Development Core Team, 2011: R: A Language and Environment for Statistical Computing. – R Foundation for Statistical Computing, Vienna.

Schlerf, M., Hill, J., Bärisch, S. & Atzberger, C., 2003: Einfluss der spektralen und räumlichen Auflösung von Fernerkundungsdaten bei der Nadelwaldklassifikation. – PFG **2003** (1): 25–34.

Schlerf, M. & Atzberger, C., 2012: Vegetation Structure Retrieval in Beech and Spruce Forests Using Spectrodirectional Satellite Data. – IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **5** (1): 8–17.

Sridharan, H., 2011: Multi-Level comparison of WorldView-2 8-Band and AISA Hyperspectral imageries for urban forest classification. 8-Band Research Challenge. – DigitalGlobe, http://dgl.

us.neolane.net/res/img/019954e1ad60acaed6b05b88e7319030.pdf (23.11.2011).

Straub, C., Weinacker, H. & Koch, B., 2010: A comparison of different methods for forest resource estimation using information from airborne laser scanning and CIR orthophotos. – European Journal of Forest Research **129** (6): 1069–1080.

Updike, T. & Comp, C., 2010: Radiometric Use of WorldView-2 Imagery. – DigitalGlobe, Longmont, http://www.digitalglobe.com/downloads/Radiometric_Use_of_WorldView-2_Imagery.pdf (7.5.2012).

Venables, W.N. & Ripley, B.D., 2002: Modern applied statistics with S. – **4th** ed., 495 S., Springer, New York.

Voss, M. & Sugumaran, R., 2008: Seasonal effect on tree species classification in an urban environment using hyperspectral data, LiDAR, and an object-oriented approach. – Sensors **8** (5): 3020–3036.

Wang, L., Gong, P. & Biging, G.S., 2004: Individual Tree-Crown Delineation and Treetop Detection in High-Spatial-Resolution Aerial Imagery. – Photogrammetric Engineering and Remote Sensing **70** (3): 351–357.

Waser, L.T., Klonus, S., Ehlers, M., Küchler, M. & Jung, A., 2010: Potential of Digital Sensors for Land Cover and Tree Species Classifications – A Case Study in the Framework of the DGPF-Project. – PFG **2010** (2): 141–156.

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): 76–85.

Wolf, B.-M. & Heipke, C., 2007: Automatic extraction and delineation of single trees from remote sensing data. – Machine Vision and Applications **18** (5): 317–330.

Wulder, M.A., Hall, R.J., Coops, N.C. & Franklin, S.E., 2004: High spatial resolution remotely sensed data for ecosystem characterization. – BioScience **54** (6): 511–521.

Adressen der Autoren:

Markus Immitzer MSc., Univ. Prof. Dr. Clement Atzberger & Dr. Tatjana Koukal, Universität für Bodenkultur, Institut für Vermessung, Fernerkundung und Landinformation, Peter-Jordan-Straße 82, A-1190 Wien, Tel.: +43-1-47654-5100, Fax: -5142, e-mail: vorname.nachname@boku.ac.at

Manuskript eingereicht: April 2012
Angenommen: Mai 2012

PFG 2012 / 5, 0589–0602
Stuttgart, October 2012

Article

Scale-specific Hyperspectral Remote Sensing Approach in Environmental Research

Angela Lausch, Leipzig, Marion Pause, Tübingen, Ines Merbach & Sarah Gwillym-Margianto, Halle, Karsten Schulz, München, Steffen Zacharias & Ralf Seppelt, Leipzig

Keywords: imaging hyperspectral remote sensing, multi-scale analyses, vegetation monitoring

Summary: Hyperspectral remote-sensing data can contribute significantly to data analysis in research, opening up a wide spectrum for fields of application due to geometrical as well as spectral characteristics, e.g. in water status analysis, in the classification of vegetation types, in the classification of physical-biochemical vegetation parameters, in classifying soil composition and structure, and in determining large-scale soil contamination. Hence, there is a tremendous demand for hyperspectral information. However the use of commercial hyperspectral data is associated with a number of problems and a great deal of time and effort is required for research using hyperspectral data that spans different spatial and/or hierarchical as well as temporal scales. As a result few investigations have been conducted on the causal relationships between imaging hyperspectral signals and meaningful vegetation variables over a longer monitoring period. At the Helmholtz Centre for Environmental Research (UFZ) Leipzig a scale-specific hyperspectral remote sensing based on the sensors AISA-EAGLE (400–970 nm) and AISA-HAWK (970–2500 nm) has been set up. On three different scales (plot, local and regional) intensive investigations are being carried out on the spatio-temporal responses of biophysical and biochemical state variables of vegetation, soil and water compared to the hyperspectral response. This paper introduces and discusses the scale approach and demonstrates some preliminary examples from its implementation.

Zusammenfassung: Hyperspektraldaten stellen für die Forschung eine sehr bedeutsame Auswertegrundlage dar, da sie aufgrund ihrer geometrischen als auch spektralen Eigenschaften eine Vielzahl unterschiedlicher Anwendungsgebiete, z.B. Gewässerzustandserfassung, Vegetationsklassifizierungen, Charakterisierung physikalisch-biochemischer Vegetationsparameter, Strukturierung und Zusammensetzung des Bodens, Erfassung von großflächigen Bodenkontaminationen, eröffnen. Es besteht somit ein sehr hoher Bedarf an Hyperspektralinformationen. Der Einsatz von kommerziellen Hyperspektraldaten ist jedoch mit einer Vielzahl von Problemen verbunden. So sind Forschungen hinsichtlich unterschiedlicher räumlich/hierarchischer als auch zeitlicher Skalen mit Hyperspektraldaten nur sehr schwer möglich, andererseits existieren nur wenige Untersuchungen zu kausalen Zusammenhängen zwischen abbildenden Hyperspektralsignalen und gesuchten Vegetationsvariablen über einen langen Monitoringzeitraum. Am Helmholtz Zentrum für Umweltforschung (UFZ) Leipzig wurde eine skalenspezifische hyperspektrale Fernerkundung auf Grundlage der Sensoren AISA-EAGLE (400–970 nm) und AISA-HAWK (970–2500 nm) etabliert. In drei unterschiedlichen Maßstabsbereichen (Grundstück (plot), kommunal (local) und regional) werden intensive Untersuchungen zum raum-zeitlichen Verhalten von biophysikalischen und biochemischen Zustandsgrößen von Vegetation, Boden und Wasser gegenüber hyperspectral response durchgeführt. Im Artikel wird der Skalenansatz vorgestellt, diskutiert und erste Umsetzungsbeispiele gezeigt.

---

1 Introduction

The application of optical remote sensing data from airborne and satellite sensors has been well established in environmental research for more than three decades (Schaepmen et al. 2009). It provides a state-of-the-art method for a variety of monitoring issues requiring

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0141
1432-8364/12/0141 $ 3.25

spatial information of the Earth's surface. Depending on the specific data product and ancillary data, improvements in ecological, hydrological and climate modelling have been possible over a wide range of spatial scales. Thereby, the reliability of final information or quantifications depends mainly on sensor-specific limitations in terms of their spectral and spatial characteristics. Modern airborne imaging hyperspectral sensors open up many new different fields of applications thanks to their high geometric (< 5 m) and spectral (~ 5 nm) characteristics, e.g. monitoring the state of aquatic ecosystems, quantifying biodiversity, retrieving biophysical-biochemical vegetation parameters, assessing soil structure and composition, recording soil contamination over larger areas. With the launch of the German hyperspectral satellite mission EnMAP (Environmental Mapping and Analyses Program) that is foreseen for 2015, a new imaging spectrometer data format in terms of its spectral (6.5–10 nm) and geometric (30 m × 30 m spatial resolution) characteristics will be available (Stuffler et al. 2009). The EnMAP data product on the regional scale and the possibility of available multi-temporal data will enable monitoring issues to be realized on the regional scale.

Knowledge about the retrieval of vegetation parameters, e.g. leaf area index, pigments, photosynthesis activity, is generally performed with analytical, simulation or experimental studies. Since, the latter requires tremendous effort in terms of cost, time and technical infrastructure, leaf and vegetation canopy radiative transfer models (RTM) became a valuable method for investigating the relationship between narrow-band spectral features and plant or vegetation canopy parameters (Jacquemoud et al. 2009). This will be able to be applied to a wide range of species and has the advantage that it can be applied from the field level upwards. However, overlapping and confounding internal factors, e.g. heterogeneity of the vegetation canopy parameters, and external factors, e.g. the observation angle, that influence the signal, can constrain the implementation of RTM with studies involving field-scale heterogeneity since RTM assume homogeneity. Empirical analyses have therefore been criticized for their lack of generality, which physical-based approaches promise to overcome. Since, quantitative reflectance data are directly applied as input variables, empirical approaches linking spectral vegetation indices (VI) and vegetation parameters are less influenced by atmospheric correction factors but are rather more site-specific with atmospheric and surface characteristics at the time of data acquisition (Houborg & Anderson 2009). However, to validate quantitative studies using empirical or physical based approaches in general, a tremendous effort is required from ground truthing campaigns. In the case of fluctuating internal factors, e.g. phenological stages, and external factors, e.g. illumination conditions, spatial and spectral observation characteristics, affecting the canopy reflectance between the imaging spectrometer campaigns, the established methods, parameter characteristics, e.g. pigment value range or coefficients are not directly transferable to other applications, study sites or phenological stages. In terms of transferring methods and awareness of important correlations, there is in particular a lack of knowledge about spatial scale-dependent information, which needs to be verified much more with real data and naturally occurring land surface heterogeneity effects.

In the setting of TERENO (terrestial environmental observatories, www.tereno.net, Zacharias et al. 2011) imaging hyperspectral airborne remote sensing plays a key role in long-term monitoring on different scales and in different regions. The Helmholtz Centre for Environmental Research UFZ delivers knowledge about complex systems and relationships in the environment by interlinking the natural, social and human sciences. In order to guarantee a comprehensive process-oriented research in landscapes and ecosystems, imaging spectrometer sensors that are able to conduct in-house optical remote sensing have been made available. The imaging hyperspectral remote sensing is based on the two sensors; AISA-EAGLE (400–970 nm) and AISA-HAWK (970–2500 nm) – Airborne Imaging Spectrometer for Applications (Mäkisara et al. 1993) developed by SPECIM (Spectral Imaging LTD., Finland). The sensors, which have a high geometric (0.5–5 m) and spectral (2.3–8.5 nm) resolution, are used on different observation platforms in order to be able to obtain multi-scale (spatial) spectrometer data. A rotating mirror device enables use of the scan line sensors on a lifting platform and in laboratory experiments. The permanent availability of the sensors guarantees data acquisition at any required time, e.g. a specific phenological stage, on the landscape scale using different aircrafts.

2 Research Issues and Background

Knowledge acquired about the sensitivity of hyperspectral reflectance data and plant or vegetation canopy parameters is generally acquired on a specific spatial observation scale. Many studies used field spectrometer data to analyze appropriate relationships or to develop new spectral vegetation indices (Bannari et al. 2007). Relationships retrieved from radiative transfer models should be interpreted or transferred with caution and the parameterization, e.g. value ranges, observation settings, applied during the experiments should be taken into account (Haboudane et al. 2004). One issue that is often not addressed is the influence of the spatial observation scale on the reflectance signal in terms of its sensitivity to retrieved biochemical and biophysical parameters. Hence, it is generally not clear with which accuracy quantitative results can be expected from potential stakeholders. Therefore the need arises to analyze real data at different spatial observation scales and to test recently proved methods, e.g. radiative transfer models, empirical models, and neural networks, in terms of their performance.

Furthermore there is a lack of knowledge about the temporal behaviour of plant and vegetation canopy reflectance characteristics over entire growing cycles from experimental data on the landscape scale. Such basic monitoring experiments promise progress in understanding the dynamics of overlaying effects in particular, i.e. different plant and vegetation parameters. Simulation experiments are very valuable in providing basic understanding, although a validation of real data cannot be substituted to finally understand the effect of combined effects under real conditions and how these will change over the growing season.

The outlined issues are very difficult to address in research practice and when focusing on the use of commercial hyperspectral data, the following obstacles arise:

- Data is generally collected at one specific spatial resolution due to costs and the time management of the partner operating the sensor.
- Multi-temporal datasets are difficult to obtain over one growing cycle in the case of limited sensor availability, e.g. competing campaigns in terms of time, with appropriate clear sky conditions. For instance, the HyMap (Hyperspectral Mapper) campaigns in Germany over recent years were generally conducted in July and August. A temporal monitoring of different vegetation parameters is therefore not possible over the year or an entire growth cycle.
- Furthermore, as a result of the research question and the key parameters of interest, the requirements of the ground truthing design are comparatively clear. Field campaigns can often take many months or even years to plan because of the high effort of organisation required in terms of staff, e.g. technicians, students, instrumentation, e.g. mobile plant and canopy analyser, laboratory resources, e.g. pigment extraction, and ancillary information, e.g. thematic maps.

To provide realistic and high quality datasets and make progress in remote-sensing analyses, imaging spectrometer data from the landscape level, airborne and ground “segment” are still not a separable union. Therefore, the key issues of the monitoring design presented in this paper are i) to analyze the sensitivity of hyperspectral data to a wide range of physiological parameters and phenological stages on different spatial scales, ii) to acquire knowledge about the temporal dynamics of the above, and iii) to improve our understanding about the overlaying effects for retrieving parameters on “real” data.

To address a wide range of these issues, a scale-specific remote sensing experiment was set up, which is presented in the following section.

3 Scale-specific Monitoring Methods

The concept of the scale-specific remote sensing experiment at the Helmholtz Centre for Environmental Research (UFZ) in Leipzig/Germany is based on data sampling from three different observation scales, which are presented in Fig. 1. Scales 1 to 3 are outlined in more detail in section 3.2. Scale 4, which represents the final landscape scale will not be discussed further in this paper.

All measurements were carried out with the same imaging hyperspectral sensors which are described in section 3.1. Selecting spatial ground resolutions enables algorithms and training datasets to be transferred between scales. During the laboratory measurements we were able to obtain the same footprint (geometric resolution 50 cm × 50 cm) as at the landscape level (scale 3).

3.1 *Sensor Characteristics*

Tab. 1 summarizes the main sensor properties of the UFZ's hyperspectral sensors. The relevant spectral and geometric sensor properties that are required depend on the research objectives and the spatial observation scale.

3.2 *Experimental Design*

Scale 1 – Long term laboratory vegetation monitoring experiments, plot scale (plant level, vegetation canopy)

To examine causal relationships between imaging hyperspectral signals and single plant or vegetation canopy parameters as well as their overlaying effects, the only appropriate method is to conduct controlled laboratory experiments. By conducting such experiments, the spectral response of vegetation under different arbitrary scenarios, drought stress, CO_2, heavy metal pollution, the effect of pesticides etc., can be studied at frequent intervals, e.g. twice a week. Consequently, there is tremendous potential for model development and validation to retrieve plant and vegetation canopy parameters (leaf area index, chlorophyll a/b ratio, photosynthesis activity, biomass, carbon/nitrogen ratio, soil moisture). Furthermore, a major advantage of this kind of laboratory experiments is that measurements are always carried out under the same basic conditions such as light source and general geometric observation properties (2 × 1000 W halogen lamps, fixed angle of incidence and distance sensor-object). A darkroom measuring approximately 3 × 3 × 3 m^3 made of light-

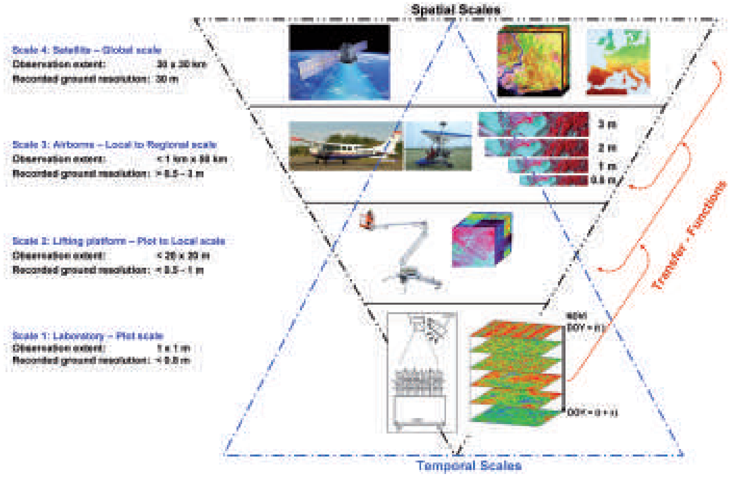


Fig. 1: The concept of monitoring the biophysical–biochemical vegetation variables on different spatial and temporal scales with the imaging hyperspectral sensors AISA-EAGLE/HAWK (modified after Lausch et al. 2012).

Tab. 1: Characteristics of the hyperspectral sensors from the UFZ – AISA-EAGLE / HAWK.

Sensor head	**AISA-EAGLE VNIR**				**AISA-HAWK II SWIR**
Weight	11 kg				18 kg
Dimensions (L/W/H)	380/220/55 mm				220/275/470 mm
Spectral range	400 – 970 nm				970 – 2500 nm
Spectral resolution	2.9 nm				8.5 nm
Max. spatial pixels	1024				320
Camera	CCD Camera				MCT Camera
SNR	350:1 – 1400:1 (depending on band configuration)				800:1 (peak)
Spectral binning options	1x	2x	4x	8x	
Spectral bands	488	252	122	60	254
Spectral sampling/ band	1.25 nm	2.3 nm	4.6 nm	9.2 nm	
Image rate	30	40	60	85	
Focal length	23 mm	18.5 mm		9 mm	22.5 mm
FOV	29.9 degrees	36.7 degrees		62.1 degrees	24.0 degrees
Swath width	0.53 x altitude	0.66 x altitude		1.20 x altitude	0.43 x altitude
Ground resolution at 1000 m altitude	0.52 m	0.65 m		1.2 m	1.34 m
Additional parts					
Mirror scanner	Mirror scanner for local applications (field plots)				Mirror scanner for local applications (field plots)
FODIS	Fiber Optic Down welling Irradiance Sensor				Fiber Optic Down welling Irradiance Sensor

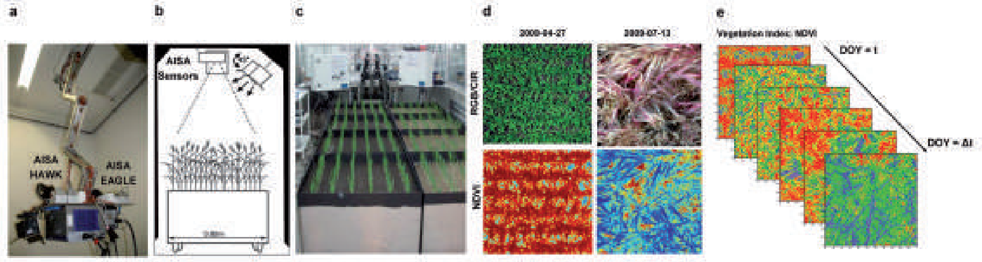


Fig. 2: Use of the imaging hyperspectral sensor AISA-EAGLE/HAWK in the laboratory, a) technical configuration of AISA-EAGLE/HAWK in a lifting platform on the ceiling, b) construction of the laboratory experiment with imaging hyperspectral sensors, c) vegetation scenarios of spring barley experiment 2009 (DOY 117-201), d) RGB, CIR and NDVI (normalized difference vegetation index) derived from the AISA-EAGLE hyperspectral image of spring barley (2009-04-27, 2009-07-13), e) quantification of vegetation indices derived from imaging hyperspectral AISA sensors – example NDVI.

proof material was set up for the hyperspectral measurements. The use of this darkroom prevents any disruptive factors from having an effect over the entire series of tests. Fig. 2 shows the test set-up in the laboratory and examples of imaging hyperspectral data from the available measurement tests on spring barley under different drought stress scenarios.

Scale 2 – Lifting platform, plot to local scale (vegetation canopy)

Test plots with a surface area of $< 20 \times 20$ m^2 can be examined using a lifting platform (Fig. 3). Both hyperspectral sensors (AISA-EAGLE/HAWK) are mounted onto the lifting platform at a height of 2–12 m above the vegetation. The aim of these tests is to record the causal relationships between spectral imaging signals and the target parameters measured, e.g. to derive biophysical and biochemical canopy state variables such as LAI (leaf area index), chlorophyll content, vegetation water content or nutrient status of vegetation, under ‘in-vitro’ conditions. With the lifting platforms a long-term monitoring of different vegetation plots is possible. Furthermore, we can test the influence of different sensor angles, any bidirectional reflectance distribution function (BRDF) effects on the imaging spectral response and establish transfer functions from scale 1 to scale 2.

Scale 3 – Airborne remote sensing, local to regional scale (vegetation, populations, ecosystems)

To provide airborne imaging hyperspectral data for innovative studies related to practice we opted to use a Cessna 207 or Piper for spatially extensive hyperspectral campaigns (Fig. 6) and a microlight aircraft (Trike, D-MUFZ, Fig. 4) as a sensor platform for small scale, e.g. field scale, hyperspectral campaigns. Fig. 4 shows the hyperspectral sensor AISA-EAGLE together with the GPS/INS unit RT3100 (Oxford Technical Solutions LTD., UK) fitted onto the microlight and the microlight itself in operation. The advantages and disadvantages of using a microlight as a sensor platform are listed below.

The advantages of using a microlight are:

- flexibility in terms of time enabling a high repetition rate of data acquisition
- independence from outside bodies in terms of project planning, since pilots and operators are members of the staff of the research institute
- economical use in terms of repairs and maintenance
- its design, enabling a use in a wide range of areas abroad (it can be dismantled and transported in containers together with the sensors)
- recording imaging of hyperspectral data with high spatial (< 0.5 m) and temporal resolution

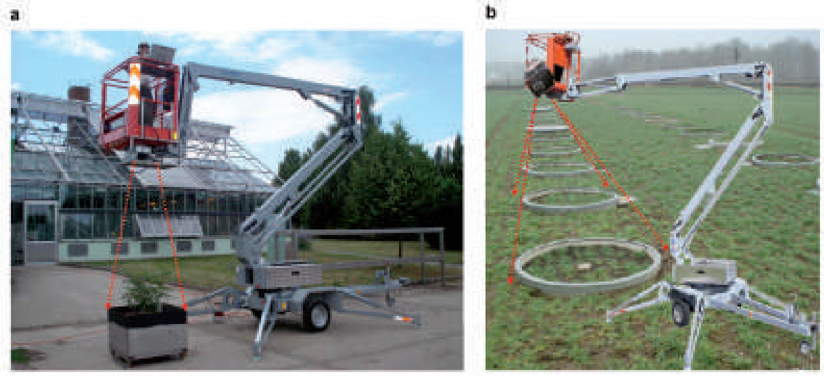


Fig. 3: Using the hyperspectral sensor AISA-EAGLE/HAWK on lifting platforms over a) test plots and b) lysimeters.

Disadvantages of using a microlight are:

- thermal lift affecting the microlight's stability and thus limiting the time of operation during the day
- sometimes difficult handling of the equipment
- limitation of the imaged area (maximum about 10 km²)

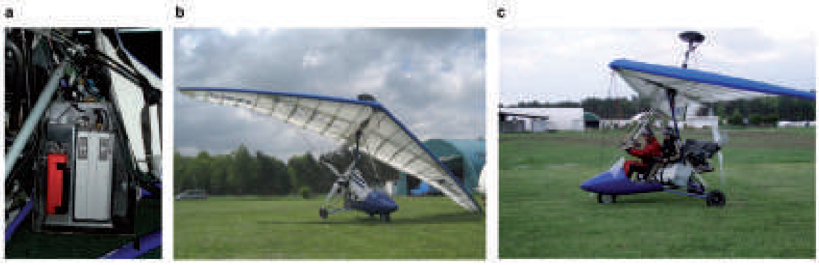


Fig. 4: a) AISA-EAGLE (400–970 nm) and GPS/INS-RT3100 mounted onto the microlight aircraft (Trike, D-MUFZ) of the UFZ, b) Microlight aircraft of the UFZ – D-MUFZ, c) Microlight aircraft for recording hyperspectral data – landscape level.

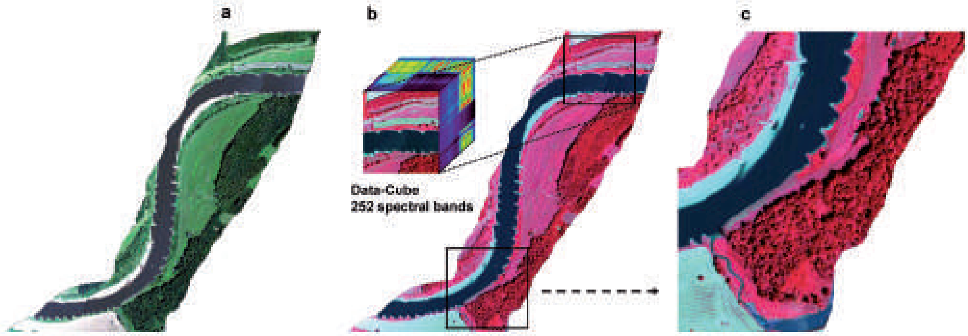


Fig. 5: "Schleusenheger Wiesen" near Dessau recorded 2008-07-03 using the microlight aircraft of the UFZ, images from the hyperspectral sensor AISA-EAGLE, 400–970 nm, 1 m ground resolution, 252 spectral bands: a) RGB-Image, b) CIR-image with data cube, c) CIR hyperspectral image subset.

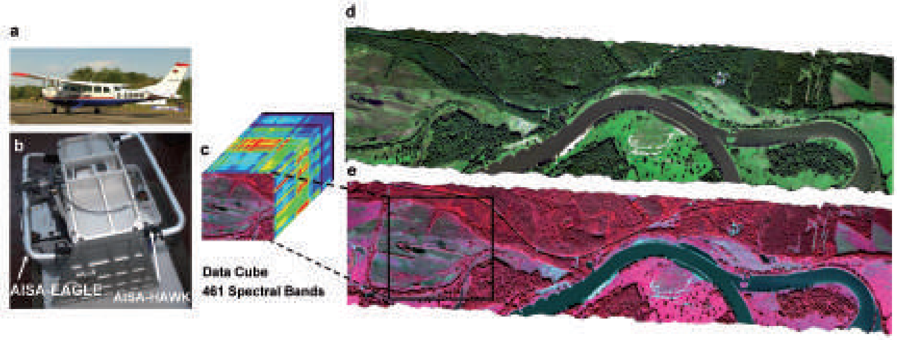


Fig. 6: a) Cessna 207, b) AISA-EAGLE/HAWK Dual sensor mount in the Cessna, c) data cube of AISA-DUAL data, d) and e) "Region Rosslau" – RGB- and CIR-image – taken from the hyperspectral sensor AISA-EAGLE/HAWK, 400–2500 nm, 2 m ground resolution, 461 spectral bands, date of recording 2010-09-23 with a Cessna 207.

4 Preliminary Results

To emphasize the importance of observation characteristics on the final results based on hyperspectral data, the following sections provide some simple examples on this issue.

4.1 *Spectral Response of Imaging vs. non Imaging Spectrometer Data*

Over a three month period from April 27, 2009 to July 20, 2009 (DOY 117-210, 84 days), spectral data (imaging hyperspectal data – AISA-EAGLE and non-imaging ASD Spectrometer) and vegetation parameter measurements (LAI, Chlorophyll SPAD-502, canopy height, vegetation water content, C/N content of vegetation) were recorded twice a week. Details on the experimental design can be found under 3.2 as well as in Lausch et al. (2012). For spring barley various vegetation indices were investigated for the AISA-EAGLE imaging spectral data as well as for the ASD non-imaging hyperspectral data with regard to their suitability for the model in terms of various biochemical and biophysical vegetation parameters over the entire vegetation period of 84 days.

We are able to assume that the differences in the model results for the vegetation index GNDVI (R800-R550) do not result from a change in biochemical or biophysical parameters to the vegetation, soil or atmosphere due to the fact that both measurements taken using the imaging and non-imaging hyperspectral sensors were carried out under the same basic conditions and within minutes of each other. We therefore assume that there are other factors influencing the spectral behaviour of both sensors. The different spectral responses from both sensors could be explained by several factors: (I) Differentiated sensor-specific mapping characteristics and the specific sensor characteristics of AISA-EAGLE (Whiskbroom-Scanner) and ASD. (II) The calibration of the non-imaging spectrometer (ASD) with the imaging spectrometer that sometimes proves to be insufficient or is not carried out at all, leading to inaccurate measurements and consequently a repetition of errors in terms of sensor models and validation with hyperspectral data (at scales 3 and 4). (III) The inner geometry, structure and pattern of the vegetation is strongly reflected by the hyperspectral response. This is much stronger compared to the spectral response of biochemical and biophysical vegetation parameters (chlorophyll-content, vegetation-water content, protein content). (IV) Both sensors take a different "footprint" of the object, e.g. vegetation, even though the field of view (FOV) from the objective is comparable. (V) The varying degree of dependency of the spectral signal on the date (DOY) or change in phenology, i.e. for the vegetation index GNDVI over the entire vegetation period of 84 days (DOY 117-201).

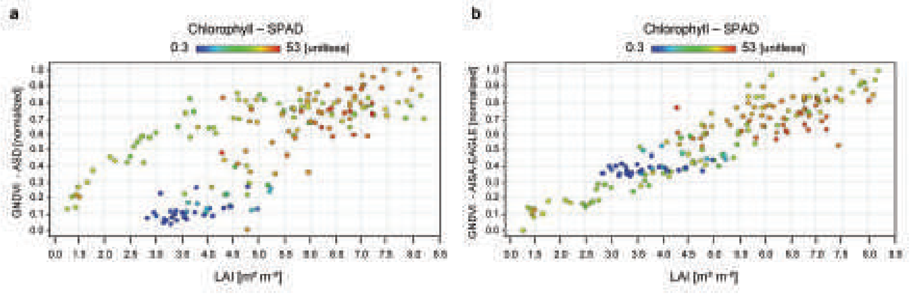


Fig. 7: Relationship between vegetation index GNDVI and LAI (leaf area index) obtained from laboratory measurements with a) non-imaging spectrometer ASD and b) imaging spectrometer AISA-EAGLE for spring barley. The colours show Chlorophyll SPAD-502 content (unitless) over the entire growing season (DOY 117-201).

4.2 *Monitoring of Vegetation Parameters and Spectral Response over an entire flowering Period with the Imaging Hyperspectral Sensor*

Another study describes the same experimental approach and the results from using the imaging hyperspectral sensor AISA (400–970 nm, 252 spectral bands) under controlled and comparable conditions in a laboratory to study the spectral response compared to different biochemical and biophysical vegetation and soil parameters (LAI, Chl-SPAD-502, CCC, GWC, vegetation height, C/N-content) over an entire flowering period of spring barley (Fig. 8 a–f). The spectrum of each hyperspectral image was used to calculate a range of vegetation indices (VI's) that have been recorded in the literature. Furthermore, all combinations of the 252 spectral bands were tested to calculate a range of vegetation difference indices (VI's$_{(xy))}$ and reflectance value indices (R$_{(X)}$) at the central wavelength (*x* nm) of each band (R$_{(x)}$). For all three index types we examined the relationship with the vegetation variables measured on the ground by using a cross-validation procedure.

4.3 *Estimating phenological Stages of Barley from Time Series Measurements with an Imaging Hyperspectral Sensor*

The aim of another application for the approach presented in this paper was to set up a model to predict the different phenological BBCH macro-stages of barley in the laboratory on the plot scale and to transfer the best model found to predict the phenological stages of barley to the landscape scale. To classify phenology eight vitality and phenology-related vegetation parameters were obtained like for example leaf area index (LAI), Chl-SPAD, C-content, N-content, C/N-content, canopy chlorophyll content (CCC), gravimetric water content (GWC) and vegetation height (VH) at the same time that all imaging spectral measurements (AISA-EAGLE) were conducted. These biochemical biophysical vegetation parameters were examined according to their suitability to record images of various phenological BBCH macro-stages of barley (**B**iologische Bundesanstalt für Land- und Forstwirtschaft, **B**undessortenamt und **CH**emische Industrie) (Biologische Bundesanstalt für Land- und Forstwirtschaft 2001, Hack et al. 1992). The predictive models were developed and

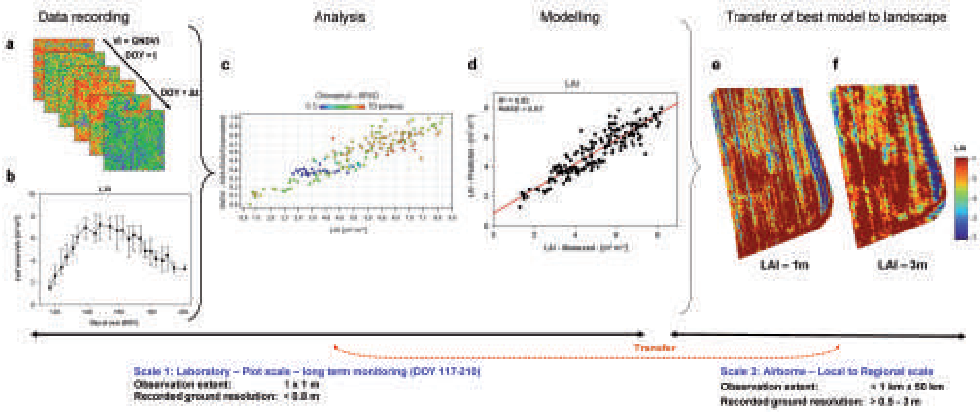


Fig. 8: Long-term vegetation monitoring experiment on scale 1 – laboratory – plot scale of spring barley, DOY 117-210, a) vegetation index GNDVI – obtained from the imaging hyperspectral data – AISA-EAGLE, b) recorded biophysical vegetation parameter – leaf area index (LAI), c) GNDVI and LAI derived from the imaging spectrometer AISA-EAGLE, d) best regression model for estimating LAI quantified from GNDVI – AISA-EAGLE, e) transfer of the best regression model for estimating LAI – GNDVI to scale 3 – airborne – regional scale, modelling LAI in 1 m f) the same for 3 m, recording date 2010-06-15.

tested using four different vegetation index types: (I) Published VIs, (II) Reflectance VIs as well as (III) $VI_{(xy)}$ formula combinations and (IV) a combination of all VIs.

To investigate a differentiation between the phenological BBCH macro-stages of spring barley, many well-known published VIs were included in the analyses. Our results show that the best prediction of the different macro-stages results from a combination of the published VIs PRI (photochemical reflectance index), renormalized difference vegetation index (RDVI) as well as the water band index (WBI) with a classification accuracy of 82.39 %. The best predictive model of the phenological BBCH macro-stages was obtained from a comprehensive model using all three VIs – Published VIs, Reflectance VIs and a combination of formula VIs with a classification accuracy of 84.80 %. The best predictive model was subsequently used on airborne AISA-EAGLE hyperspectral data to model the phenological macro-stages of barley on the landscape level (Fig. 9 e,f).

4.4 *Spatial Heterogeneity Analysis using an Imaging Hyperspectral Sensor*

There are very few empirical studies that use hyperspectral data to support the hypothesis of deriving land surface variables from different spatial scales. The goal of the current study was to investigate the influence of differently recorded spatial scale hyperspectral data on the reflectance behaviour and heterogeneity of the vegetation. The hyperspectral sensors AISA-EAGLE/HAWK were mounted onto an aircraft to record spectral signatures over a very short time sequence of a particular day. The reflectance measurements were collected at three different spatial resolutions ranging from 1 m to 3 m. The NDVI was ascertained from all image data. The NDVI heterogeneity of all images was compared based on methods of variography. Variogram models of the NDVI heterogeneity were obtained from the recorded spatial resolutions 1 m, 2 m and 3 m for grassland and deciduous forest (Fig. 10 a–e).

The results showed that the spatial NDVI patterns of different recorded and scaled data do not correspond among each other. The

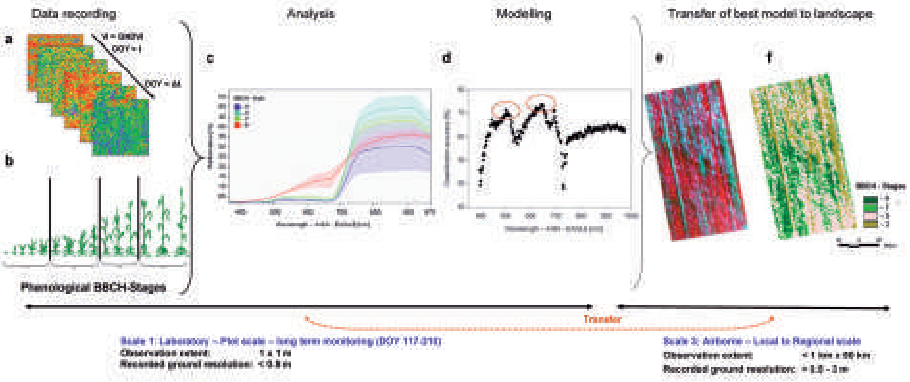


Fig. 9: Long term vegetation monitoring experiment on scale 1 – laboratory – plot scale of spring barley, DOY 117-210, a) quantification of different vegetation indices based on imaging hyperspectral data – AISA-EAGLE, b) BBCH-macro-stages of vegetation, c) spectral response from AISA-EAGLE for the BBCH macro-stages 2, 5, 7 and 9 for barley, d) predictive power of each spectral reflectance value ($R_{(X)}$) at the central wavelength (*x* nm) of each band of the imaging AISA-EAGLE spectrometer to classify the phenological BBCH stages 2, 5, 7 and 9 of spring barley, transfer of the best model for predicting BBCH macro-stages to scale 3 – airborne – regional scale, e) airborne AISA-EAGLE hyperspectral data – 1 m spatial resolution, recording date 2010-06-15, shown as a CIR-image, f) modelling the BBCH macro-stages based on the best model from scale 1 – plot scale.

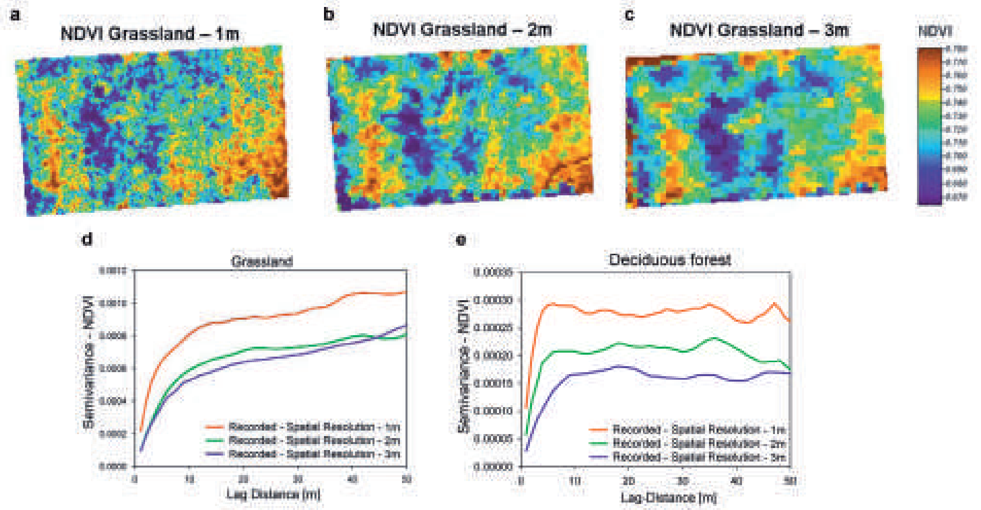


Fig. 10: Calculation of NDVI from the AISA-EAGLE/HAWK (DUAL) hyperspectral data recorded with different ground resolutions, recording date 2010-09-11, a) 1 m, b) 2 m, c) 3 m, Semivariance – NDVI for d) grassland and e) deciduous forest – derived from hyperspectral data (AISA-EAGLE/HAWK) recorded at a spatial resolution of 1 m, 2 m and 3 m.

NDVI patterns of different spectral data that were recorded showed slight changes. The implications behind these findings are that we need to exercise caution when interpreting and combining spatial structures and spectral indices derived from satellite images with differently recorded geometric resolutions.

5 Conclusion and Outlook

The objective of this paper was to propose our scaling method of combining hyperspectral remote sensing data from different spatial and temporal scales and to point out the potential of using only one and the same imaging hyperspectral sensor as the input to plot the local, regional and landscape level.

The use of one hyperspectral sensor at different temporal and spatial scales (“One Sensor At Different Scales” – OSADIS Approach, Lausch et al. 2012) offers the unique advantage of a true comparison of data at different spatial scales as well as the transfer of process information obtained from long-term *in-situ* monitoring investigations. It is possible to investigate the effect of different spatial, temporal, spectral and directional scales of land surfaces i.e. heterogeneity, vegetation phenology or wavelength. Confounding factors such as the phenology of vegetation, BRDF measurements from vegetation geometry, and a number of dynamic atmospheric effects etc. can specifically be eliminated, as such parameters can be considered to be relatively constant within the time frame of 2–3 hrs.

With our preliminary results we were able to show that it is not only possible to combine sensors with different characteristics, e.g. AISA-EAGLE and ASD, geometrically, spectrally as well as temporally but also to apply these process investigations over different scales.

A comparability of measuring equipment with different imaging optics and sensors is extremely difficult. Spectral sensors have different (I) recording characteristics, (II) spectra and spatio-temporal recording characteristics as well as (III) FOV. Moreover, an attempt to introduce some conformity results in BRDF as well as species-specific spectral responses, which result from a difference in sensors and not from a difference in processes. It is therefore extremely difficult to separate these effects.

By using the One Sensor At Different Scales Approach we are able: a) to develop suitable stress-controlled long-term vegetation indicators for selected target variables like for example LAI, chlorophyll, photosynthesis activity, water content, nutrient content etc. b) to realistically transfer the models obtained to the landscape level c) to record imaging hyperspectral information at different spatial scales, whereby we are able to achieve a true comparison of the structural and process results obtained d) to minimize the existing magnitude of errors from geometrical, spectral and temporal effects due to sensor- and temporal-specific differences e) to carry out a top-up and down-scaling through the determination of scale-dependent correction factors and transfer functions. f) With our scale approach (OSADIS) we attempt to understand scales, structures, patterns and their temporal changes better and in more depth and are able to describe or quantify them at all.

References

Bannari, A., Khurshid, K.S., Staenz, K. & Schwarz, J.W., 2007: A comparison of hyperspectral chlorophyll indices for wheat crop chlorophyll content estimation using laboratory reflectance measurements. – IEEE Transactions On Geoscience And Remote Sensing **45:** 3063–3074.

Biologische Bundesanstalt für Land- und Forstwirtschaft, 2001: Entwicklungsstadien mono- und dikotyler Pflanzen. – BBCH Monographie, http://www.jki.bund.de/index.php?id=1075&q=bbch (05.06.2012).

Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J. & Strachan, I.B., 2004: Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. – Remote Sensing of Environment **90:** 337–352.

Hack, H., Bleiholder, H., Buhr, L., Meier, U., Schnock-Fricke, U., Weber, E. & Witzenberger, A., 1992: Einheitliche Codierung der phänologischen Entwicklungsstadien mono- und dikotyler Pflanzen. – Erweiterte BBCH-Skala, Allgemein. – Nachrichtenblatt Deutscher Pflanzenschutzdienst **44:** 265–270.

Houborg, R. & Anderson, M.C., 2009: Utility of an image-based canopy reflectance modeling tool for remote estimation of LAI and leaf chlorophyll content at regional scales. – Journal of Applied Remote Sensing **3:** 259–274.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., Francois, C. & Ustin, S.L., 2009: PROSPECT plus SAIL models: A review of use for vegetation characterization. – Remote Sensing of Environment **113:** 56–66.

Lausch, A., Pause, M., Merbach, I., Zacharias, S., Doktor, D., Volk, M. & Seppelt, R., 2012: A new multi-scale approach for monitoring vegetation using remote sensing-based indicators in laboratory, field and landscape. – Environmental Monitoring and Assessment, doi: 10.1007/s10661-012-2627-8).

Mäkisara, K., Meinander, M., Rantasuo, M., Okkonen, J., Aikio, M., Sipola, K., Pylkkö, P. & Braam, B, 1993: Airborne Imaging Spectrometer for Applications (AISA). – Digest of IGARSS'**93** (2): 479–481, Tokyo, Japan.

Schaepman, M.E., Ustin, S.L., Plaza, A.J., Painter, T.H., Verrelst, J. & Liang, S.L., 2009: Earth system science related imaging spectroscopy. – An assessment. – Remote Sensing of Environment **113:** 123–137.

Stuffler, T., Forster, K., Hofer, S., Leipold, M., Sang, B., Kaufmann, H., Penne, B., Mueller, A. & Chlebek, C., 2009: Hyperspectral imaging. – An advanced instrument concept for the EnMAP mission (Environmental Mapping and Analysis Programme). – Acta Astronautica **65:** 1107–1112.

Zacharias, S., Bogena, H., Samaniego, L., Mauder, M., Fuss, R., Pütz, T., Frenzel, M., Schwank, M, Baessler, C., Butterbach-Bahl, K., Bens, O., Borg, E., Brauer, E., Dietrich, P., Hajnsek, I., Helle, G., Kiese, R., Kunstmann, H., Klotz, S., Munch, J.C, Papen, H., Priesack, E., Schmid, H.P., Steinbrecher, R., Rosenbaum, U., Teutsch, G. & Vereecken, H., 2011: A network of terrestrial environmental observatories in Germany. – Vadose Zone Journal **10:** 955–973.

Addresses of the Authors:

Dr. Angela Lausch, Helmholtz Centre for Environmental Research – UFZ, Department Computational Landscape Ecology, Permoserstraße 15, D-04318 Leipzig, Tel.: +49-341-235-1961, e-mail: angela.lausch@ufz.de

Dr. Marion Pause, University of Tübingen, Water & Earth System Science Competence Centre, Keplerstrasse 17, D-72074 Tübingen, Tel.: +49-7071-2975240, e-mail: marion.pause@uni-tuebingen.de

Dr. Ines Merbach , Helmholtz Centre for Environmental Research – UFZ, Department of Community Ecology, Theodor-Lieser-Str. 4, D-06120 Halle, Tel.: +49-345 558 5393, e-mail: ines.merbach@ufz.de

Sarah Gwillym-Margianto, Helmholtz Centre for Environmental Research – UFZ, Department of Community Ecology, Theodor-Lieser-Str. 4, D-06120 Halle, Tel.: +49-179-4784912, e-mail: sarah.gwillym@yahoo.co.uk

Prof. Dr. Karsten Schulz, Department of Geography, Ludwig-Maximilians-Universität München, Luisenstr. 37, D-80333 München, Tel.: +49-89-2180-6681, e-mail: k.schulz@lmu.de

Dr. Steffen Zacharias, Helmholtz Centre for Environmental Research – UFZ, Department Department Monitoring and Exploration Technologies, Permoserstraße 15, D-04318 Leipzig, Tel.: +49-341-235-1381, e-mail: steffen.zacharias@ufz.de

Prof. Dr. Ralf Seppelt, Helmholtz Centre for Environmental Research – UFZ, Department Computational Landscape Ecology, Permoserstraße 15, D-04318 Leipzig, Tel.: +49-341-235-1250, e-mail: ralf.seppelt@ufz.de

Manuskript eingereicht: November 2011
Angenommen: Mai 2012

PFG 2012 / 5, 0603–0612
Stuttgart, October 2012

Article

Comparing Classification Results of Multi-Seasonal TM against AVIRIS Imagery – Seasonality more Important than Number of Bands

Sylvio Mannel, Nevada, MO & Maribeth Price, Rapid City, SD, USA

Keywords: AVIRIS, Landsat TM, multitemporal, forest, classification

Summary: We classified forest cover and tree density in the Black Hills, SD, in twenty spatially contiguous AVIRIS scenes. Results were compared to those derived from two-season Landsat TM imagery. A decision tree classifier was used to analyze the TM data as well as the over two hundred bands of the twenty AVIRIS scenes. The classification based on summer AVIRIS data was more accurate than the classification based on the comparable early fall TM data. However, classification of spring and especially, two-season TM data resulted in higher accuracies than the classification based on summer hyperspectral data. These results indicate that seasonality is more important than the number of spectral bands.

Zusammenfassung: Die Waldfläche und Baumdichte in den Black Hills, South Dakota wurde in zwanzig räumlich zusammenhängenden AVIRIS Szenen klassifiziert. Diese Resultate wurden mit jahreszeitlich verschiedenen Landsat TM Bildern verglichen. Die TM Daten und die über zweihundert Bänder der zwanzig AVIRIS Szenen wurden anhand einer Entscheidungsbaum-Klassifizierung (decision-tree) analysiert. Aus den Ergebnissen lässt sich zeigen, dass die im Sommer aufgenommenen AVIRIS Klassen eine höhere Genauigkeit also die Frühherbst TM Daten aufweisen. Allerdings sind die Ergebnisse für TM besser, wenn Frühlingsdaten herangezogen werden. Die TM Kombination von Frühling und Herbst hat insgesamt die höchste Genauigkeit. Daraus lässt sich ableiten, dass Jahreszeit wichtiger als die Anzahl der Spektralbänder ist.

1 Introduction

Hyperspectral data offer the opportunity to explore the differences of land cover types without being restricted to a few wavelengths. AVIRIS (airborne visible/infrared imaging spectrometer) was the first hyperspectral sensor that measured over 200 bands between the 400 nm and 2500 nm spectrum with individual band widths of ~10 nm. AVIRIS employs four spectrometers in the following ranges: 400–710 nm, 670–1290 nm, 1250–1870 nm and 1830–2450 nm (Green et al. 1990). The flight altitude is about 20 km with a rate of 7300 spectra per second (Green et al. 1990). The covered area of a high altitude flight ranges from 11 km × 9 km (Williams & Hunt Jr. 2002) to 12.3 km × 10.2 km (Riano et al. 2002). The resulting area of about 120 km^2 is much smaller than for example a Landsat TM scene, which usually covers about 20,000 km^2.

Vegetation studies can take advantage of the continuously available reflectance; for example, Niemann et al. (2002) note that narrow spectral bands are necessary to detect some forest related parameters whose spectral range may be small. Hyperspectral data have been used to map land cover types such as woody vegetation (Wylie et al. 2000, Ustin & Xiao 2001), leafy spurge (Williams & Hunt Jr. 2002), shrub recovery after fire (Riano et al. 2002), vegetation in semi arid ecosystems (Asner & Heidebrecht 2002, Okin et al. 2001) or lake water quality (Hoogenboom et al. 1998, Thiemann & Kaufmann 2002).

The nearly continuous spectrum also has its costs, in an economical, computational, and spatial sense. A typical AVIRIS scene holds

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0142 1432-8364/12/0142 $ 2.50

about 400 MB. Most classifications, for example, maximum likelihood, require all flight lines fused into one image, severely restricting their utility for hyperspectral classification of large areas with more than 10 AVIRIS scenes. However, a decision tree classification, allows the user to work with each AVIRIS scene separately. To our knowledge this might be one of the first studies to classify a medium scale forest of about 2,000 km^2, utilizing twenty AVIRIS contiguous scenes. In preliminary tests decision trees also showed to be comparable or even slightly better than maximum likelihood classifications of Landsat TM data (Mannel et al. 2002).

Decision trees use a binary recursive partitioning algorithm to divide the data into smaller subsets with increasing statistical homogeneity (Swain & Hauska 1977, Breiman et al. 1984, Clark & Pregibon 1993). These divisions can be represented as branches and nodes, where nodes are connected to a set of possible answers that will lead to a classification. This process is often referred to as data mining (Read 2000).

Projects that compared AVIRIS data to other data sources are not always in agreement. Ustin & Xiao (2001) found AVIRIS about 20 % more accurate than SPOT data in classifying forest regions. Lee & Cohen (2002) had better success mapping leaf area index (LAI) with AVIRIS than with Landsat ETM+ data. Lefsky et al. (2001), on the other hand, found that multi-seasonal TM data performed better than AVIRIS for quantifying forest biomass and basal area, although their study used 1994 AVIRIS data in which the signal-to-noise ratio was lower than today. They used six temporally different TM scenes but also showed improved results with only two TM scenes. Different seasons have been found to be important for vegetation classifications (Schriever & Congalton 1995, Wolter et al. 1995, Lefsky et al. 2001).

In this study we compare twenty scenes of single-date AVIRIS data and two-season TM imagery to classify a forest in the Black Hills, South Dakota, USA. These two sensors were chosen to allow us to compare expensive hyperspectral data to a low-cost multispectral alternative. We received AVIRIS data via a grant by NASA and then selected freely available Landsat TM imagery over the same area. A decision trees classifier was used because of slightly better results in preliminary tests (Mannel et al. 2002) and because it offers the best alternative for processing large areas composed of many AVIRIS scenes (by being able to work with separate scenes without the need to fuse them).

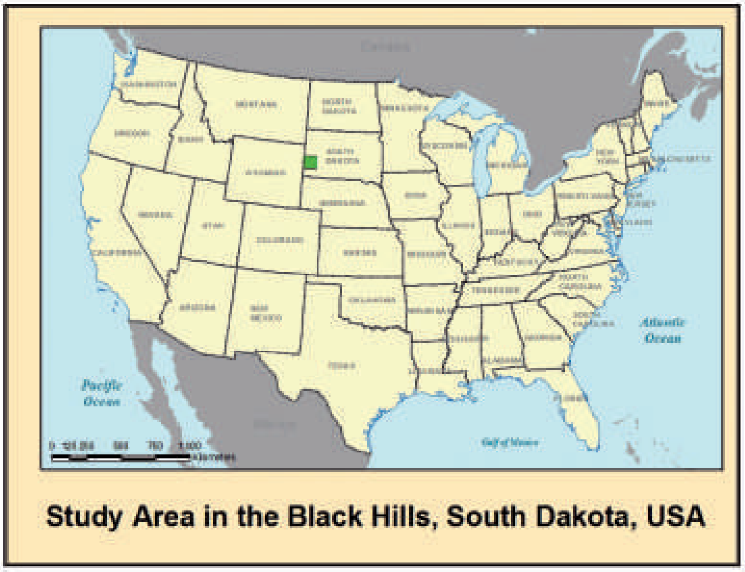


Fig. 1: Study area.

2 Methods

2.1 *Study Area*

The study site lies in the northern Black Hills, SD. The Black Hills form an oval uplift approximately 200 km × 100 km in size with elevations reaching 2200 m in contrast to the surrounding plains at an elevation of about 900 m. The forested hills rise above the sea of prairie grass and form an orogenic and ecologic island in the plains. Ponderosa pine (*Pinus ponderosa*) covers approximately 84 % of the Black Hills (Bennett 1984), supplemented by other trees such as white spruce (*Picea glauca*), aspen (*Populus tremuloides*), paper birch (*Betula papyrifera*) and oak (*Quercus macrocarpa*). Ponderosa pine and white spruce each form forest stands typically dominated by one species, and they form two of the primary mapping classes in this study. Birch is hard to distinguish from aspen even from a few yards away. The spectral similarities between aspen and birch were confirmed by Singhroy et al. (2000) who measured both species with a field spectrometer. In addition the two trees usually occur as a species association in medium to dense stands. For these reasons, birch and aspen were mapped as a single class. Bur oak stands are isolated and too small to justify a separate land cover based on remote sensing. Open meadows are common in the Black Hills, and are dominated by mixed shortgrass prairie grasses such as western wheatgrass (*pascopyrum smithii*), blue grama (*bouteloua gracilis*), little bluestem (*schizachyrium scoparium*), and others, with a variety of forbs (Larson & Johnson 1999). Non-vegetated classes include water and bare rock or soil.

2.2 *Field Data and Cover Classes*

Field data plots were sampled in the summer of 2000. Points were randomly distributed and stratified by cover class and forest density based on the Rocky Mountain Resource Information System (RIS) compiled by the Black Hills National Forest. Density refers to the percentage of ground covered by tree crowns when viewed from directly overhead. We established and measured 135 15-m radius plots. Overstory canopy cover was measured with a sighting tube (Ganey & Block 1994, Cook et al. 1995, Mannel et al. 2006). The sighting tube made it possible to record species while taking canopy measurements. Based on the forest service protocols, we distinguished the following densities: open, medium (40 %–60 % canopy cover), and dense (RIS, unpublished data 2000). Cover type was designated based on the dominant species that covered more than 70 % of the plot. If no species covered more than 70 %, the plot was labeled “mixed”.

Tree species differ in their abundance and densities. Ponderosa pine is abundant in the Black Hills but rarely forms very dense stands. We only encountered two field plots with a canopy cover of more than 70 %. The spectral reflectance characteristics are therefore close to the medium class (40 %–60 %). We, therefore, combined medium and dense pine to one class. Aspen and white spruce are much less abundant than pine. Both tree classes usually grow in medium or dense stands in the Black Hills. We combined all densities for these two species into one aspen and one white spruce class.

Trees in the Black Hills typically form distinct stands dominated by one species, and areas with mixtures of different species (other than aspen/birch) were rare and of small size. As a result, an insufficient number of reference data for mixed classes could be collected. Water was identified using Digital Orthophoto Quadrangles (DOQs). DOQs are essentially georeferenced aerial photographs. Our DOQs were black and white with a ground resolution of one metre.

Training data for the nonvegetated class was based on an open sand pit, which was also the calibration site for AVIRIS atmospheric correction. Additional reference data for bare areas were difficult to collect, since bare areas large enough to buffer possible georeferencing errors and mixed pixels are rare in the Black Hills. Our final land cover scheme included water, nonvegetated bare areas, meadow, aspen (including birch), spruce, open pine, and medium/dense pine.

The field-sampled plots were mapped on AVIRIS, TM, and DOQ imagery and were visually inspected for potential mislocations prior to including them in the reference data set. Checking the spatial validity of remotely sensed data with respect to reference data location is necessary in boundary regions between classes. Removing reference data from boundary regions, such as roads or along meadow/forest boundaries, can increase the accuracy significantly (Foody 2002, Mannel et al. 2011).

In order to provide sufficient training and test data we used DOQs to collect additional points spatially close to field data (Mannel et al. 2006). Around all field-sampled plots we visually identified homogeneous areas using 1-m resolution DOQs, making the assumption that the vegetation characteristics within those regions were similar to those measured at the field site (Mannel et al. 2006, Mannel et al. 2011). Tab. 1 shows the final reference data distribution based on field data and data clusters sampled with DOQs. There are fewer clusters than field data for two reasons. First, some field data sites were spatially invalid and were omitted from consideration. Second, some of the original randomly sampled sites were close to each other, and were considered to belong to a single cluster.

It is obvious that the DOQ-collected reference data would be spatially autocorrelated within each cluster. Moran's index of all the data including the clustered additional points was 0.67 (Mannel et al. 2011). However, autocorrelation between clusters was not present because the field-sampled plots at the core of the clusters were randomly stratified, and spatially close sites were aggregated to a single cluster. When assigning data points for either training or testing during analysis, the reference data were always divided by clusters such that a single cluster was assigned as either all training data or all testing data, but never both (Mannel et al. 2011). This technique was applied to negate the effects of spatial autocorrelation on the accuracy assessment.

2.3 *Image Preprocessing*

The high-altitude AVIRIS flight took place in the summer of 2000. The first flight on 26 June 2000 was repeated due to 5 %–50 % cloud cover within the flight rows. The second flight on 6 July 2000 covered six flight lines resulting in 30 scenes. The weather was clear with the exception of light haze and one cloud. We selected twenty scenes that covered an area containing most of our field data sites.

During the AVIRIS overflight we measured the reflectance of our calibration site, a bare sand pit, with a handheld spectrometer (ASD VNIR Dual, with a spectral range of ~370 nm–1100 nm). The atmospheric correction program ACORN (Analytical Imaging and Geophysics, LLC, 2001) was used in conjunction with the calibration site measurements to correct for atmospheric effects and to convert radiance to reflectance. The following AVIRIS bands were removed: 374 nm–394 nm (bands 1–3), 1354 nm–1424 nm (bands 107–114), 1812 nm–1951 nm (bands 153–168), and 2489 nm–2509 nm (bands 222–224) because of strong atmospheric effects. In these wavelengths atmospheric gases, such as water vapor increase the signal noise to unaccepta-

Tab. 1: Number of sites for each land cover class. The field measured plots were supplemented with additional points based on DOQs resulting in large point clusters.

Land cover	Number of field data plots	Number of field- and DOQ-based reference data	Number of clusters
Meadow	19	64	10
Bare	0	5	1
Open pine	26	104	14
Dense pine	26	212	19
Spruce	21	126	12
Aspen	17	84	9
Water	0	106	2
Mixed	25	Not utilized	–

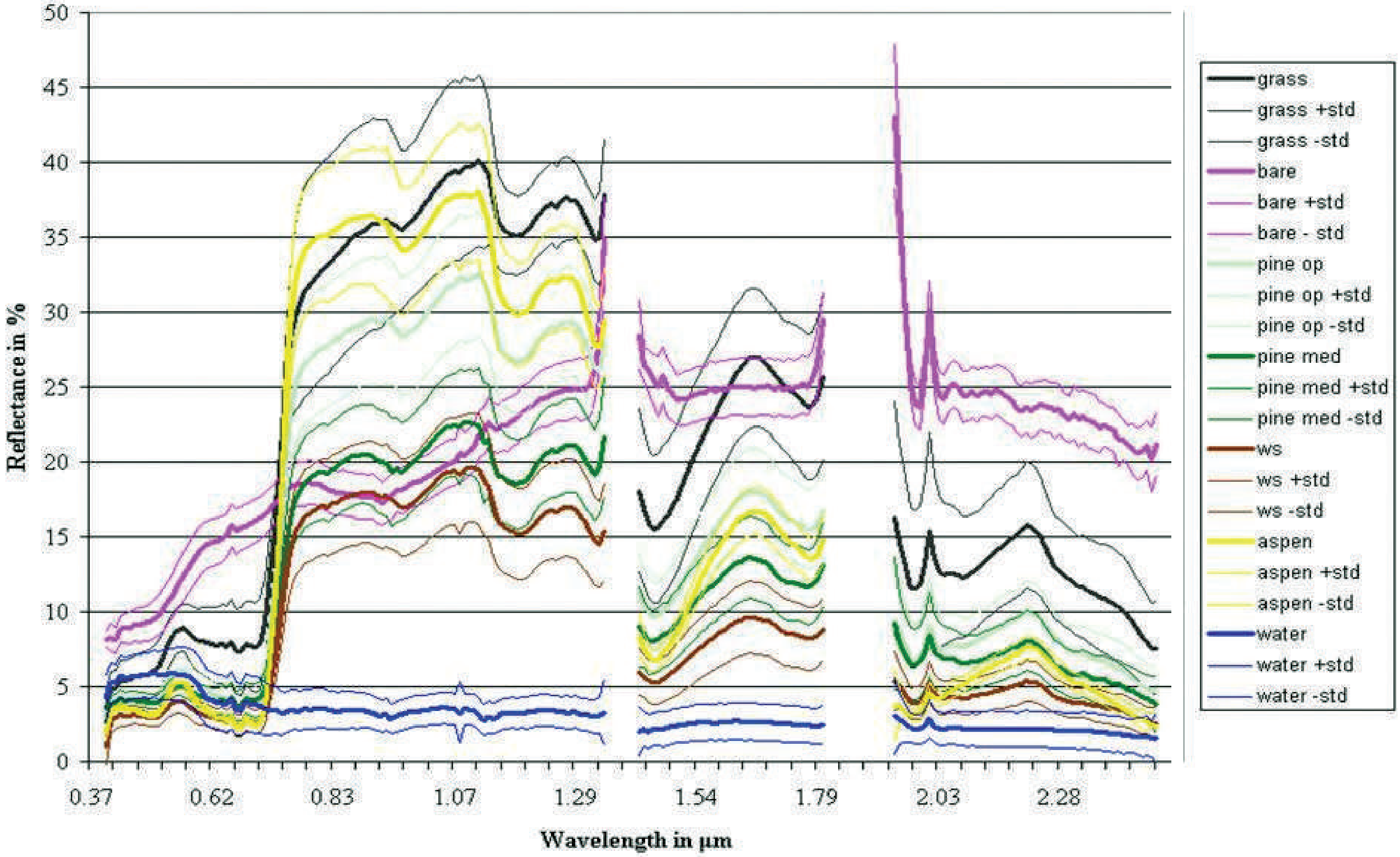


Fig. 2: AVIRIS spectra and standard deviation for land cover classes. The following AVIRIS bands were removed: 374 nm–394 nm (bands 1–3), 1354 nm–1424 nm (bands 107–114), 1812 nm–1951 nm (bands 153–168), and 2489 nm–2509 nm (bands 222–224) because of strong atmospheric effects.

ble levels. Some of the noise is still visible as spikes in the remaining bands (Fig. 2). After adjusting for atmospheric noise 194 bands (out of the 224) were still available for the actual classifications, somewhat lowering the available spectrum for classifications, but still significantly higher than the number of Landsat TM bands.

AVIRIS scenes were georeferenced using DOQs. The root-mean-square (RMS) error of the transformation was 1.3 pixels, and the maximum RMS error for a single GCP was 4.6 pixels. We then overlaid our reference data and identified the matching AVIRIS pixels. Fig. 2 shows the spectral distribution of our reference data along the AVIRIS bands. Vegetation like grass, deciduous trees, and coniferous trees show a distinctive reflectance spectrum (Fig. 2). However, classification challenges become evident through the high standard deviation throughout the different vegetation types (Fig. 2). The high standard deviation for water is due to two very different lakes, one was clear, while the other one contained high levels of sediments.

The Landsat TM5 images were from early spring; leaves not yet fully developed (May 5, 1998) and early fall, "leaves on" (September 24, 1998). We cut the TM scene to match the approximate area of the AVIRIS overflight. Both seasons of the TM data were classified separately, as well as, a combined two-season scene to investigate the influence of seasonality. The same reference data were applied to AVIRIS and TM. Remotely sensed data and field data were collected within 4 years. We did not notice any significant change in landcover, e.g. via logging, during this time period.

2.4 *Decision Tree Classification*

We classified all data using the decision tree program See5 (also known as C5.0), distributed by RuleQuest Research Pty Ltd (Quinlan 2002). See5 is largely based on the technology used by its predecessor C4.5, whose algorithms are further explained in Quinlan (1996). According to Kotsiantis (2007) C4.5 is "the most well-known algorithm in the literature for building decision trees". Further algorithms are regarded as proprietary (Quinlan 2002, personal communication). See5 assigns the classes by weighting the quantity of the input data, i.e. more training data of pine means there is more pine in the study area.

See5 allows for the building of multiple decision trees to improve accuracy by utilizing several classifiers that predict a class. These predictions are counted and determine the final class. One type of multiple trees is "boosting", which identifies the difficulties and mistakes made by the previous iteration and concentrates on them in the next iteration (Friedl et al. 1999, Chan et al. 2001, Wu et al. 2006). Boosting lasts until the predefined number of iterations is reached. We utilized boosting, because our data (Mannel 2003) and other studies showed an accuracy improvement by about 5 % (Friedl et al. 1999, Chan et al. 2001).

One point of caution in using decision trees is "overfitting" the decision tree to the reference data visible in large "trees" (Kotsiantis 2007). A smaller "pruned" tree, with fewer branches is usually more robust than a larger tree (Quinlan 1996, Friedl et al. 1999). Therefore, we used the smallest tree with the highest accuracy for our final results.

2.5 *Accuracy Assessment*

We divided the reference data into different test and training data based on spatially unrelated clusters, rather than on individual points. An entire cluster was either training or test data (Mannel et al. 2006). In addition, we sought to reduce bias in our accuracy assessment by performing a 4-fold cross-validation (Mannel et al. 2011). For that we manually created four trial sets each with different test and training data that were chosen by randomly selecting 1/3 of the reference clusters as test data and the remaining 2/3 clusters as training data. Producer and user accuracies were averaged to calculate overall accuracy.

3 Results

The two pine densities were hardest to distinguish. Confusion mainly existed between the pine densities and between medium pine and white spruce (Fig. 2). On the other hand bare soil and water were easiest to distinguish, because of their unique spectra and homogeneity of the reference pixels. This was followed by grass with almost 100 % accuracy when a spring image was included (Tab. 2). Misclassifications of meadow occurred with aspen/birch especially in the summer (AVIRIS) and early fall (fall TM).

Classification success depended on 1) the type of landcover, 2) sensor type and 3) season:

1) Landcover characteristics affecting classification success are: a) spectral properties and b) the homogeneity of the respective 20 m (AVIRIS) or 30 m (TM) mixed pixels. On the one hand, birch and aspen are virtually indistinguishable, because they have very similar spectral properties (Singhroy et al. 2000) and occur in mixed stands. Distinguishing different densities of the same forest composition is also tricky since the border between open and dense is fluent and artificial. On the other hand, we expected and achieved 100 % accuracies for homogeneous and spectrally unique bodies, such as water and bare soil. For meadow we also expected and partly achieved very high accuracies because of their relatively homogeneity (in comparison to forested areas) and

Tab. 2: Comparison of vegetation classification success of spring TM, fall TM, two-seasonal TM and summer AVIRIS based on the 4fold holdout method (Mannel et al. 2011).

Land cover	Spring TM (%)	Fall TM (%)	Two-season TM (%)	Summer AVIRIS (%)
Meadow	99	86	99	88
Open pine	68	57	79	71
Dense pine	87	83	88	79
Spruce	86	83	89	81
Aspen	86	68	90	83
Average	85	75	89	81

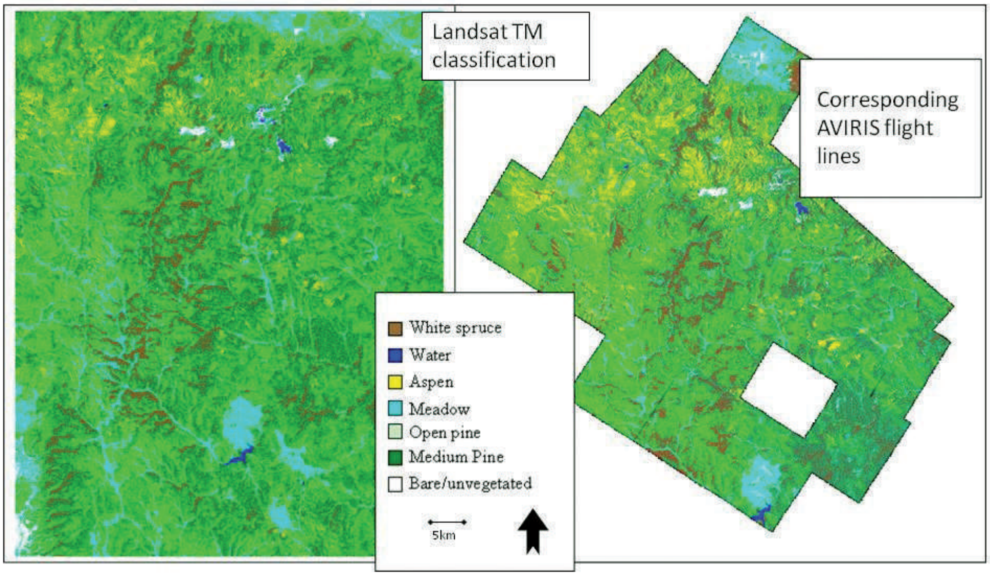


Fig. 3: a) Classification of combined spring-fall Landsat TM data (left) and b) AVIRIS classification (right). The missing scene contained errors that the data provider was unable to fix. No reference data was used from the area of the missing AVIRIS scene.

their spectral properties. The higher wavelengths (>1.5 nm) provided grounds for distinguishing meadow from aspen/birch (Fig. 2).

2) Summer AVIRIS data led to 6 % better results than the comparable early-fall TM image.
3) Seasons clearly influence success of tree classifications. Classifying just the TM spring image showed an accuracy increase of about 10 % in comparison to the fall image (Tab. 2). Utilizing both spring and fall increased the TM classification accuracy by 4 % compared to just the spring season (Tab. 2). We found the two-seasonal TM classification to be about 8 % more accurate than the classification based on summer AVIRIS data (Tab. 2). All vegetation covers were better when using the two-seasonal TM approach, including meadow, aspen, spruce and the two pine densities (Tab. 2).

Fig. 3 shows the classified area based on two-seasonal TM classification and the corresponding AVIRIS flight lines.

4 Discussion and Conclusion

We found that the decision tree classification of summer-AVIRIS data provided 6 % better results than the comparable early-fall TM image. However, the summer AVIRIS classification fell short of accuracies reached via the spring TM scene. This result confirms the importance of seasonality for vegetation analysis. The AVIRIS data were taken in the middle of the summer when leaves were fully developed, which may not be the best season for a forest classification (Schriever & Congalton 1995). Accuracy can be further improved by combining several seasons. The two-seasonal TM data showed the best classification accuracy in all cover classes. Furthermore, multi-seasonal multispectral data is less demanding in terms of processing time and operator effort than hyperspectral AVIRIS data.

The advantages of hyperspectral remote sensing cannot be doubted, but for large-area land cover mapping applications, inexpensive multispectral data would appear to remain the sensor of choice. Multispectral data covers a larger area in a single scene, is less expensive to obtain, and requires less effort to process and analyze. Our study seems to indicate that seasonality is crucial and exceeds advantages of hyperspectral data.

Acknowledgements

The authors would like to acknowledge the efforts of Dong Hua for his collaborative input on this research, Mark Rumble and the USDA Rocky Mountain Forest Service Research Station for providing vegetation data and support with the reference data collection, Bruce Wylie and Chengquan Huang for assistance with the decision tree analysis, and Doug Baldwin and Krystal Price for collecting and processing field data. Furthermore, we would like to thank Teri Peterson for her help with the statistical analysis and Paul Cook for his review and suggestions. This research was supported by National Aeronautics and Space Administration (NASA) Food and Fiber Applications of Remote Sensing Program, Grant NAG13-99021, South Dakota School of Mines and Technology, National Science Foundation/Experimental Program to Stimulate Competitive Research (EPSCoR) Grant EPS-0091948, EPS-9720642 and the South Dakota Space Grant Consortium.

References

Analytical Imaging and Geophysics, LLC, 2001: ACORN (TM) Tutorial – Stand Alone Version. – ImSpec, LLC, Boulder, CO, USA.

Asner, G. & Heidebrecht, K., 2002: Spectral unmixing of vegetation, soil and dry carbon cover in arid regions: comparing multispectral and hyperspectral observations. – International Journal of Remote Sensing **23** (19): 3939–3958.

Bennett, D., 1984: Grazing Potential of Major Soils within the Black Hills of South Dakota. – Master of Science Thesis, South Dakota State University, Brookings, SD, USA.

Breiman, L., Friedman, J., Olshen, R. & Stone, C., 1984: Classification and Regression Trees, Wadsworth, Belmont, CA, USA.

Chan, J., Huang, C. & DeFries, R., 2001: Enhanced Algorithm Performance for Land Cover Classification from Remotely Sensed Data Using Bagging and Boosting. – IEEE Transactions on Geoscience and Remote Sensing **39** (3): 693–695.

Clark, L.A. & Pregibon, D., 1993: Tree-Based Models. – Chambers, J.M. & Hastie, T.J. (eds.): Statistical Models in S. – Chapman & Hall Inc., London, Great Britain.

Cook, J., Stutzman, T., Bowers, C., Brenner, K. & Irwin, L., 1995: Spherical Densiometers Produce Biased Estimates of Forest Canopy Cover. – Wildlife Society Bulletin **23** (4): 711–717.

Foody, G., 2002: Status of land cover classification accuracy assessment. – Remote Sensing of Environment **80:** 185–201.

Friedl, M., Brodley, C. & Strahler, A., 1999: Maximizing Land Cover Classification Accuracies Produced by Decision Trees at Continental to Global Scales. – IEEE Transactions on Geoscience and Remote Sensing **37:** 969–977.

Ganey, J. & Block, W., 1994: A Comparison of Two Techniques for Measuring Canopy Cover. – Western Journal for Applied Forestry **9** (1): 21–23.

Green, R., Conel, J., Carrere, V., Bruegge, C., Margolis, J., Rast, M. & Hoover, G., 1990: Inflight Validation and Calibration of the Spectral and Radiometric Characteristics of the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). – SPIE Conference on Aerospace Sensing, Imaging Spectroscopy of the Terrestrial Environment, Orlando, Florida, USA.

Hoogenboom, H., Dekker, A. & Althuis, I., 1998: Simulation of AVIRIS Sensitivity for Detecting Chlorophyll over Coastal and Inland Waters. – Remote Sensing of Environment **65:** 333–340.

Kotsiantis, S., 2007: Supervised Machine Learning: A Review of Classification Techniques. – Informatica **31:** 249–268.

Larson, G. & Johnson, J., 1999: Plants of the Black Hills and Bear Lodge Mountains, South Dakota State University, Brookings, SD, USA.

Lee, K. & Cohen, W., 2002: Comparison of AVIRIS and Landsat ETM+ for the estimation of leaf area index. – **11th** JPL Airborne Earth Science Workshop.

Lefsky, M., Cohen, W. & Spies, T., 2001: An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. – Canadian Journal of Forest Research **31:** 78–87.

Mannel, S., Hua, D., Price, M., Knobloch, A. & Chen, X., 2002: Decision Tree Classification of a Forest Using AVIRIS and Multi-Seasonal TM Data. – **11th** JPL Airborne Earth Science Workshop.

Mannel, S., 2003: Contribution of remote sensing to forest inventories in the Black Hills, SD – a focus on reference data acquisition, decision, tree classification and hyperspectral versus multispectral data. – South Dakota School of Mines and Technology, Ph.D. Dissertation.

Mannel, S., Price, M. & Hua, D., 2006: A method to multiply training and test data. – International Journal of Remote Sensing **27** (3–4): 623–627.

Mannel, S., Price, M. & Hua, D., 2011. Impact of reference datasets and autocorrelation on classification accuracy. – International Journal of Remote Sensing **32** (19): 5321–5330.

Niemann, K., Goodenough, D. & Bhogal, A., 2002: Remote sensing of relative moisture status in old growth Douglas fir. – International Journal of Remote Sensing **23** (2): 395–400.

Okin, G., Roberts, D., Murray, B. & Okin, W., 2001: Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. – Remote Sensing of Environment **77:** 212–225.

Quinlan, J., 1996: Improved Use of Continuous Attributes in C4.5. Journal of Artificial Intelligence Research **4:** 77-90.

Quinlan, J., 2002: See5/C5.0. RuleQuest Research Pty Ltd, St Ives NSW 2075, Australia.

Read, B., 2000: Data mining and science? CLRC Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, UK, http://www.ercim.org/publication/ws-proceedings/12th-EDRG/EDRG12_Re.pdf (30.5.2012).

Riano, D., Chuvieco, E., Ustin, S., Zomer, R., Dennison, P., Roberts, D. & Salas, J., 2002: Assessment of vegetation regeneration after fire throughout multitemporal analysis of AVIRIS images in the Santa Monica Mountains. – Remote Sensing of Environment **79:** 60–71.

RIS, 2000: Rocky Mountain Resource Information System (RIS) User Manual.

Schriever, J. & Congalton, R., 1995: Evaluating Seasonal Variability as an Aid to Cover-Type Mapping from Landsat Thematic Mapper Data in the Northeast. – Photogrammetric Engineering & Remote Sensing **61** (3): 321–327.

Singhroy, V., Saint-Jean, R., Levesque, J. & Barnett, P., 2000: Reflectance spectra of the boreal forest over mineralized sites. – IEEE Geoscience and Remote Sensing Symposium, IGARSS **2000** (4): 1379–1381.

Swain, P. & Hauska, H., 1977: The Decision Tree Classifier: Design and Potential. – IEEE Transaction on Geoscience Electronics **15** (3): 142–147.

Thiemann, S. & Kaufmann, H., 2002: Lake water quality monitoring using hyperspectral airborne data – a semiempirical multisensor and multitemporal approach for the Mecklenburg Lake District, Germany. – Remote Sensing of Environment **81:** 228–237.

USTIN, S. & XIAO, Q., 2001: Mapping successional boreal forests in interior central Alaska. – International Journal of Remote Sensing **22** (9): 1779–1797.

WILLIAMS, A. & HUNT JR., E., 2002: Estimation of leafy spurge cover from hyperspectral imagery using mixture tuned matched filtering. – Remote Sensing of Environment **82:** 446–456.

WOLTER, P., MLADENOFF, D., HOST, G. & CROW, T., 1995: Improved Forest Classification in the Northern Lake States Using Multi-Temporal Landsat Imagery. – Photogrammetric Engineering & Remote Sensing **61** (9): 1129–1143.

WU, X., KUMAR, V., QUINLAN, J., GHOSH, J., QIANG, Y., MOTODA, H., MCLACHLAN, G., ANGUS, N., BING, L., PHILIP, S., ZHI-HUA, Z., STEINBACH, M., HAND, D. & STEINBERG, D., 2006: Top 10 algorithms in data mining. – Knowledge and Information Systems **14:** 1–37.

WYLIE, B., MEYER, D., CHOATE, M., VIERLING, L. & KOZAK, P., 2000: Mapping Woody Vegetation and Eastern Red Cedar in the Nebraska Sand Hills Using AVIRIS. – Ninth JPL Airborne Earth Science Workshop **9:** 491–500.

Addresses of the Authors:

Dr. SYLVIO MANNEL, Cottey College, Environmental Studies, USA-64772 Nevada, MO, Tel.: +1-417-667-6333, e-mail: smannel@cottey.edu

Dr. MARIBETH H. PRICE, Professor, Dept of Geology and Geological Engineering, South Dakota School of Mines and Technology, 501 E. St. Joseph St., Rapid City, SD 57701, USA, Tel. 1-605-394-2468, Fax: 1-605-394-6703, e-mail: maribeth.price@sdsmt.edu

Manuskript eingereicht: Dezember 2011
Angenommen: Mai 2012

PFG 2012 / 5, 0613–0630
Stuttgart, October 2012

Article

Validation and Automatic Repair of Planar Partitions Using a Constrained Triangulation

Ken Arroyo Ohori, Hugo Ledoux & Martijn Meijers, Delft, The Netherlands

Keywords: triangulation, repair, planar partition, topological data structure, polygonal coverage

Summary: Planar partitions are frequently used to model, among others, land cover, cadastral parcels and administrative boundaries. In practice, they are often stored as a set of individual polygons to which attributes are attached, e.g. with the *Simple Features* paradigm, causing different errors and inconsistencies, e.g. gaps, overlaps and disconnected polygons, which are introduced during their creation, manipulation and exchange. These errors severely hamper the use of planar partitions in other software, e.g. due to false assumptions causing erroneous calculations. Existing approaches to validate planar partitions involve first building a planar graph of the polygons and enforcing constraints, then repair is done by snapping vertices and edges of this graph. We argue that these approaches have many shortcomings in terms of complexity, numerical robustness and difficulty of implementation, and do not guarantee valid results. Furthermore, they are semi-automatic, requiring manual user intervention. We propose in this paper a novel method to validate and *automatically* repair planar partitions. It uses a constrained triangulation of the polygons as a base – which by definition is a planar partition – and only simple operations are needed, i.e., labelling of triangles, to both validate and repair. Perhaps the biggest advantage of our method is that we can guarantee that a planar partition is valid after repair. In the paper we describe the details of our method, our implementation, and the experiments we have done with real-world datasets. We show that our implementation scales to big datasets and that it offers better capabilities and overall performance than existing solutions.

Zusammenfassung: *Validierung und automatische Korrektur von planaren Graphen unter Benutzung von bedingter Triangulation*. Planare Graphen werden neben anderen Methoden oft für die Modellierung von Landnutzung, Kataster und Verwaltungsgrenzen verwendet. In der praktischen Anwendung werden die Flächeneinheiten oft als selbstständige Polygone, denen Attribute zugeordnet sind, gespeichert, z.B. als Simple Features. Die Anwendung dieser Methode verursacht oft Inkonsistenzen, z.B. Lücken, Überlappungen und nicht verbundene Polygone. Diese Fehler behindern die Weiterverarbeitung der Flächenaufteilung in anderen Anwendungen, z.B. wenn falsche Annahmen den Folgeberechnungen zu Grunde gelegt werden müssen. Existierende Verfahren zur Validierung von planaren Graphen nutzen oft einen bedingten planaren Graphen. Die Korrektur wird dann durch Snapping auf Knoten und Kanten des Graphen durchgeführt. Wir sind der Meinung, dass dieses Verfahren Nachteile bezüglich Komplexität, numerischer Stabilität und Implementierung hat. Vor allem kann ein richtiges Ergebnis nicht garantiert werden. Darüber hinaus sind die Algorithmen oft nur halbautomatisch. Daher schlagen wir in diesem Artikel ein neues Verfahren zur Validierung und automatischen Korrektur von planaren Graphen vor. Es beginnt mit einer bedingten Triangulation der gegebenen Polygone, die per Definition ein planarer Graph ist. Danach sind nur einfache Operationen zur Beschreibung der Dreiecke, zur Validierung und zur Korrektur erforderlich. Wahrscheinlich ist der größte Vorteil unserer Methode, dass eine gültige Raumaufteilung im Ergebnis garantiert wird. Im Artikel beschreiben wir unsere Methode, die Implementierung und die Anwendung auf reale Datensätze. Wir zeigen, dass unsere Methode auch für große Datensätze geeignet und sogar leistungsfähiger als andere Methoden ist.

---

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0143 1432-8364/12/0143 $ 4.50

1 Introduction

Planar partitions are frequently used in GIS to model concepts such as land cover, the cadastre, or the administrative boundaries of a given country. As shown in Fig. 1, a planar partition is a subdivision of a polygonal subset of the plane into non-overlapping polygons. In practice, planar partitions are often represented, and stored in a computer, as a set of individual polygons to which one or more attributes are attached, and the topological relationships between polygons are not explicitly stored, i.e., shared boundaries are thus represented and stored twice. The preferred method of practitioners is representing polygons according to the *Simple Features* specifications (OGC 2006), for instance as an Esri *shapefile* (Esri 1998) or in a database, such as PostGIS (PostGIS 2012).

If a planar partition is stored as a set of individual polygons, then in practice errors, mistakes and inconsistencies will often be introduced when the planar partition is built, updated or exchanged. Examples of common errors are: overlapping polygons, gaps between polygons, and polygons not connected to the others. This can be, among others, due to human error, the use of floating-point arithmetic, or limited precision (Schirra 1997). These errors can have catastrophic consequences for practitioners since most software and algorithms using planar partitions as input assume that this input is valid. At best erroneous results are returned, at worst it causes a software failure, often without any warning to the user. Moreover, such problems are often not visible at the scale that the data is usually viewed, exacerbating the problem (Laurini & Milleret-Raffort 1994).

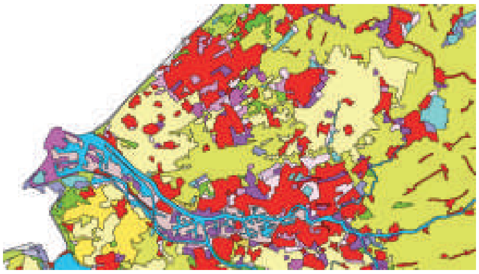


Fig. 1: Part of the Corine Land Cover dataset for the region around Delft, The Netherlands.

Solving that issue entails working on two related problems: (1) how to identify errors in a planar partition; and (2) how to repair these errors. As described in section 2, both problems have been tackled in the past with the creation of a planar graph of the input. The validation, the simpler of the two problems, is usually implemented as a set of topological and geometrical constraints that the planar graph must have. As for the repair, it is usually performed by *snapping* together the vertices and edges of the graph, or by using topological information. As we discuss in section 2, both approaches have drawbacks for users: the former method is error-prone, i.e., topological inconsistencies can be created, and the latter is only semi-automatic, because in practice real-world datasets can easily contain several hundred errors.

We present in this paper a novel method to both validate and *automatically* repair planar partitions stored according to the *Simple Features* specifications. Our method, which is an extension of our preliminary results (Ledoux & Meijers 2010), uses a constrained triangulation (CT) of the polygons as a support – which is by definition a planar partition – and both the validation and the repair functions are performed with relatively simple operations. These are the labelling of triangles, and standard graph traversal algorithms (such as depth-first search). Since errors are repaired by relabelling triangles (vertices are never moved), we can guarantee that a given repair operation will preserve the topological consistency of the whole planar partition. We describe in section 3 how the CT is used, how the polygons are labelled, how the validation is performed, and how we can automatically repair a planar partition. Moreover, we describe six different repair operations that can be used to obtain different output.

We have implemented the method in C++, and its most relevant details are discussed in section 4. Our software takes as input polygons stored according to the *Simple Features*s specification, validates them, repairs them if they contain errors, and returns a new set of polygons that is guaranteed to be a valid planar partition. We also report in that section our experiments with several real-world datasets, some of them rather large, and we compare our method and its implementation to alternatives, both for validation and for repair. Finally, we discuss the advantages and disadvantages of using our method and the conclusions drawn from this in section 5.

2 Related Work

Since a planar partition is formed by a set of individual polygons, we first discuss what a valid polygon is in our context, and then we review existing methods to validate and repair planar partitions.

2.1 *Simple Features and Validity of Simple Polygons*

While there are several definitions of what constitutes a valid polygon (van Oosterom et al. 2004), we use in the following the standard *Simple Features* (OGC 2006), with the addition of the ISO 19107 Spatial schema (ISO 2003) polygon orientation rules. *Simple Features* defines a polygon as follows: "A Polygon is a planar Surface defined by 1 exterior boundary and 0 or more interior boundaries. Each interior boundary defines a hole in the Polygon.". In the specification, six assertions are given that together define a valid polygon. Essential for a valid polygon is that the boundaries of the polygon must define one connected area. Additionally, a polygon can contain holes. We say that the exterior boundary of the polygon is the *outer ring*, and a hole is an *inner ring*. These holes can be filled by one or more polygons, which can recursively contain holes, which are filled by other polygons. Observe also that holes are allowed to interact with each other and the outer boundary under certain conditions, e.g. they are allowed to touch at one point, as long as the interior of the polygon stays one connected area. Each polygon is stored independently from other polygons, and it is not possible to store topological relationships between the polygons. The ISO 19107 standard (ISO 2003) is more ambiguously defined, but it does establish orientation rules (counter clockwise for the outer ring, clockwise for the inner ones), which we use in our output.

The validation of a single polygon is possible with different libraries, GEOS (GEOS 2012) and JTS (Vidid Solutions 2012) being two widely used open-source examples.

The repair of single polygons is a less documented topic than their validation. Different software vendors offer tools to help identify and semi-automatically repair broken polygons. Examples are ST_MakeValid() from PostGIS and the constraints in 1Spatial Radius Topology. The method we present in this paper has been adapted to automatically repair common errors in individual polygons, e.g. wrong ring orientation, or holes that split the interior of a polygon (Ledoux et al. 2012). However, we focus in this paper on the validation and repair of planar partitions only and we assume that the input polygons are individually valid.

2.2 *Validation of a Planar Partition Using a Planar Graph*

Assuming that individual polygons have been deemed to be valid, it is possible to test the validity of a planar partition by identifying the two types of invalid configurations: overlaps and gaps.

If individual polygons are checked without building a planar graph or an indexing structure, finding overlaps involves checking whether any possible pair of polygons overlap. This is a computationally expensive operation to make because of its quadratic behaviour, even when heuristics to speed up the process are used (Badawy & Aref 1999, Kirkpatrick et al. 2002). Additionally, robustness issues are significant in polygon intersection tests (Hoffmann et al. 1988). Finding the potential gaps in a planar partition is even more problematic. For this, computing the union of the entire set of polygons is required, which is also computationally expensive (Margalit & Knott 1989, Rivero & Feito 2000).

The validation process can be sped up by first building a planar graph of the input polygons, which is afterwards checked for consistency. It should first be noticed that while different approaches are available to construct a planar graph (Shamos & Hoey 1976, van Roessel 1991), it is still sometimes difficult, especially if the polygon contains holes. The graph of the boundary can then be unconnected and extra machinery is necessary to represent the knowledge of holes in the graph structure. The fact that holes are also allowed to touch complicates the task of validation even further, since holes cannot be assumed to form an unconnected planar graph.

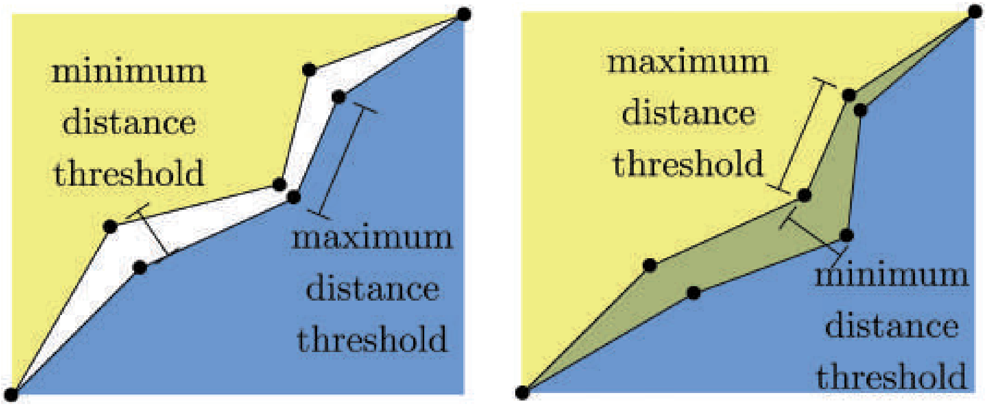
Based on this graph, Plümer & Gröger (1997) specify a list of minimal mathematical axioms that can be used to check the validity of a planar partition: no dangling edges, no zero-length edges, planarity, no holes, no self-intersections, no overlaps, and having a connected graph. It is important to note that Plümer & Gröger base their axioms on concepts from graph theory, but they also highlight the fact that a graph-based approach alone is not enough: the graph has to be augmented with geometrical knowledge (each vertex has geometry attached, i.e., the coordinates of points have to be stored). Validation is thus underpinned by both geometrical and topological concepts and systems thus have to deal with those two concepts at the same time. The method we propose in this paper – using a constrained triangulation – permits us to do exactly this: to embed both geometry and topology in the same structure.

2.3 *Repair Using Point and Edge Snapping and Splitting*

The most common method for planar partition repair is based on the assumption that polygons *approximately* match each other at their common boundaries. This implies that they should be within a certain distance of each other along those edges. If, additionally, all parts further apart than this value are known not to be common boundaries, it is possible to "snap" together polygons that are closer to each other than this threshold, while keeping the rest untouched. This method of planar partition repair is available in many GIS packages, including ArcGIS (ArcGIS 2012), FME (FME 2012), GRASS (GRASS 2012) and Radius Topology (1Spatial 2012).

Since thresholds are central to this method, it is of utmost importance to select a good threshold value, something that is completely different in each dataset. For planar partition repair to be successful using this method, such a threshold should be chosen in a careful manner, and always comply with a few conditions. These have been summarised as follows:

- Adjacent polygons should not be further apart than this threshold along any part of their common boundaries (shown as the maximum threshold in Fig. 2 (a). Otherwise, gaps cannot be fixed.
- Adjacent polygons should not overlap each other in areas which are further inwards



(a) Gaps between polygons (b) Overlapping polygons

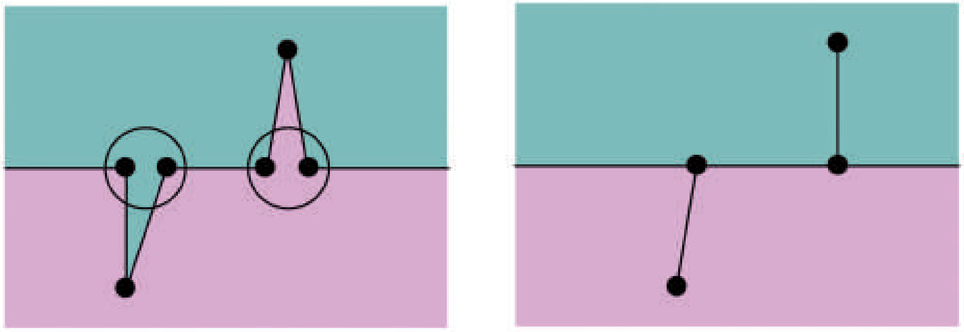
Fig. 2: Defining a threshold for vertex, edge and face snapping. The threshold to be used should be larger than the largest minimum distance between the matching boundaries, and smaller than the minimum distance between vertices.

than this threshold from their common boundaries shown as the maximum threshold in Fig. 2 (b). Otherwise, overlaps cannot be fixed.

- None of the vertices of a polygon should be closer to each other than this threshold, including non consecutive vertices shown as the minimum thresholds in Fig. 2. Otherwise, they might be snapped together, creating repeated vertices, disjoint regions, or various topological problems.
- None of the vertices of a polygon should be closer than this threshold to any non incident edge. Otherwise, they might be snapped together, creating disjoint regions or various topological problems.

This threshold value is usually manually determined, either by trial and error, or by analysing certain properties of the dataset(s) involved, e.g. point spacing, precision, or map scale. However, it is often hard to find an optimal threshold for a certain dataset, since ensuring that it works well for every part of a dataset is unrealistic. Moreover, sometimes such a threshold does not even exist, e.g. because point spacing in some places might be smaller than the width of the gaps and overlaps present.

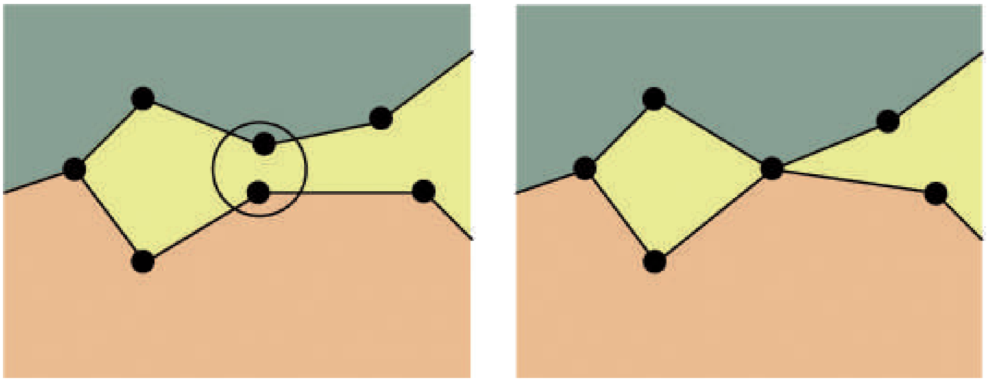
Since the aforementioned conditions are frequently not met or are not checked beforehand, and since it is still necessary to perform repair of a dataset, snapping is often per-



(a) Before snapping

(b) After snapping (invalid topology)

Fig. 3: Spikes and punctures can be created by snapping, since the bases of these elongated forms (encircled) might be narrower than the threshold, but its length is not.



(a) Before snapping

(b) After snapping (yellow peninsula cut off)

Fig. 4: Polygons can be split by snapping, since some parts of them might be narrower than the threshold (encircled). While this result does not create an invalid planar partition, it can change the number of polygons present and their topological relations, and can therefore be undesirable.

formed nevertheless, possibly creating invalid polygons and/or planar partitions, or significantly changing the topology of the existing features. Two examples of this phenomenon are shown in Figs. 3 and 4.

While these examples show that snapping is not problem-free, it is important to note that commercial GIS packages often implement more complex snapping options such as point-to-edge, edge-to-edge, or using a reference dataset. These options can help to solve a problematic case, but can also have undesired consequences, such as changing the topology of the polygons. Another problem is that post-processing operations to clean resulting polygons might be required, e.g. disposing of polygons with small areas, removing redundant lines, thresholds for minimum angles etc., which could again create invalid configurations, requiring iterative validation or repair processes.

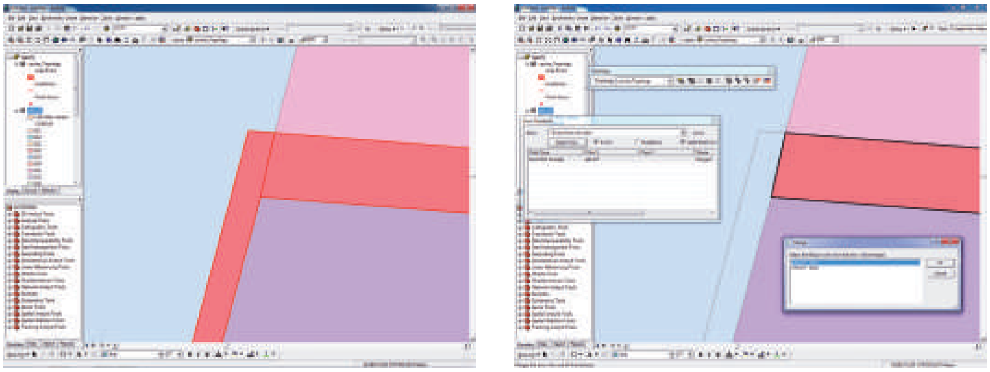
2.4 *Repair Using Topological Information*

A different approach for planar partition repair, based on topological information, is available in some software.

GRASS also creates a graph, using edges as a base structure instead of triangles, and could be used to (manually) detect overlaps and gaps based on the number of labels using the functions v.what and d.what.vect. However, there is no simple automated procedure to get the number of labels at a certain point, which makes it very cumbersome and time consuming to use GRASS for this purpose.

Meanwhile, ArcGIS provides a more complete solution, using a method similar in some ways to the one developed and described in section 3. It involves using the Geodatabase feature of the software with some combined validation rules, e.g. must not overlap and must not have gaps. However, fixing every problematic area in an appropriate manner requires extensive user intervention (Fig. 5), since the best choice for each case depends on the specific configuration of the error.

Since both the aforementioned programs do not offer an automated process to correctly solve this problem and GRASS lacks the ability to visualise problem areas, they are not really comparable to our solution. A planar partition can easily contain tens of thousands of polygons, possibly generating thousands of errors, which need to be checked and repaired semi-automatically.



(a) Viewing a topology error in ArcGIS

(b) Assigning an overlapping region to one of the polygons involved

Fig. 5: Planar partition repair in ArcGIS. The user is expected to zoom in to a particular error, analyse the situation, e.g. by looking at the properties from the surrounding polygons, and make a decision to assign the problematic region to a certain polygon. More than 11,000 errors were detected in this tile of the Corine dataset.

3 Validation and Automatic Repair Using a Constrained Triangulation

Our approach to validation and automatic repair of planar partitions uses a constrained triangulation (CT) as a supporting structure because it has many good properties, including the following:

- It is by definition a planar partition. Therefore, as long as we keep the information about which polygon each triangle belongs to, the reconstructed polygons will be either a valid planar partition, or multiple ones.
- It can be built quickly, in $\boldsymbol{O}(\boldsymbol{n} \log \boldsymbol{n})$ with a variety of approaches (Clarkson et al. 1992, Guibas & Stolfi 1985, Mücke et al. 1999). The actual computational complexity can be $\boldsymbol{O}(\boldsymbol{n} \log \boldsymbol{n}) + \boldsymbol{k}$, with $\boldsymbol{k}$ being the number of edge-edge intersections, which could conceivably even be $\boldsymbol{n}^2$. However, $\boldsymbol{k} \ll \boldsymbol{n}$ for most GIS datasets.
- Changes to the triangulation, e.g. adding a new constrained edge is local, and therefore fast.
- Constrained edges can usually be added in constant time, being only significantly slower and more complex when there is an intersection with an existing constrained edge (Shewchuk 1997b).
- Implementation-wise, several stable and fast triangulation libraries exist, including CGAL (2012), Triangle (Shewchuk 1997a) and GTS (2006).

The general workflow of our approach to both validate and repair a planar partition is as follows:

- the CT of the input segments forming the polygons is constructed;
- each triangle in the CT is labelled with the label of the polygon inside which it is located;
- problems are detected by identifying triangles having no or multiple labels, and by verifying the connectivity between triangles;
- repairing of the problems is made by re-labelling triangles to ensure that each triangle has exactly one label;
- extracting the polygons from the triangulation (polygons modelled with the *Simple Features* specifications).

As mentioned previously, for this workflow we assume that each input polygon is individually valid. We describe in the following section the concepts needed and we give a detailed description of the different steps.

3.1 *Triangulation of a Polygon and Constrained Triangulation*

A triangulation subdivides an area into non-overlapping triangles. Using a constrained triangulation, every line segment that defines the boundary of a polygon, is ensured to appear as an edge in the triangulation. It is known that

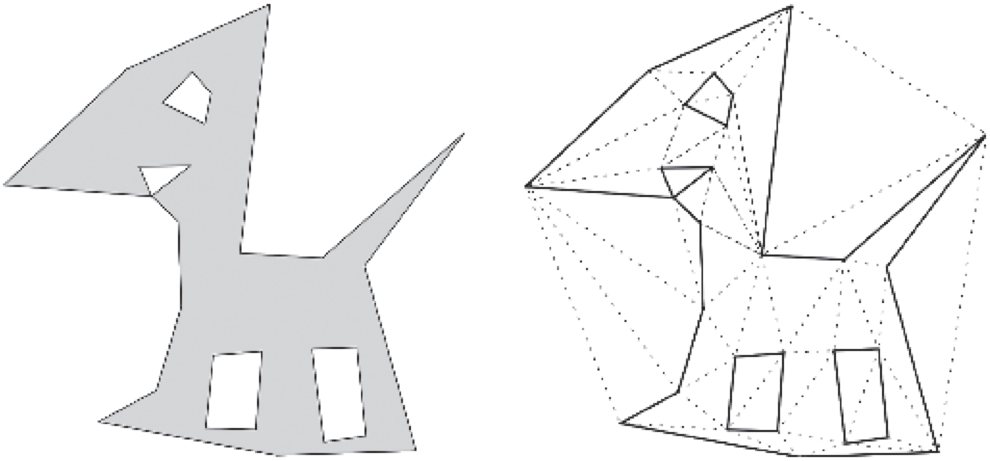


Fig. 6: (a) A polygon with 4 holes; (b) the constrained triangulation of the segments of this polygon.

any polygon (also with holes) can be triangulated without adding extra vertices (de Berg et al. 2008, Shewchuk 1997a). Fig. 6 shows an example.

In our approach, the triangulation is performed by constructing incrementally a CT of all the segments representing the boundaries (outer + inner) of each polygon. If the set of input polygons forms a planar partition, then each segment will be inserted twice except those forming the outer boundary of the set of input polygons. This is usually not a problem for triangulation libraries because they ignore points and segments at the same location as is the case with the solution we use, see section 4. When segments are found to intersect, they are split with a new point created at the intersection. This is the only situation in which the generation of new points is required.

Notice that our approach requires only a constrained triangulation, which not necessarily includes the Delaunay criterion (Shewchuk 1997a). However, having well-shaped triangles is useful for repair purposes and does not significantly increase the processing time. Therefore our implementation actually constructs and uses a constrained *Delaunay* triangulation.

3.2 *Labelling the Triangles of a Planar Partition*

Labelling a triangle means assigning to it the label(s) of the polygon(s) that it belongs to. If two input polygons overlap, each triangle in the overlapping region should have the labels of the two polygons.

In our previous work on validation (Ledoux & Meijers 2010), we used the centroid of a polygon to start the labelling process, but this method is prone to errors if for instance the calculated centroid is outside or on the boundary of the polygon and does not allow us to differentiate between gaps and overlaps.

To solve these problems, we store information about the constrained edges of the CT. Since it is known that the input rings are closed and have a known orientation according to the ISO 19107 orientation rules, it is also known on which side of a certain line segment the interior of the polygon lies. This property is used for robust labelling of each polygon. Triangles adjacent to the outer ring of the polygons are labelled first, and this is later expanded to triangles further in the interior of the polygon, recursively labelling adjacent unlabelled faces as long as no constrained edges are crossed. After this operation, all triangles that are part of any polygon have been labelled, with overlapping regions having multiply labelled triangles. However, holes are then indistinguishable from triangles outside the planar partition, since both have zero labels. Therefore, a special label is created for all triangles outside the planar partition, referred to as the "universe" label, which are labelled by recursively labelling adjacent triangles from any triangle known to lie in the exterior of the planar partition. To achieve this we exploit the concept of the "far-away point" (Liu & Snoeyink 2006); which is used by several implementations and is also known as the "big triangle" (Facello 1995).

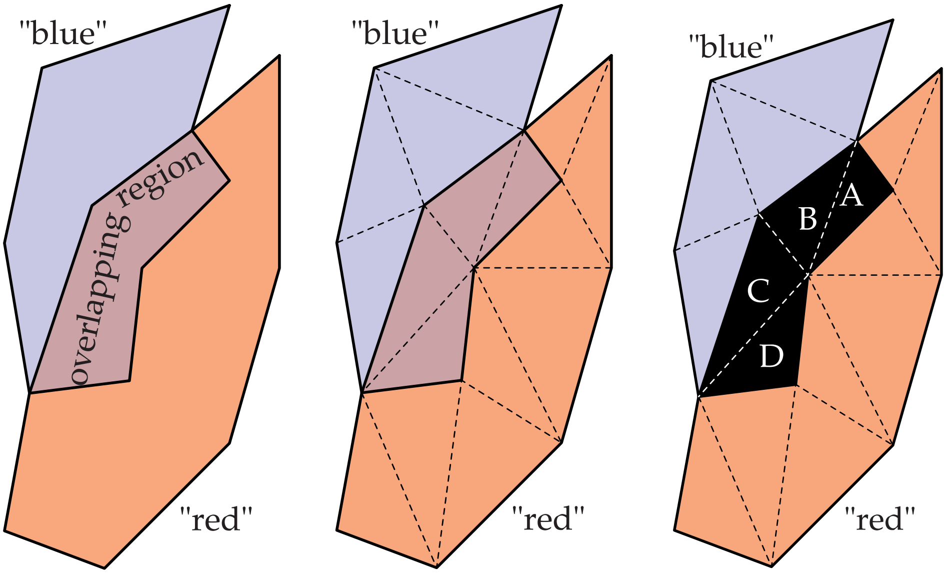
3.3 *Validation*

If the set of input polygons forms a planar partition then all the triangles will be labelled with one and only one label. The problems are easily detected:

- **Gaps** are detected by finding triangles without any labels.
- **Overlaps** are detected by finding triangles with two or more labels.
- **Disjoint regions** are detected by identifying regions separated by the "universe" label. This is done by starting at a given triangle, and doing a breadth-first search on the dual of the triangulation, without visiting the triangles labelled as "universe". If all triangles can be reached, no polygons of the planar partition are disjoint.

3.4 *Repair Operations*

The greatest benefits of using a labelled triangulation for planar partition repair stem from the fact that while repair operations are performed, the validity of the planar partition is always kept, together with the integrity of the data. Unlike snapping, vertices are not re-



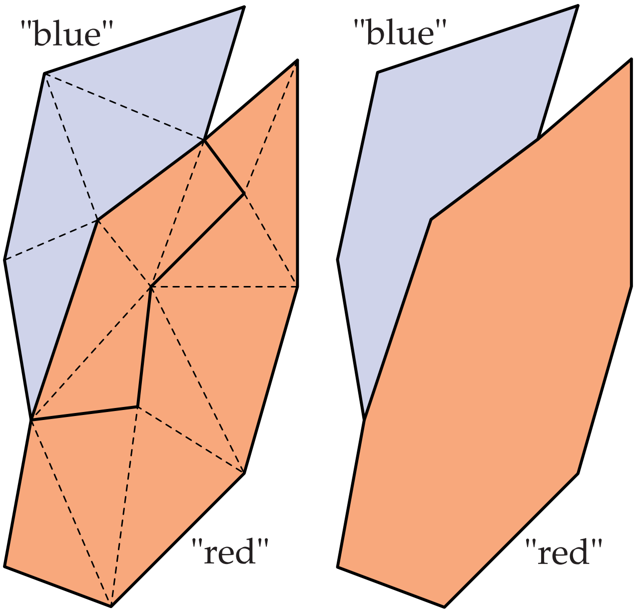
(a) A set of overlapping polygons that needs to be repaired.

(b) The polygons are triangulated and each triangle is labelled with the polygons it belongs to.

(c) An invalid region is found, defined as a set of contiguous polygons with an equivalent set of 0 or ≥2 labels.

Triangle	Repair
A	"red"
B	"red"
C	"red"
D	"red"

(d) The neighbours of the region are analysed to see how it should be re-labelled, and the result is saved in a list.



(e) The triangles are relabelled according to the list.

(f) The output polygons, which now have no gaps or overlaps.

Fig. 7: The steps in a generic repair operation.

Tab. 1: The repair operations currently implemented in our software. The types of operations are defined according to the map algebra classification by Tomlin (1994).

Repair operation	Type	Criteria
Triangle by priority list	Varies	Label that has the highest priority according to a predefined priority list
Triangle by number of neighbours	Focal	The label present in the largest number of adjacent faces, overlaps included
Triangle by absolute majority	Focal	Label present in two or more valid adjacent faces
Triangle by longest boundary	Focal	Label present along the longest portion of the boundary of the adjacent faces
Regions by longest boundary	Focal of zonal	Label present along the longest portion of the boundary of the adjacent faces
Regions by random neighbour	Focal of zonal	Random label from the adjacent faces

(a) The original polygons

(b) Repaired each triangle using the label adjacent along the longest boundary from the neighbouring triangles

(c) Repaired each region using the label adjacent along the longest boundary from the neighbouring triangles

(d) Repaired each region using a random label from the neighbouring triangles

Fig. 8: Different repair operations used in the two polygons for the Arribes del Duero Natural Park in Spain (red) and the International Douro Natural Park in Portugal (green). All of them can be considered best by a certain criterion, like (b) preserving the area ratio between the two polygons, (c) smoothness of the boundary, or (d) a balance between the two.

quired to be moved during the process, and unlike snapping, repair is performed using *local* criteria, instead of global ones because the snapping-threshold is usually the same for the whole dataset. This comes in contrast to other methods, where care needs to be taken to ensure that the geometric or topologic validity is kept.

Fig. 7 shows the standard steps required in a repair operation. In order to avoid order dependency when repairing, the repair operation is always performed after all choices for label-assigning have been made.

In particular, we propose six different repair operations that can be used to fix gaps and overlaps. These are shown in Tab. 1 and all imply re-labelling triangles.

Four of them use triangles as a base, i.e., the label assigned is based only on that of its three neighbouring triangles, which is fast and modifies the area of each input polygon the least. Despite their simplicity, they offer substantial control over the results. For instance, the first two operations only differ from each other in their handling of overlapping faces; but *triangle by number of neighbours* is better for large overlapping regions, while *triangle by absolute majority* is better in fixing small problems. Having well-shaped (Delaunay) triangles is most useful for triangle-based repair functions. Two of them use regions of adjacent triangles with equivalent sets of labels, which is slower than a triangle-based method but yields results that can be cartographically more pleasing. An interesting repair operation for practitioners is the one in which a *priority of labels* is used, i.e., in case of gaps/overlaps the labels in the triangle (overlap) or in the adjacent polygons (gap) are ordered according to a user-defined priority, and the highest priority is assigned to the problematic triangles. Notice that these repair operations can be used one after the other (in a hierarchical manner), for instance if first the repair according to the longest boundary is used but one zone has two or more boundaries with exactly the same length, then the deadlock can be solved by choosing one randomly. A sample of the results obtained with different repair operators is also shown in Fig. 8.

More repair operations based on extensions of the idea of labelling triangles/regions can be further developed. For instance, triangles could be split to subdivide an area with problems as in Bader & Weibel (1997), or sliver triangles/regions could be discarded (and then filled during repair).

3.5 *Extraction of Polygons from a Triangulation*

Starting from a labelled triangulation, it is possible to reconstruct the individual polygons of a planar partition to conform to valid polygons according to the ISO 19107 and the *Simple Features* specifications, which allows users to incorporate automatic validation and repair in their workflow.

We do this operation polygon by polygon and start at an unprocessed triangle. Then we visit all the connected triangles that have the same label, reconstruct the polygon, and mark the triangles as processed. Note that since all of these triangles are connected, the outer and inner boundaries of a polygon are all simple, i.e., non self-intersecting. We repeat this operation until all triangles have been processed.

For each polygon, we have to recover not only its outer boundary, but also its inner boundaries, which are not connected. Observe that we cannot simply follow the original constrained edges as these do not have any meaning after a planar partitions was repaired; the boundaries of the repaired polygons are instead formed by edges incident to two triangles having different labels.

For each polygon, we start at a triangle, and move on to triangles having the same label. As the process goes on, a single polyline that runs along all boundaries of the polygon is generated. This involves a depth-first search (clockwise) that recursively reaches until the boundary of a polygon, returning a long chain of edges in a procedure similar to following the boundary edge by edge. The procedure is shown step by step in Fig. 9.

The polyline created with this method has "bridges", which allow us to keep all inner boundaries (holes) connected with the outer boundary, in a manner that keeps the interior connected as well. These help to preserve connectivity and the relations between different boundary types (outer and inner), but are re-

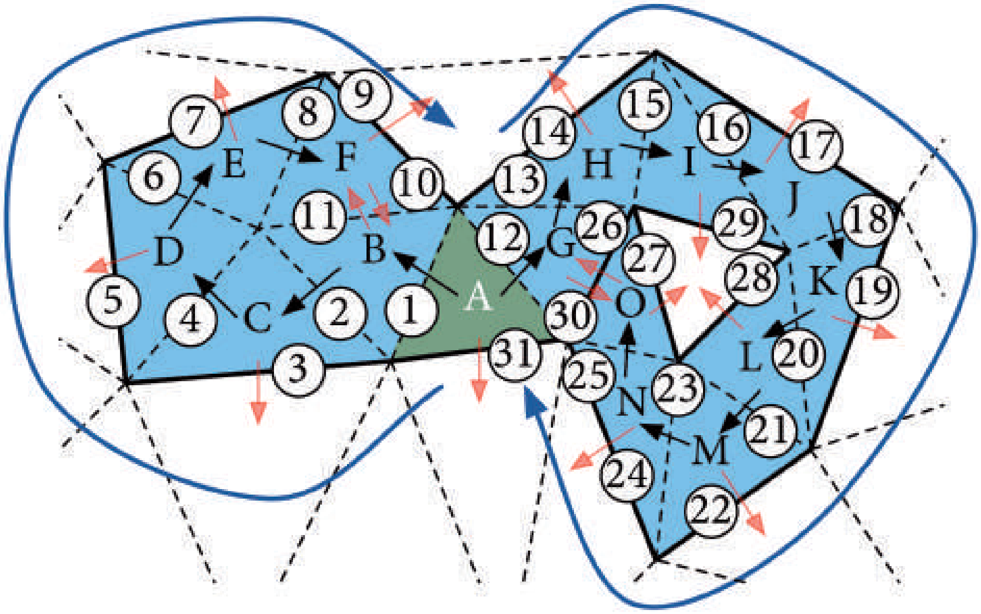


Fig. 9: The traversal order when navigating through triangles in a clockwise manner. Starting at △ A and the edge between △ A and △ B, the operations occur according to the encircled numbers. Black arrows denote when an unprocessed triangle is found, red arrows when it is not. Notice how the traversal is performed clockwise (shown in dark blue arrows) for both cases, despite starting from different sides.

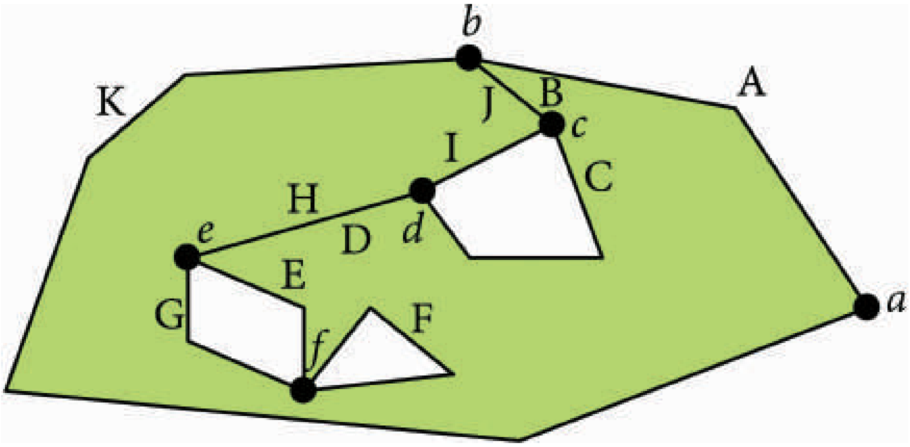


Fig. 10: Processing the polyline shown starting from vertex $\boldsymbol{a}$ and moving towards polyline $\overline{A}$, the following operations occur: (1) $\overline{A}$ unclosed → push, (2) $\overline{B}$ unclosed → push, (3) $\overline{C}$ unclosed → push, (4) $\overline{D}$ unclosed → push, (5) $\overline{E}$ unclosed → push, (6) $\overline{F}$ closed → store, (7) pop → $\overline{EG}$ closed → store, (8) pop → $\overline{DH}$ degenerate → erase, (9) pop → $\overline{CI}$ closed → store, (10) pop → $\overline{BJ}$ degenerate → erase, (11) pop → $\overline{AK}$ closed → store.

moved later in the process to conform to the *Simple Features* specifications. Also, its orientation conveys the information of whether a section of it is part of an inner or an outer boundary.

This polyline is processed with a stack-based algorithm that generates separate closed rings for the outer boundary and each of the inner boundaries, collapsing the "bridges" that were generated. In order to do this, the polyline is cut at the positions where more than two edges join, and these are joined in the correct order by keeping track of (yet) unclosed rings. When a new segment is processed, it can be one of three options: one that completes a ring, one that is part of a bridge, or one that is not yet closed. Closed rings are stored and bridges are removed, while unclosed rings are saved in the stack until they can be popped to form a closed ring together with a new segment. This is shown in Fig. 10.

4 Implementation and Experiments

4.1 *Implementation*

An implementation of the algorithms described in section 3 was written in the C++ programming language, using external libraries for some functionality. The developed software is called **pprepair**, and is open source and freely available at http://tudelft-gist.github.com/pprepair/. C++ was selected in order to have good control with regards to low level details and to achieve good performance, which makes it possible to compare it with existing solutions. The libraries used are: the OGR Simple Features library (GDAL 2012) which allows input and output from a large variety of data formats common in GIS; and CGAL (CGAL 2012) which has support for many robust spatial data structures and the related operations; we use its constrained triangulation module.

4.2 *Experiments with Real-World Planar Partition Datasets*

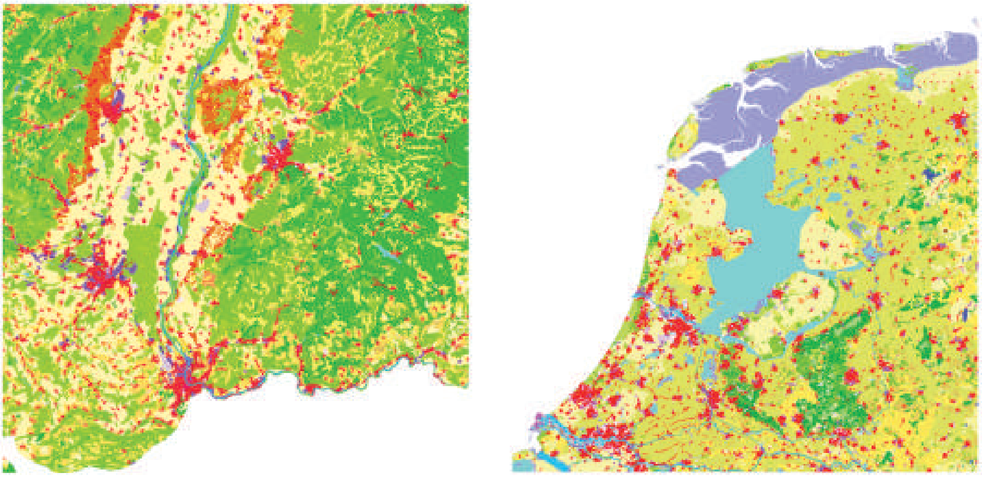
We have made experiments with four freely available real-world datasets, i.e., we have validated and automatically repaired them with our implementation; the overview of these datasets is shown in Fig. 11 and their properties in Tab. 2. The datasets are the following:

- **E41N27** Corine 2000 (Corine 2000) tile E41N27, which contains a shifted polygon (by about 10 cm), creating many small gaps and overlaps in the dataset. The snapping threshold has been set to 1 m.
- **4tiles** Corine 2000 tiles E39N32, E39N33, E40N32 and E40N33, which are known to have long and thin overlapping regions (1 mm) with each other. The snapping threshold has been set to 1 cm.
- **16tiles** 16 adjacent Corine 2000 tiles: E39N30, E39N31, E39N32, E39N33, E40N30, E40N31, E40N32, E40N33, E41N30, E41N31, E41N32, E41N33, E42N30, E42N31, E42N32, E42E33. Between some of them are gaps, others are overlapping. However, all match within a few centimetres. The snapping threshold has been set to 10 cm.
- **Mexico** 1:1,000,000 scale land cover dataset from INEGI consists of over 26,000 polygons. It is mostly valid according to the Shapefile specification, but contains some very large polygons, with tens of thousands of vertices, which presents difficulties to most existing software.

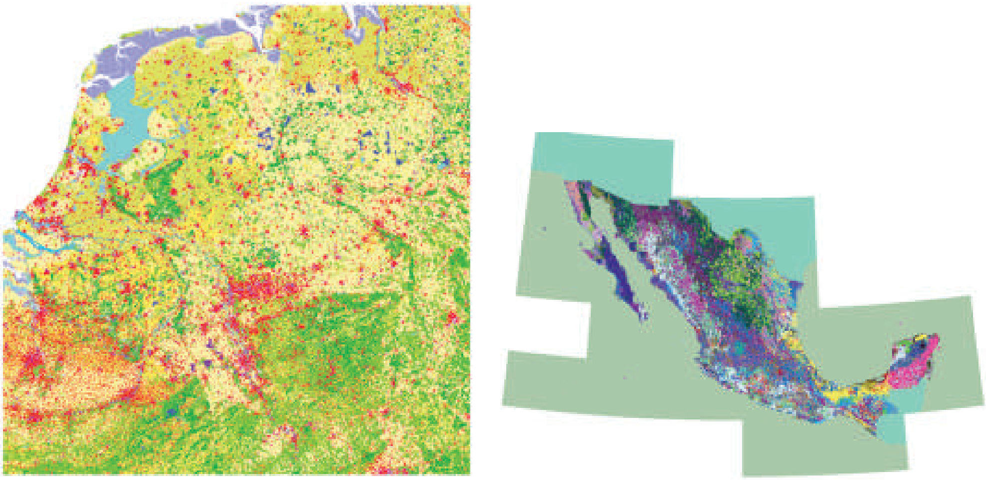
As a comparison, we have also tried to perform the same operations with other available software. While the capabilities for pla-

Tab. 2: Properties of the datasets used for the experiment.

	# polygons	# pts	# pts in largest polygon	avg # pts per polygon
E41N27	14,969	496,303	26,740	33.7
4tiles	4,984	365,702	16,961	74.7
16tiles	63,868	6,622,133	95,112	103.7
Mexico	26,866	4,181,354	117,736	155.6



(a) E41N27 (b) 4tiles



(c) 16tiles (d) Mexico

Fig. 11: Overview of the four datasets.

nar partition repair among the software tested vary considerably, with full topological repair only available in ArcGIS (using manual operations only), it is also important to consider how different repair implementations scale to large datasets. For this, a few performance tests were made in our implementation, and three planar partition repair tools that perform this process using snapping and splitting. The testing methodology for each tool is as follows:

ArcGIS In ArcCatalog, a multiple feature dataset is created in a Geodatabase, set with XY resolution values equal to the snapping threshold. The features are imported into it and the merge and dissolve operations are used to merge adjacent polygons with the same ID. Topology is then generated to check that the planar partition is valid. Everything is finally exported to a single Shapefile. The individual parts of the process are timed and the total is recorded. Memory usage is calculated as the difference between the just loaded ArcCatalog application and its maximum memory usage throughout the process.

FME A reader is created for each input file, which serves as input to a Snapper transformer; features with the same IDs are then dissolved, and finally they are output into a new Shapefile writer. The topology generator is used to be able to tell whether the result is a valid planar partition. Results are timed and the maximum memory allocation of the fme.exe process is recorded.

GRASS Input files are imported with v.in.ogr, with all polygon cleaning operations performed and snapping set to the correct values. Boundaries between features with the same IDs are then dissolved using v.dissolve. Files are then exported with v.out.ogr. Times reported by GRASS are added together to give the total, while memory usage by the v.in.ogr.exe, v.dissolve.exe and v.out.ogr.exe are monitored and their maximum is recorded.

pprepair Files are read and put into the triangulation, the triangulation is labelled, the repair is performed with the longest boundary first (see Tab. 1), with ambiguous cases resolved with a random choice. Polygons are then extracted from the triangulation and output to a single Shapefile. The entire process is timed, and the maximum amount of memory used is recorded.

Notice that we are somewhat comparing apples with oranges here since our implementation is able to repair more cases than other methods, keeps topological consistency, and does not require finding out the appropriate threshold value (if it exists). More importantly, it is able to directly state whether the result is a planar partition, unlike the three other solutions (with these snapping is performed but that does not guarantee that the output is a valid planar partition). The results of the experiments are shown in Tab. 3.

We have made the experiments in order to have an idea of the processing time and the memory usage involved. As Tab. 3 shows, our approach uses somewhat more memory than other solutions. This is explained by the extra (unconstrained) edges that are added to the input over when triangulating them. It should however be noticed that both ArcGIS and GRASS crashed with the biggest dataset 16tiles. Our implementation is the fastest of the four tested, being for instance around three times faster than FME for the biggest datasets. Only for the Mexico dataset is our implementation slower than FME. This is (probably) explained by the fact that its polygons already form a valid planar partition, and therefore very few snapping operations have been performed by FME; the planar graph of the input was basically built, and then the polygons saved back to disk. We believe that ArcGIS and GRASS struggled with the dataset because it contains several very large polygons (with more than 10,000 vertices). Our implementation took 3m31s, but, as Tab. 4 shows, it took CGAL 3m02s to simply triangulate the input edges.

This table also demonstrates the efficiency of our approach: for the four datasets, around 85 % of the time of our approach was used – by the CGAL library – to read the input from disk and triangulate the input edges.

During the implementation, we took several engineering decisions to optimise our pro-

Tab. 3: Planar partition repair comparison using large datasets.

	pprepair		ArcGIS		FME		GRASS	
	memory	time	memory	time	memory	time	memory	time
E41N27	145 MB	19s	145 MB	1m3s	158 MB	31s	59 MB	3m09s
4tiles	116 MB	17s	113 MB	37s	105 MB	31s	49 MB	53s
16tiles	1.45 GB	4m47s	crashes	–	636 MB	15m48s	crashes	–
Mexico	1.01 GB	3m31s	216 MB	>1d	264 MB	2m45s	408 MB	11m38s

Tab. 4: Timed steps of the planar partition repair procedure, rounded to the nearest second. The percentage is the triangulation time over the total time.

	E41N27	4tiles	16tiles	Mexico
Triangulate	0:17	0:15	4:00	3:02
Label	0:01	0:01	0:27	0:16
Repair	0:00	0:00	0:00	0:00
Reconstruct	0:01	0:01	0:11	0:07
Output	0:00	0:00	0:10	0:06
Total	0:19	0:17	4:47	3:31
triangulate %	89%	88%	83%	86%

gram. One of them was to favour disk space over computation time, as we wanted to be able to process large datasets. This is why we always reconstruct polygons, even if they were not modified by the repair process. Although possible, it would require us to keep in memory the original polygons (which would significantly increase the memory consumption) and to keep track of which labels have been modified. And, as Tab. 1 shows, the reconstruction is very efficient as it only takes around 3–4 % of the total time for the 16tiles and Mexico datasets.

5 Discussion and Conclusions

We have presented a new method for repairing polygonal area partitions, ensuring that the output partition is valid, i.e., all individual polygons are conforming to the ISO 19107 (ISO 2003) and *Simple Features* specification (OGC 2006) and neither gaps nor overlaps are present between any pair of polygons. The novelty of our method lies in the fact that repair is performed according to user defined criteria, but then takes place without any human intervention. Automatic repair is becoming increasingly an important topic due to data integration, e.g. data collected for the different themes of the European INSPIRE Initiative will finally have to fit together and data could eventually be matched fully automatically by dedicated web service components (INSPIRE 2009). Our approach could be at the heart of such a web service.

The proposed approach excels in automated repair at the cost of increased memory usage compared to a pure graph-based approach – this difference is mainly caused by the unconstrained edges introduced by the triangulation. While it would be possible to use our repair rules together with a (primal/dual) graph-based approach, these additional edges in the triangulation give fine-grained control over the repair operations, and ensures that the graph is connected, which facilitates the reconstruction of polygons.

We have implemented our algorithm over a numerically robust triangulator (CGAL) and since repair operations are expressed solely in terms of re-labelling of triangles (no geometric computation is involved), the approach is also fully robust. Since, during our experiments, most of the time was used to compute the constrained triangulation, another library could also be tested to improve the implementation.

For the future, we plan to:

- Improve the scalability of the approach and process datasets with more than 10 million polygons. It is known that using divide-and-conquer techniques triangulation algorithms can handle big datasets (Amenta et al. 2003, Blandford et al. 2005). We will investigate whether it is possible to automatically repair each divided part individually and 'glue' the repaired parts together.
- Investigate snap rounding (de Berg et al. 2007, Hobby 1999) as a pre-processing step – or embedded directly in the triangulation – to guarantee that repaired planar partitions have no vertices that are closer than a certain threshold. However, snap rounding may change the topology of the input, but the output will nevertheless be a valid partition as the topology will be repaired. Apart from topological changes, snap rounding can also lead to removed polygon parts that are too small to be preserved based on the chosen ε.
- Add more advanced repair operations to our repair toolkit, e.g. repair could take place based on splitting a collection of triangles.
- Extend our work to include the third dimension to validate and repair 3D city models using a constrained tetrahedralisation (Si 2008). Notice that the tetrahedralisation of a given polyhedron does not always exist, and thus extra (Steiner) points might need

to be added. The main concepts of our approach, (re)labelling and reconstruction, extend naturally to 3D. However, appropriate repair operations for 3D city models still need to be defined, and some application-specific constraints, e.g. right angles at corners are not trivial to implement in our current approach.

References

1Spatial, 2012: www.1spatial.com (24.5.2012).

Amenta, N., Choi, S. & Rote, G., 2003: Incremental constructions con BRIO. – 19th Annual Symposium on Computational Geometry: 211–219, ACM Press.

ArcGIS, 2012: www.esri.com/software/arcgis/index.html (24.5.2012).

Badawy, W.M. & Aref, W.G., 1999: On local heuristics to speed up polygon-polygon intersection tests. – 7th ACM International Symposium on Advances in Geographic Information Systems: 97–102, ACM Press.

Bader, M. & Weibel, R., 1997: Detecting and resolving size and proximity conflicts in the generalization of polygonal maps. – 18th International Cartographic Conference: 1525–1532, Stockholm.

Blandford, D.K., Blelloch, G.E., Cardoze, D.E. & Kadow, C., 2005: Compact representations of simplicial meshes in two and three dimensions. – International Journal of Computational Geometry and Applications **15** (1): 3–24.

CGAL, 2012: Computational Geometry Algorithms Library: www.cgal.org/ (24.5.2012).

Clarkson, K.L., Mehlhorn, K. & Seidel, R. 1992: Four results on randomized incremental constructions. – Finkel, A. & Jantzen, M. (eds.): 9th Annual Symposium on Theoretical Aspects of Computer Science **577:** 461–474, Lecture Notes in Computer Science, Springer.

Corine, 2000: www.eea.europa.eu/data-and-maps/data/corine-land-cover-2000-clc2000-seamless-vector-database (24.5.2012).

de Berg, M., Halperin, D. & Overmars, M. 2007: An intersection-sensitive algorithm for snap rounding. – Computational Geometry **36** (3): 159–165.

de Berg, M., van Kreveld, M., Overmars, M. & Schwarzkopf, O., 2008: Computational Geometry: Algorithms and Applications. – 3rd ed., Springer.

Esri, 1998: Shapefile technical description. – White paper, Esri.

Facello, M.A., 1995: Implementation of a randomized algorithm for Delaunay and regular triangulations in three dimensions. – Computer Aided Geometric Design **12** (4): 349–370.

FME, 2012: www.safe.com (24.5.2012).

GDAL, 2012: gdal.org/org (11.8.2012)

GEOS, 2012: Geometry Engine Open Source. – http://trac.osgeo.org/geos/ (24.5.2012).

GRASS, 2012: http://grass.osgeo.org (24.5.2012).

GTS, 2006: GTS Library Reference Manual, gts.sourceforge.net/reference/book1.html (24.5.2012).

Guibas, L. & Stolfi, J., 1985: Primitives for the manipulation of general subdivisions and the computation of Voronoi. – ACM Transactions on Graphics **4** (2): 74–123.

Hobby, J.D., 1999: Practical segment intersection with finite precision output. – Computational Geometry **13** (4): 199–214.

Hoffmann, C.M., Hopcroft, J.E. & Karasick, M.S., 1988: Towards implementing robust geometric computations. – **4th** Annual Symposium on Computational Geometry: 106–117.

INSPIRE, 2009: D2.5: Generic conceptual model, version 3.2. – Technical report, Drafting Team "Data Specifications", http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/D2.5_v3.2.pdf, (24.5.2012).

ISO, 2003: ISO 19107, Geographic information – Spatial schema.

Kirkpatrick, D., Snoeyink, J. & Speckmann, B., 2002: Kinetic collision detection for simple polygons. – International Journal of Computational Geometry and Applications **12** (1–2): 3–27.

Laurini, R. & Milleret-Raffort, F., 1994: Topological reorganization of inconsistent geographical databases: A step towards their certification. – Computers & Graphics **18** (6): 803–813.

Ledoux, H. & Meijers, M., 2010: Validation of planar partitions using constrained triangulations. – Joint International Conference on Theory, Data Handling and Modelling in GeoSpatial Information Science: 51–55, Hong Kong.

Ledoux, H., Ohori, K.A. & Meijers, M., 2012: Automatically repairing invalid polygons with a constrained triangulation. – **15th** AGILE International Conference on Geographic Information Science.

Liu, Y. & Snoeyink, J., 2006: Faraway point: A sentinel point for Delaunay computation. – International Journal of Computational Geometry and Applications **18** (4): 343–355.

Margalit, A. & Knott, G.D., 1989: An algorithm for computing the union, intersection or difference of two polygons, Computers & Graphics **13** (2): 167–183.

Mücke, E.P., Saias, I. & Zhu, B., 1999: Fast randomized point location without preprocessing in two- and three-dimensional Delaunay triangulations. – Computational Geometry. – Theory and Applications **12:** 63–83.

OGC, 2006: OpenGIS Implementation Specification for Geographic Information – Simple feature access – Part 1: Common architecture, 1.2.0 ed.

Plümer, L. & Gröger, G., 1997: Achieving integrity in geographic information systems – maps and nested maps. – GeoInformatica **1** (4): 345–367.

PostGIS, 2012: http://postgis.refractions.net/ (24.5.2012).

Rivero, M. & Feito, F.R., 2000: Boolean operations on general planar polygons. – Computers & Graphics **24** (6): 881–896.

Schirra, S., 1997: Precision and Robustness in Geometric Computations. – Algorithmic Foundations of Geographic Information **9:** 255–287, Lecture Notes in Computer Science, Springer.

Shamos, M.I. & Hoey, D., 1976: Geometric intersection problems. – IEEE, FOCS: 208–215.

Shewchuk, J.R., 1997a: Delaunay Refinement Mesh Generation. – PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburg.

Shewchuk, J.R., 1997b: Adaptive precision floating-point arithmetic and fast robust geometric predicates. – Discrete & Computational Geometry **18:** 305–363.

Si, H., 2008: Three Dimensional Boundary Conforming Delaunay Mesh Generation. – PhD thesis, Technische Universität Berlin.

Tomlin, C.D., 1994: Map algebra: One perspective. – Landscape and Urban Planning **30** (1–2): 3–12.

van Oosterom, P., Quak, W. & Tijssen, T., 2004: About invalid, valid and clean polygons. – Fisher, P.F. (ed.): Developments in Spatial Data Handling. – **11th** International Symposium on Spatial Data Handling: 1–16.

van Roessel, J.W., 1991: A new approach to planesweep overlay: Topological structuring and linesegment classification. – Cartography and Geographic Information Science **18** (1): 49–67.

Vivid Solutions, 2012: Java Topology Suite. – http://www.vividsolutions.com/jts/jtshome.htm (24.5.2012).

Address of the Authors:

Ken Arroyo Ohori, Dr. Hugo Ledoux & Dr. Martijn Meijers, Section GIS Technology, OTB Research Institute for the Built Environment, Delft University of Technology, Jaffalaan 9, 2628 BX Delft, The Netherlands, e-mail: {g.a.k.arroyoohori}{h.ledoux}{b.m.meijers}@tudelft.nl

Manuskript eingereicht: November 2011
Angenommen: April 2012

PFG 2012 / 5, 0631 – 0646
Stuttgart, October 2012

Article

Automatisierte Verkehrszeichenkartierung aus mobil erfassten Stereobilddaten unter Verwendung der Tiefeninformation aus Dense-Stereo-Matching

Stefan Cavegn & Stephan Nebiker, Muttenz, Schweiz

Keywords: mobile mapping, road signs, depth maps, dense stereo matching

Summary: *Automated 3D Road Sign Mapping with Stereovision-based Mobile Mapping exploiting Depth Information from Dense Stereo Matching.* This paper presents algorithms and investigations on the automated detection, classification and mapping of road signs which systematically exploit depth information from stereo images. This approach was chosen due to recent progress in the development of stereo matching algorithms enabling the generation of accurate and dense depth maps. In comparison to mono imagery-based approaches, depth maps also allow 3D mapping of the objects. This is essential for efficient inventory and for future change detection purposes. Test measurements with the mobile mapping system by the Institute of Geomatics Engineering of the University of Applied Sciences and Arts Northwestern Switzerland demonstrated that the developed algorithms for the automated 3D road sign mapping perform well, even under difficult to poor lighting conditions. Approximately 90 % of the relevant road signs with predominantly red, blue and yellow colours in the standard and small format in Switzerland can be detected, and 85 % can be classified correctly. Furthermore, fully automated mapping with a 3D accuracy of better than 10 cm is possible.

Zusammenfassung: In diesem Beitrag werden Algorithmen und Untersuchungen zur automatischen Detektion, Klassifizierung und Kartierung von Verkehrszeichen vorgestellt, die systematisch die Tiefeninformation aus Stereobildaufnahmen ausnutzen. Dieser Ansatz bietet sich an, da in den letzten Jahren neue Stereo-Algorithmen entwickelt wurden, welche die Generierung von dichten Tiefenkarten aus Stereobilddaten ermöglichen. Im Vergleich zu den auf Monobildern basierenden Verfahren erlauben die Tiefenkarten auch eine genaue 3D-Kartierung der Objekte, was für die effiziente Inventarisierung und für eine zukünftige Veränderungsanalyse wesentlich ist. Testmessungen mit dem mobilen Messfahrzeug des Instituts Vermessung und Geoinformation der Fachhochschule Nordwestschweiz haben gezeigt, dass die entwickelten Algorithmen für die automatisierte Verkehrszeichenkartierung auch bei schlechten Beleuchtungsverhältnissen sehr leistungsfähig sind. Es können etwa 90 % der relevanten Verkehrszeichen mit überwiegend roten, blauen und gelben Farbanteilen im Normal- und Kleinformat in der Schweiz detektiert und 85 % korrekt klassifiziert werden. Zudem ist eine vollautomatische Kartierung mit einer 3D-Genauigkeit von unter 10 cm möglich.

1 Einleitung

Entlang von Straßen in westeuropäischen Ländern sind sehr viele Verkehrszeichen angebracht. In der Schweiz sind es zurzeit etwa 5 Millionen. Vielfach fehlen den Eigentümern digitale Informationen über Position, Zustand sowie Zweckmäßigkeit dieser Verkehrszeichen. Nach Meinung verschiedener Experten sind beispielsweise viele Verkehrszeichen überflüssig. So wird geschätzt, dass die unnötige Beschilderung in der Schweiz den Bund, die Kantone und die Gemeinden jährlich mehrere Millionen Franken kostet (Schweizer Fernsehen 2008).

Zur Analyse und Lösung dieser Probleme und Fragestellungen bietet sich ein digitales Verkehrszeicheninventar an. Um ein solches Inventar anzulegen, wurden die Sachdaten und dokumentarische Bildinformation

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0144 1432-8364/12/0144 $ 3.75

der Verkehrszeichen in der Regel aufwändig vor Ort manuell erfasst und die Position mit einem GNSS-Empfänger im Meter- bis Dezimeterbereich bestimmt. In den letzten Jahren erfolgte die Kartierung und Inventarisierung aber auch zunehmend auf Grund von Daten, welche mit mobilen Messfahrzeugen erfasst wurden. Der Vorteil liegt darin, dass die 3D-Objekte sehr effizient kartiert werden können, ohne den Verkehrsfluss zu stören und das Vermessungspersonal zu gefährden. In Belgien konnten Verkehrszeichen über das ganze Landesgebiet anhand von Laserscanning-Daten detektiert werden; die Ableitung der Sachdaten erfolgte aber größtenteils manuell (Trimble 2009). In den Niederlanden wurde die Verkehrszeichenkartierung auf der Basis von Panoramabildern, welche im Abstand von 5 m vorliegen, manuell vorgenommen (de With et al. 2010). Können Verkehrszeichen aus georeferenzierten Bilddaten größtenteils automatisch extrahiert werden, reduziert sich der manuelle Kartierungsaufwand massiv. Algorithmen für die Verkehrszeichenextraktion aus Monobildern sind in Maldonado-Bascón et al. (2008) beschrieben. In diesem Beitrag werden Algorithmen präsentiert, die es erlauben, Verkehrszeichen aus mobil erfassten Stereobildsequenzen automatisch und benutzerunterstützt zu extrahieren, die 3D-Position im gewünschten geodätischen Referenzrahmen zu bestimmen und diverse Sachdaten aus den

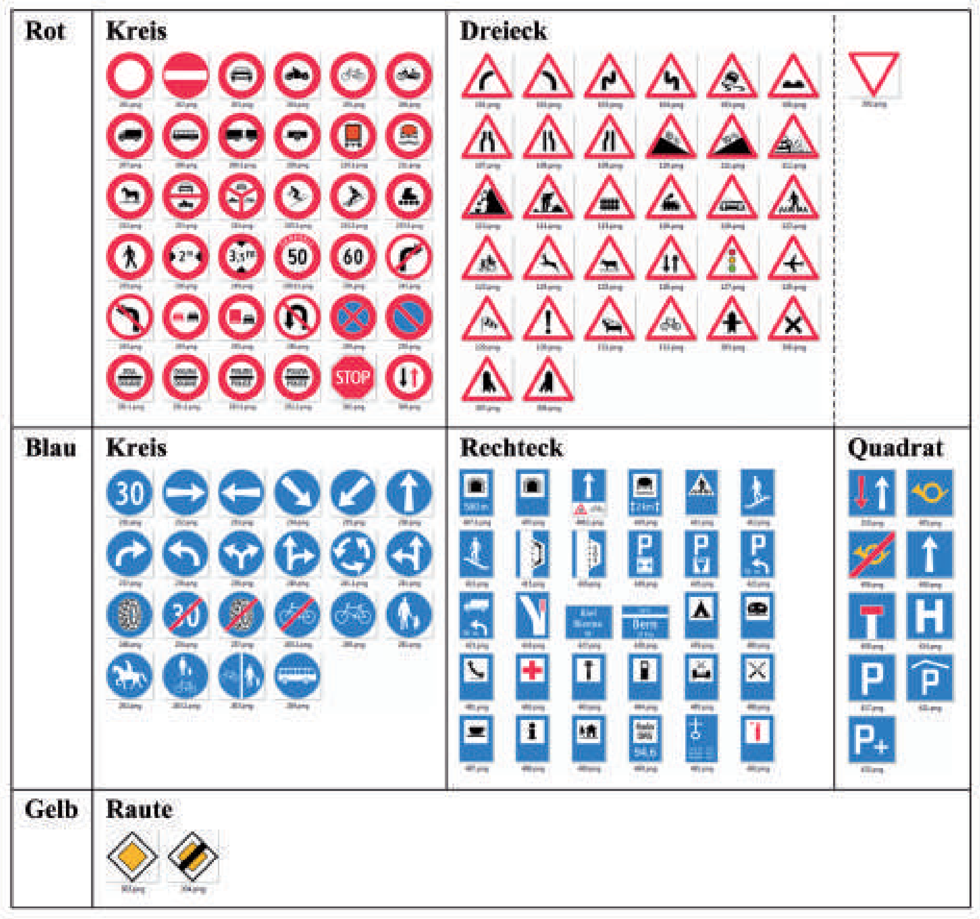


Abb. 1: Verkehrszeichen in der Schweiz, welche mit den entwickelten Algorithmen automatisch detektiert, klassifiziert und kartiert werden können. Gliederung nach dominanter Farbe (Zeilen) und nach Form (Spalten).

Bilddaten automatisch abzuleiten. Diese Algorithmen wurden vorerst auf die Verkehrszeichen in der Schweiz optimiert, welche unterschiedliche Farben und Formen aufweisen sowie je nach Straßentyp in vier verschiedenen Normgrößen vorkommen können (siehe Abb. 1). Sie lassen sich aber auch auf andere Länder übertragen. Da die Anwendung nicht für Fahrerassistenzsysteme oder intelligente autonome Fahrzeuge ausgelegt ist, sondern das Ziel explizit in der Kartierung, Inventarisierung und Bewirtschaftung von Verkehrszeichen liegt, hat die Echtzeitfähigkeit der Algorithmen nicht erste Priorität. Stattdessen liegt der Fokus auf Vollständigkeit, Korrektheit und geometrischer Genauigkeit.

Im folgenden Kapitel werden bestehende Ansätze für die Detektion, Klassifizierung und Kartierung von Verkehrszeichen erläutert. Nach der Vorstellung des Lösungsansatzes in Abschnitt 3 erfolgt die Präsentation der entwickelten Algorithmen und Softwaremodule in Abschnitt 4. Die Untersuchungen und Resultate sind in Abschnitt 5 beschrieben. Im letzten Abschnitt wird nach einem Fazit ein Ausblick auf laufende und zukünftige Entwicklungen gegeben.

2 Verwandte Arbeiten

Die meisten Ansätze zur Thematik der bildbasierten Verkehrszeichenextraktion lassen sich in zwei Hauptschritte unterteilen. Zuerst erfolgt die Detektion der Verkehrszeichen in der Bildszene mit dem Ziel, mögliche Kandidaten zu lokalisieren. Um für jeden Kandidaten die korrekte Verkehrszeichenart zu bestimmen, ist danach eine Klassifizierung erforderlich. Einige der maßgebenden existierenden Ansätze für die Detektion und Klassifizierung der Verkehrszeichen sind in diesem Abschnitt dokumentiert. Eine detaillierte Zusammenstellung von unterschiedlichen Methoden mit ihren Merkmalen ist in Nguwi & Kouzani (2008) gegeben. Ist zusätzlich die absolute Position der detektierten Zeichen von Interesse, fällt als dritter Schritt noch die Kartierung der Verkehrszeichen an.

2.1 *Detektion von Verkehrszeichen*

Das Ziel der Detektion besteht darin, in einem Kamerabild Verkehrszeichenkandidaten zu finden. Ausgangsbasis ist dabei vielfach die Farbinformation. Wird eine Farbsegmentierung mit Schwellwerten vorgenommen, kann schnell auf die Suchregionen fokussiert werden. Da der RGB-Farbraum empfindlich auf Änderungen der Lichtverhältnisse durch Schatten, Beleuchtungs- und Betrachtungsgeometrie sowie auf starke Reflektionen reagiert, erfolgt die Segmentierung in der Regel im HSV-Farbraum mit den Farbton- und Sättigungs-Komponenten (Fleyeh 2006, Maldonado-Bascón et al. 2008, Fu & Huang 2010). Jau et al. (2008) haben umfangreiche Untersuchungen zu den Unterschieden zwischen dem RGB- und HSV-Farbraum vorgenommen und festgestellt, dass für die Verkehrszeichendetektion die Farbsegmentierung im HSV-Farbraum besser als im RGB-Raum geeignet ist. Madeira et al. (2005) verwenden für die Farbsegmentierung die Farbton- und die chromatische RGB-Komponente, welche nach de la Escalera et al. (2004) berechnet wird. Dies weil die Sättigungs-Komponente im Gegensatz zur chromatischen RGB-Komponente für kleine Werte sehr empfindlich auf Rauschen reagiert (Madeira et al. 2005). Nach García-Garrido et al. (2006) treten Segmentierungsprobleme bei vorwiegend weißen Verkehrszeichen und bei Beleuchtungsunterschieden auf. Formbasierte Methoden zeigen eine größere Robustheit bezüglich variabler Beleuchtungsverhältnisse. Obwohl diese Methoden schnell in der Ausführung und ziemlich einfach zu implementieren sind, produzieren sie viele Kandidatenobjekte, was für die folgenden Schritte mehr Aufwand bedeutet (Gómez-Moreno et al. 2010).

2.2 *Klassifizierung von Verkehrszeichen*

Bei der Klassifizierung geht es darum, für den beim Detektionsprozess ermittelten Verkehrszeichenkandidaten den richtigen Typ zu bestimmen. Dabei kommen in vielen Fällen neuronale Netzwerke zum Einsatz (de la Escalera et al. 2003, Nguwi & Kouzani 2008, Hossain et al. 2010). Diese erfordern zusätzlich zur Implementierung ein umfassendes Training der Algorithmen, das Eingangsbild muss aber nicht in eine andere Raumrepräsentation transformiert werden (Fleyeh & Dougherty 2005). Ren et al. (2009) zeigen, dass der Trainingsprozess im Fall einer umfassenden Verkehrszeichenklassifizierung sehr viele Trainingsbilder mit unterschiedlicher Skalierung, Orientierung und Beleuchtung benötigt. Neuronale Netzwerke werden daher vielfach nur für wenige bestimmte Verkehrszeichen wie Geschwindigkeitsbegrenzungen verwendet. Gao et al. (2006) schreiben, dass die neuronalen Netzwerke neu trainiert werden müssen, wenn ein weiteres Referenzmuster hinzugefügt wird, was viel Zeit in Anspruch nimmt. Beim Template Matching handelt es sich um ein intensitätsbasiertes Bildzuordnungsverfahren zwischen Referenz- und Suchbild. Für die Klassifizierung wird es unter anderem von Piccioli et al. (1996) und Malik et al. (2007) eingesetzt. Ein neues Referenzmuster kann problemlos hinzugefügt werden. Das Template Matching in seiner Grundform weist aber keine Robustheit auf Skalierung, Rotation oder allgemein auf affine Transformationen auf und reagiert empfindlich auf Beleuchtungsunterschiede (Ren et al. 2009).

2.3 *Weitere Ansätze für die Detektion und Klassifizierung von Verkehrszeichen*

Viele Ansätze lassen sich nicht nur für die Detektion oder Klassifizierung von Verkehrszeichen einsetzen, sondern sind für beide Aufgaben geeignet. Einige davon werden nachfolgend diskutiert.

Die Hough-Transformation reagiert gemäß Nguwi & Kouzani (2008) tolerant auf Lücken und ist nicht sehr anfällig auf Rauschen. Durch die unterschiedlichen Normgrößen und Formen der Verkehrszeichen, sowie durch die teilweise ungleiche Ausdehnung in den Koordinatenrichtungen, müssen viele Maßstäbe geprüft werden, was sich negativ auf die Rechenzeit und den Speicherbedarf auswirkt. Somit können für Echtzeitanwendungen nur schnellere abgeänderte Methoden eingesetzt werden (Nguwi & Kouzani 2008). Auch Chutatape & Guo (1999) zählen die lange Rechenzeit und den hohen Speicherbedarf der standardisierten Hough-Transformation zu den Nachteilen. Um dies zu verbessern, wird eine modifizierte Variante der Hough-Transformation vorgeschlagen (Chutatape & Guo 1999). Diese Variante verwenden Kim et al. (2006) für die Verkehrszeichendetektion, nachdem vorgängig Kanten mit dem Canny-Operator aus den Bilddaten extrahiert wurden. Barrile et al. (2007) führen die Formdetektion mit der standardisierten Hough-Transformation durch; für die Klassifizierung verwenden sie die generalisierte Hough-Transformation. Diese wird auch von Habib et al. (1999) auf Kanten angewandt, welche mit dem Canny-Filter extrahiert wurden.

Für die Detektion und Klassifizierung von Verkehrszeichen werden zunehmend die Ansätze von Support Vector Machines (SVM) und Scale Invariant Feature Transform (SIFT) verfolgt. Beim SIFT-Ansatz nach Lowe (2004) sind die extrahierten Merkmale invariant gegen Translationen, Rotationen und Skalierung und robust gegen Beleuchtungsänderungen, Bildrauschen und kleine geometrische Deformationen (Reiterer et al. 2009, Ren et al. 2009). Fleyeh & Dougherty (2008) stellen eine Echtzeit-Anwendung für schwedische Verkehrszeichen vor. Mehr als 3400 Bilder kamen für das Training und diverse Tests zum Einsatz. Dabei wurden zwei Stufen von SVM trainiert. Die erste Stufe dient der Formdetektion und die zweite der Erkennung der Piktogramme (Fleyeh & Dougherty 2008). Maldonado-Bascón et al. (2007) verwenden für die Detektion und anschliessende Formklassifizierung eine lineare SVM und für die Erkennung des Inhalts eine SVM mit Gauss-Kernel. Ein Verkehrszeichen wird nur dann korrekt detektiert, wenn es in mehr als einem Bild vorkommt. Somit ergeben sich wenige ‘False Positives’. Die implementierten Algorithmen sind invariant bei Translation, Rotation, Skalierung und in vielen Fällen bei partiellen Verdeckungen (Maldonado-Bascón et al. 2007).

2.4 *Extraktion unterschiedlicher Verkehrszeichen*

In der Regel wird entweder auf die Detektion oder die Klassifizierung von Verkehrszeichen fokussiert; praxistaugliche Gesamtlösungen mit zusätzlicher automatischer Kartierung sind kaum implementiert. In mehreren Fällen wird auch nur eine semantische Kategorie wie Geschwindigkeitszeichen behandelt. Im Gegensatz dazu haben Ren et al. (2009) eine Echtzeit-Anwendung für fast alle Verkehrszeichen in Neuseeland implementiert. Die Detektion wird mit einer Farbsegmentierung im HSV-Farbraum vorgenommen. Es folgt eine Hough-Transformation und ein anschliessendes Matching mit SIFT- oder SURF-Merkmalen. Mit den in Maldonado-Bascón et al. (2008) vorgestellten Algorithmen lassen sich mehr als 300 verschiedene Verkehrszeichen in Spanien erkennen. Die Anwendung funktioniert sowohl für urbane als auch für ländliche Umgebungen. Anpassungen für Verkehrszeichen in anderen Ländern können ohne weiteres vorgenommen werden (Maldonado-Bascón et al. 2008). Ein System, welches Verkehrszeichen verschiedener Länder detektieren und klassifizieren kann, wurde von Khan et al. (2011) entwickelt. Sie haben diverse Farben und Formen berücksichtigt, und die Algorithmen sind invariant bei Translation, Rotation, Skalierung und partiellen Verdeckungen.

2.5 *Kartierung von Verkehrszeichen*

Um Verkehrszeichen zu inventarisieren, wurden bis vor kurzem die Koordinaten mit einem handgeführten GNSS-Empfänger gemessen, das Zeichen mit einer Digitalkamera fotografiert und die Sachdaten mit Hilfe eines GIS-Clients vor Ort erfasst. Anhand eines mobilen PDA-basierten Geoinformationssystems können all diese Arbeiten mit nur einem Gerät durchgeführt werden. Ein solches wurde von Benesova et al. (2007) für die Inventarisierung von Verkehrszeichen eingesetzt, wobei die Klassifizierung halbautomatisch erfolgte.

Die bisherige Forschung und Entwicklung im Bereich der Detektion und Klassifizierung von Verkehrszeichen war größtenteils getrieben durch die Bedürfnisse der Automobilindustrie (vgl. Abschnitte 2.1 bis 2.4). Dabei sind die Aspekte Fahrerassistenz und Sicherheit zentral, womit in der Regel auf rote Gefahren-, Vorschrifts- oder Vortrittssignale fokussiert wurde. Die automatisierte exakte Lokalisierung bzw. Kartierung von sämtlichen Verkehrszeichen spielte demgegenüber eine deutlich geringere Rolle, was sich in einer kleineren Anzahl entsprechender Forschungsarbeiten niederschlägt. Zu den besonders relevanten Arbeiten im Bereich der automatisierten Kartierung von Verkehrszeichen gehören sowohl Madeira et al. (2005), Kim et al. (2006) und Baró et al. (2009), welche die absolute 3D-Koordinatenbestimmung von Objektpunkten aufgrund von Stereobilddaten vornehmen, als auch Shi et al. (2008), die einen kombinierten Ansatz von Bild- und Laserscanning-Daten verwenden. Während Shi et al. (2008) eine Genauigkeit von circa 30 cm erreichen, können Madeira et al. (2005) Kartierungen nur im Meterbereich vornehmen. Die exakte 3D-Kartierungsmöglichkeit in einem globalen geodätischen Bezugssystem ist wesentlich und gewinnt zunehmend an Bedeutung, da dadurch digitale Inventare für eine Reihe allgemeiner GIS-basierter Aufgabenstellungen, z.B. Verkehrsplanung, Change Detection, Simulationen, Überlagerung und Nachführung von existierenden oder projektierten Geodaten in Mixed-Reality Umgebungen, angelegt werden können.

3 Lösungsansatz auf Basis der Stereodisparität

Beim entwickelten und nachfolgend vorgestellten Ansatz bildet die Nutzung der Tiefenkarten aus Stereobildaufnahmen ein zentrales Element. Obwohl die Tiefeninformation aus Dense-Stereo-Matching ein riesiges Potential hat, wurde sie für die Extraktion von Verkehrszeichen noch gar nicht oder wie von Cyganek (2008) nur als optionaler Bestandteil für die Suchraumeinschränkung (Unterteilung in nahe Regionen und Hintergrund) verwendet. Anhand der Stereobilder können die Verkehrszeichen auch dreidimensional kartiert werden, was mit Monobildern nur sehr umständlich und mit beschränkter Genauigkeit möglich ist.

Für die Datenerfassung kann ein Stereobild-basiertes Mobile Mapping System wie dasjenige des Instituts Vermessung und Geoinformation (IVGI) der Fachhochschule Nordwestschweiz (FHNW) eingesetzt werden (siehe Abb. 2, links). Bei diesem Messsystem stehen für die Bildaufnahme zurzeit zwei Stereosysteme mit einer Stereobasis von je circa 90 cm und Industriekameras unterschiedlicher Auflösung (Full-HD und 11MPixel) zur Verfügung. Zur direkten Georeferenzierung der Stereobildpaare dient ein Navigationssystem POS LV 210 von Applanix, bestehend aus einem GNSS-Empfänger, einer Inertialmesseinheit und einem Distanzmessindikator. Diverse Testfahrten mit dieser multifunktionalen Forschungsplattform haben ergeben, dass für gut signalisierte Punkte bei einer ausreichenden GNSS-Konstellation 3D-Genauigkeiten von 2–3 cm möglich sind.

Um vereinfachte Verfahren für die Folgeoperationen zu ermöglichen, sind die mit dem mobilen Messfahrzeug erfassten und genähert parallelen Stereobilder mittels Rektifizierung in den Stereonormalfall umzurechnen, nachdem die Verzeichnungskorrektur angebracht wurde. Auf Grund der resultierenden Normalbilder wird mit einem Stereo-Matching-Algorithmus für jedes Pixel die Disparität bzw. x-Parallaxe ermittelt. Für jede Disparität kann danach anhand der Stereogeometrie (Basis und Kamerakonstante) ein Tiefenwert berechnet werden, deren Gesamtheit eine Tiefenkarte bildet. Für die vorliegenden Untersuchungen erfolgt das Dense Matching mit dem in OpenCV implementierten Semi-Global-Block-Matching-Algorithmus (OpenCV 2012), der einige Unterschiede zu dem von Hirschmüller (2008) entwickelten Semi-Global-Matching-Algorithmus aufweist (z.B. Berechnung der Matchingkosten). Die semiglobale Zuordnung basiert auf einer global zu optimierende Kostenfunktion, welche durch Aufsummierung der Kosten auf ausgewählten Pfaden approximiert wird. Dank dieser Approximation kann dieser semi-globale Zuordnungsalgorithmus für große Bilder, insbesondere auch für hochaufgelöste Luftbilder, eingesetzt werden. Ein Vergleich des SGM-Algorithmus mit einer Vielzahl alternativer Stereo-Algorithmen ist auf der Middlebury Stereo Webseite (Scharstein & Szeliski 2012) gegeben.

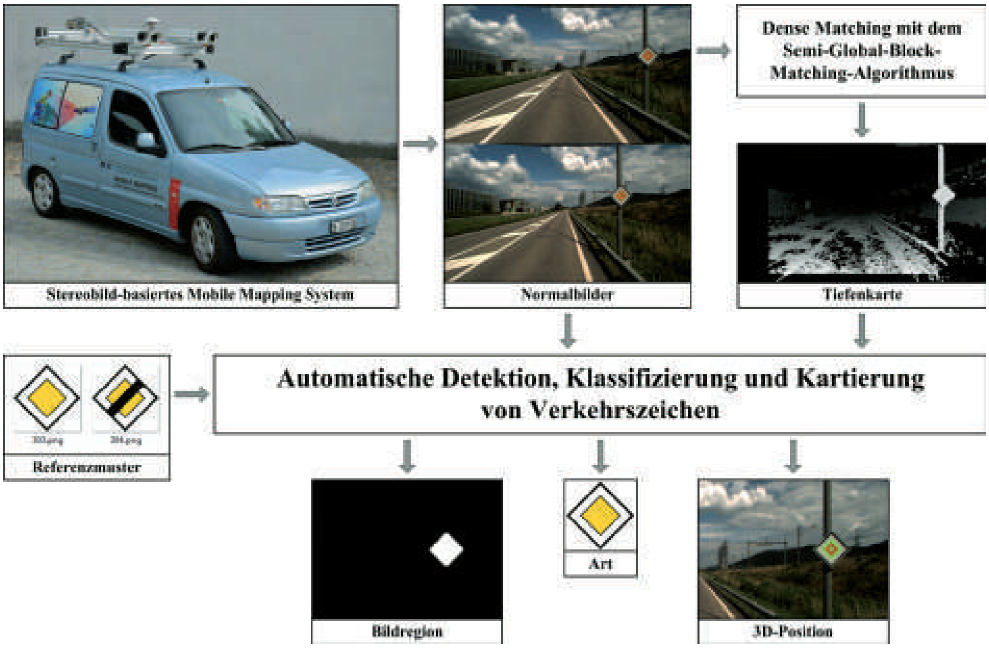


Abb. 2: Eingangs- und Ausgabedaten für die automatische Detektion, Klassifizierung und 3D-Kartierung von Verkehrszeichen.

Sowohl Normalbilder als auch Tiefenkarten sind für die anschliessende automatische Detektion und Kartierung von Verkehrszeichen notwendig (siehe Abb. 2). Für die Klassifizierung werden zusätzlich Referenzmuster von allen möglichen Verkehrszeichen benötigt. Diese können im Falle der Schweiz beispielsweise von der Webseite Wikipedia (2012) im PNG-Format heruntergeladen werden. Nach erfolgreicher Detektion, Klassifizierung und Kartierung sind die Bildregion, die Sachdaten sowie die 3D-Position des Verkehrszeichens bekannt.

Bei den entwickelten Algorithmen werden die Stereodisparitäten bzw. die daraus abgeleiteten Tiefenkarten für unterschiedliche Aufgaben verwendet:

- Einschränkung des Suchraums durch den vordefinierten Distanzbereich.
- Definition von distanzabhängigen Bedingungen für die Farbsegmente.
- Bestimmung von Regionen mit ähnlichen Tiefenwerten.
- Berechnung der 3D-Position.
- Einschränkung des Wertebereichs für die Maßstäbe bei der Hough-Transformation.

4 Entwickelte Algorithmen und Softwaremodule

Der vorgestellte Lösungsansatz auf der Basis von Stereobilddaten und Tiefenkarten wurde mit einer Reihe von Algorithmen und Softwaremodulen in Matlab umgesetzt. Diese decken den gesamten Ablauf von der automatischen Detektion über die Klassifizierung bis zur Kartierung von Verkehrszeichen ab (siehe Abb. 3) und werden nachfolgend erläutert.

4.1 *Automatische Detektion von Verkehrszeichen*

Als Eingangsdaten sind von jedem Stereobildpaar das linke Normalbild und die entsprechende Tiefenkarte erforderlich (siehe Abb. 4). Im unteren Drittel des Normalbildes sind keine permanenten Verkehrszeichen zu erwarten, weshalb dieser mit einem schwarzen Band ausmaskiert wird. Da die Farbton- und Sättigungs-Komponenten kaum auf unterschiedliche Beleuchtungsverhältnisse reagieren und mit diesen somit bessere Ergebnisse bei der Detektion von Verkehrszeichen erzielt werden können, wird das RGB-Normalbild in den HSV-Farbraum transformiert. Um den Suchraum im Normalbild einzuschränken, wird die Tiefenkarte auf einen vordefinierten Distanzbereich reduziert. Das gewählte Basis-Tiefenverhältnis liegt zwischen 0.06 und 0.25, womit auch Verkehrszeichen detektiert werden können, die an einer benachbarten Fahrbahn anliegen. Zudem kann mit einer hohen Bildaufnahmerate das gleiche Zeichen redundant detektiert und klassifiziert werden. Die Segmentierung von roten, blauen und gelben Farbregionen erfolgt anhand von Schwellwerten für die Farbton- und Sättigungs-Komponenten, welche empirisch auf Grund von mehreren Bildern verschiedener Messfahrten ermittelt wurden. Für blaue Segmente ergab sich beispielsweise ein Farbtonbereich zwischen 0.52 und 0.72 sowie ein Sättigungsbereich zwischen 0.20 und 0.80. Liegt der Farbtonwert zwischen 0.04 und 0.19 und ist der Sättigungs-

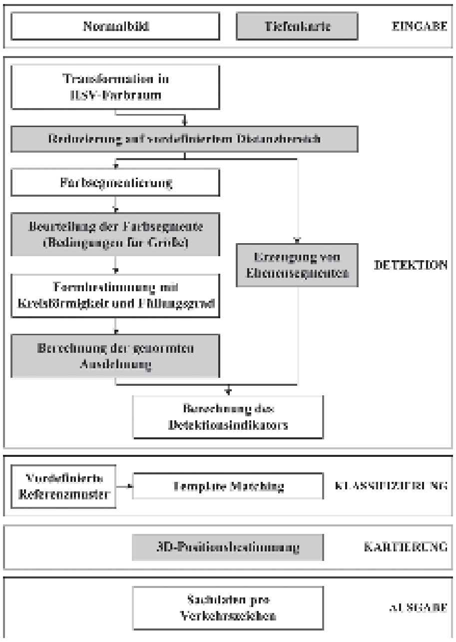


Abb. 3: Entwickelte Algorithmen für die automatische Detektion, Klassifizierung und 3D-Kartierung von chromatischen Verkehrszeichen (graue Felder: Operationen nutzen die Disparitäts- bzw. Tiefeninformation).

wert größer als 0.50 und kleiner als 0.98, handelt es sich um Pixel eines gelben Segments. Nach den morphologischen Bildoperationen Öffnen, Schließen und Füllen wird geprüft, ob die Farbsegmente einer bestimmten Flächengröße, welche von der jeweiligen Aufnahmedistanz des Segments abhängig ist, entsprechen. Ist dies der Fall, erfolgt die Bestimmung der Form mit den zwei Merkmalen Kreisförmigkeit und Füllungsgrad. Diese lassen sich wie folgt berechnen:

$$\text{Kreisförmigkeit} = \frac{4 \times \pi \times \text{Fläche des Segments}}{\left(\text{Umfang des Segments}\right)^2} \qquad (1)$$

$$\text{Füllungsgrad} = \frac{\text{Fläche des Segments}}{\text{Fläche des umschliessenden Rechtecks}} \qquad (2)$$

Die Differenzen zwischen der Ausdehnung der Segmente und den genormten Seitenlängen, welche in der Signalisationsverordnung (Schweizer Bundesrat 1979) definiert sind, dürfen ein Toleranzintervall von beispielsweise 10 cm nicht überschreiten. Die Tiefenkarten dienen dabei der Bestimmung der metrischen Breite und Höhe des Segments im Objektraum. Die gerundeten Seitenlängen sollten dem Klein- oder Normalformat in der Schweiz entsprechen. Nachdem auf der distanzreduzierten Tiefenkarte ein Schliessen und eine Dilatation ausgeführt wurden, erfolgt die Generierung von Ebenensegmenten. Dies sind Regionen, in denen die Bildpunkte sehr geringe Tiefenunterschiede aufweisen. Die Segmente werden wiederum mit morphologischen Bildoperationen (Öffnen, Schließen und Dilatation) generalisiert. Der Quotient aus Fläche des Ebenensegments innerhalb des Farbsegments (Verschnitt Abb. 4g und Abb. 4k) und Gesamtfläche des Farbsegments (Abb. 4g) bildet den Detektionsindikator. Dieser dient zur Beurteilung der Güte des Detektionsprozesses.

4.2 *Automatische Klassifizierung und Kartierung von Verkehrszeichen*

Für die Klassifizierung der Verkehrszeichen wird ein kreuzkorrelationsbasiertes Template Matching mit vordefinierten Referenzmustern

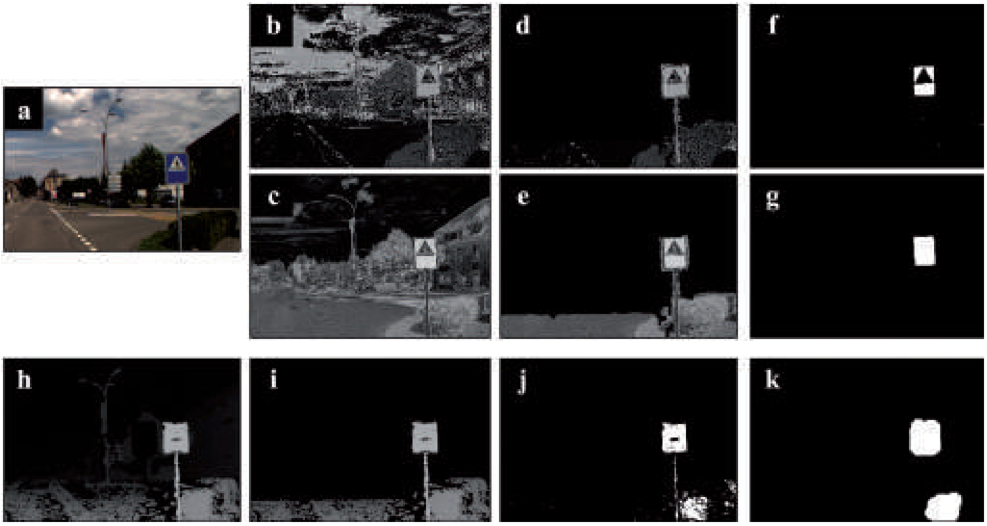


Abb. 4: Automatische Detektion eines rechteckigen Verkehrszeichens mit vorwiegend blauen Farbanteilen (a: linkes Normalbild, b: Farbton-Komponente, c: Sättigungs-Komponente, d: distanzreduzierte Farbton-Komponente, e: distanzreduzierte Sättigungs-Komponente, f: blaue Farbsegmente (binarisiert), g: blaue Farbsegmente nach morphologischen Operationen, h: Tiefenkarte, i: distanzreduzierte Tiefenkarte, j: Ebenensegmente, k: Ebenensegmente nach morphologischen Operationen).

durchgeführt (siehe Abb. 5). Dabei werden nicht alle Muster geprüft, denn durch einen hierarchischen Klassifizierungsansatz kann die Kandidatenmenge auf Grund von Farbe und Form wesentlich reduziert werden. Verkehrszeichenarten, welche auf den erfassten Straßenabschnitten nicht vorkommen, können vom Klassifizierungsprozess ausgeschlossen werden.

Es wird das Referenzmuster bestimmt, mit welchem über den Suchbereich der größte normierte Kreuzkorrelationskoeffizient ermittelt wird. Dieser Wert dient auch als Klassifizierungsindikator. Überschreitet er einen vordefinierten Schwellwert, gilt die Klassifizierung als erfolgreich.

Die Größe des Suchbildes ist durch den Verkehrszeichenausschnitt im Normalbild zuzüglich eines definierbaren Bereichs (z.B. 10 Pixel) auf jeder Seite definiert. Das Referenzmuster wird auf die Ausdehnung des Farbsegments innerhalb des Suchbildes skaliert. Die Berechnung der Korrelation erfolgt auf Grund desjenigen Kanals, welcher bei empirischen Untersuchungen die größte Ähnlichkeit zwischen einem realen Verkehrszeichenbild und dem entsprechenden synthetischen Referenzmuster zeigte. Bei roten Verkehrszeichen ist dies der blaue Kanal, bei blauen der rote und bei gelben die Sättigungs-Komponente. Pixel des Suchbildes, welche im Referenzmuster weiß sind, weisen je nach Bildqualität einen zu tiefen Grauwert auf (siehe Abb. 5). Um die Matching-Ergebnisse zu verbessern, werden somit alle Pixelwerte des Suchbildes auf weiß (maximaler Grauwert) oder schwarz (Null)

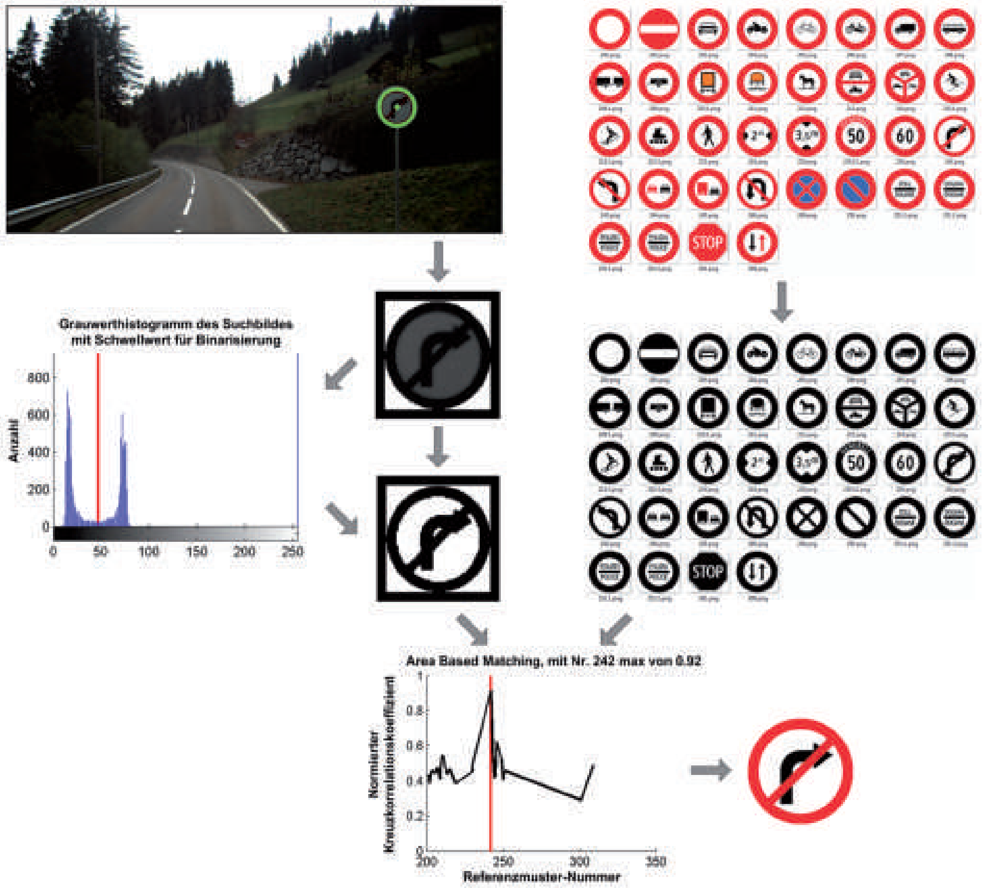


Abb. 5: Automatische Klassifizierung eines roten runden Verkehrszeichens mit kreuzkorrelationsbasiertem Template Matching.

gesetzt. Die Ermittlung des dafür benötigten Schwellwertes erfolgt dynamisch auf Grund der Grauwertverteilung des Suchbildes. Dabei wird jener Grauwert bestimmt, bei welchem der Wert des Summenhistogramms größer als ein bestimmter prozentualer Anteil der betrachteten Pixel ist. Bei Dreiecken ist dieser Anteil 40 % und bei den restlichen Formen 75 %. Die Grauwerte vom Maximum des linken und rechten Bereichs im Grauwerthistogramm werden ermittelt, der Mittelwert dient als Schwellwert.

Wenn ein Verkehrszeichen detektiert und klassifiziert werden konnte, erfolgt die Bestimmung der 3D-Objektkoordinaten. Für die Berechnung der Modellkoordinaten sind die Tiefeninformation und die Bildkoordinaten des Schwerpunktes sowie die Parameter der inneren Orientierung notwendig. Die anschließende Transformation in den gewünschten geodätischen Referenzrahmen setzt voraus, dass die äusseren Orientierungsparameter des linken Normalbildes bekannt sind.

Unterschiedliche Sachdaten wie die Landeskoordinaten des Schwerpunktes, die Referenzmusternummer und die genormten Seitenlängen werden automatisch bestimmt und in eine Textdatei ausgegeben. Diese können für den Aufbau oder die Aktualisierung eines Geoinformationssystems genutzt werden.

4.3 *Benutzerunterstützter Ansatz*

Verkehrszeichenkandidaten, welche einen zu niedrigen Detektions- oder Klassifizierungsindikator aufweisen, müssen nachträglich durch den Benutzer verifiziert werden. Dafür steht eine grafische Benutzeroberfläche zur Verfügung. Fälschlicherweise als Verkehrszeichen segmentierte Regionen können verworfen und eine falsche Klassenzuweisung korrigiert werden. Wurde ein Verkehrszeichen nicht automatisch detektiert, kann es interaktiv aus den Bilddaten extrahiert werden, indem der Schwerpunkt digitalisiert wird. Die manuelle Digitalisierung übernimmt die Rolle der Suchraumeinschränkung aus Abschnitt 4.1, ansonsten kommen die gleichen Algorithmen wie beim automatischen Ansatz zum Zuge. Dies ist vor allem dann hilfreich, wenn schlechte Tiefenkarten vorliegen. Sind die Schwellwerte für die Farbsegmentierung nicht gut gewählt, kann die ebenfalls implementierte Hough-Transformation zu besseren Ergebnissen führen. Dabei werden je nach Farbe die Formen Kreis, Rechteck, Quadrat, Dreieck und Raute getestet.

4.4 *Detektion von achromatischen Verkehrszeichen*

Für die Farbsegmentierung von weißen und grauen Verkehrszeichen erwies sich die Bestimmung von robusten Schwellwerten für die Farbton- und Sättigungs-Komponenten als untauglich, da für entsprechende Flächen in den Bilddaten keine klaren Grauwert-Intervalle ausgemacht werden konnten. Zudem weisen viele Bildpunkte in der Umgebung der Verkehrszeichen ähnliche Grauwerte auf. Der Ansatz von Paclík et al. (2000) konnte auch nicht verfolgt werden, da die Bestimmung eines festen Intervalls für die Hellwert-Komponente nicht erfolgreich war.

Nach Gómez-Moreno et al. (2010) kann die Segmentierung von achromatischen Farben durch die Verwendung der chromatischen/achromatischen Dekomposition verbessert werden. Dafür sind mehrere Ansätze denkbar, welche von ihnen vorgestellt und evaluiert werden. Der chromatische/achromatische Index, welcher von Liu et al. (2002) entwickelt wurde, ist einer davon und lässt sich wie folgt berechnen:

$$\mathrm{CAD}(R,G,B)=\frac{|R-G|+|G-B|+|B-R|}{3D} \qquad (3)$$

R, G, B repräsentieren die Farbkomponenten für ein gegebenes Pixel und D ist der Extraktionsgrad einer achromatischen Farbe. Ist der CAD-Wert kleiner oder gleich 1, handelt es sich um eine achromatische Farbe. Bei weißen Verkehrszeichen muss zusätzlich die Summe von R, G, B größer als ein definierter Schwellwert sein (Gómez-Moreno et al. 2010). Bei Liu et al. (2002) und Maldonado-Bascón et al. (2007) werden die besten Ergebnisse mit D = 20 erzielt, bei Gómez-Moreno et al. (2010) mit D = 30 und einem Schwellwert für weiß von 180. Diese Resultate konnten mit den vorliegenden Bilddaten nicht reproduziert werden. Somit wurde für die automatische Detektion von weißen und grauen Verkehrszeichen die generalisierte Hough-Transformation um die Tiefeninformation erweitert. Dabei werden zunächst Kanten mit dem Canny-Operator auf Grund der Hellwert-Komponente extrahiert. Alle Kantenpixel, welche sich in den Ebenensegmenten befinden, dienen als Eingabe für die Hough-Transformation. Geprüft werden die Formen Kreis und Rechteck. Die Einschränkung des Wertebereichs für die Maßstäbe erfolgt mit der Distanz zum Verkehrszeichen, wofür wiederum die Tiefenkarte zur Anwendung kommt. Für die anschließende Klassifizierung wird erneut das Template Matching aus Abschnitt 4.2 durchgeführt.

5 Untersuchungen und Resultate

Für die Güteuntersuchungen der entwickelten Algorithmen wurden mehr als 4000 georeferenzierte Stereobildpaare von einer Messfahrt mit dem mobilen Kartierungsfahrzeug des IVGI in Muttenz bei Basel verwendet. Die dabei eingesetzten Kameras, welche in einem Abstand von 92,7 cm montiert waren, weisen eine Kamerakonstante von 7,9 mm auf. Die Pixelgröße beträgt 7,4 µm bei einer Sensorgröße von 1926 x 1082 Pixel. Die Messfahrt erfolgte Mitte November 2010 bei bewölktem Himmel und entsprechend ungünstigen Beleuchtungsverhältnissen, was dunkle Bildaufnahmen zur Folge hatte. Pro Sekunde wurden 5 Full-HD-Stereobildpaare bei einer Fahrgeschwindigkeit von circa 40 km/h erfasst, was einem Stereobildabstand von etwa 2 m entspricht.

5.1 *Detektions- und Klassifizierungsgüte*

Für die Beurteilung der Detektions- und Klassifizierungsergebnisse erfolgte eine Unterteilung in Hauptstraßen und Quartierstraßen. Relevante Verkehrszeichen mit überwiegend roten, blauen und gelben Farbanteilen wurden bestimmt. In diesem Fall waren dies vor allem rechts an der Straße anliegende Verkehrszeichen, welche normalerweise senkrecht zur Fahrtrichtung ausgerichtet sind. Dies repräsentiert die Verkehrszeichen, welche in der Regel entlang von Hauptstraßen anzutreffen sind.

Bei der Prozessierung im Distanzbereich von 4 bis 14 m konnten mit den entwickelten Algorithmen 89 % der relevanten Verkehrszeichen automatisch detektiert werden (siehe Tab. 1). Korrekt klassifiziert wurden auf einer Hauptstraße 86 % und auf Quartierstraßen 82 %. Wie bei Höferlin & Zimmermann (2009) wurde auch der Indikator für die Präzision der Detektion berechnet:

$$\text{Präzision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4}$$

Dabei ist die Anzahl der 'True Positives' TP durch die Summe der 'True Positives' und 'False Positives' FP zu dividieren. Die hohen Werte in Tab. 1 sagen aus, dass sehr wenige Segmente fälschlicherweise als Verkehrszeichen ermittelt wurden.

Dass einige Verkehrszeichen nicht detektiert oder klassifiziert werden konnten, beruht auf unterschiedliche Ursachen. Bei der Detektion im Bereich von Baustellen ist es möglich, dass aufgrund von Abschrankungen und Warnvorrichtungen viele rote Farbsegmen-

Tab. 1: Detektions- und Klassifizierungsgüte der entwickelten Algorithmen für eine Messfahrt in Muttenz bei Basel (Alle: Verkehrszeichen mit überwiegend roten, blauen und gelben Farbanteilen; Relevante: rechts an der Straße anliegende Verkehrszeichen, welche vielfach senkrecht zur Fahrtrichtung ausgerichtet sind).

		Anzahl	**Detektionsrate**	**Klassifizierungsrate**	**Präzision**
Hauptstraße	Alle	90	72%	70%	94%
	Relevante	**63**	**89%**	**86%**	
Quartierstraßen	Alle	152	55%	47%	98%
	Relevante	**65**	**89%**	**82%**	

te ausgeschieden werden. Sind die Flächen nicht allzu groß, kann dies zu ‘False Positives’ führen. Obwohl die Verkehrszeichen in der Schweiz in der Regel einen guten Zustand aufweisen, sind gewisse Verkehrszeichen vergilbt. Dabei weist die Sättigungs-Komponente sehr niedrige Werte auf, was auch bei im Schatten liegenden Verkehrszeichen der Fall ist. Da der definierte Schwellwert für diese Komponente nicht überschritten werden kann, ist die Detektion nicht möglich. Verkehrszeichen, welche fast parallel zur Fahrtrichtung ausgerichtet sind, haben in den Normalbildern teilweise eine zu geringe Breite im Vergleich zur Höhe, womit diese Farbsegmente mit der aktuellen Lösung verworfen werden. Schwierigkeiten bereitet die automatische Detektion auch bei schlechten bzw. unvollständigen Tiefenkarten. Für mehrere Verkehrszeichen existierten keine vordefinierten Referenzmuster, was zu keiner oder zu einer falschen Klassifizierung führte. Ein nicht optimaler Schwellwert für die Binarisierung des Suchbildes kann die Ursache für einen zu niedrigen Korrelationskoeffizienten sein.

5.2 *Kartierungsgenauigkeit*

Für die Berechnung der Kartierungsgenauigkeit wurden die Referenzkoordinaten von 22 Verkehrszeichen mittels Tachymeter bestimmt. Die Differenzen zu den mit den entwickelten Algorithmen automatisch ermittelten Koordinaten wurden ausgewertet. Tab. 2 zeigt, dass für die Komponente quer zur Fahrtrichtung und für die Höhe eine empirische Standardabweichung von etwa 5 cm erreicht werden konnte und für die Komponente in Fahrtrichtung 6 cm. Erwartungsgemäß streut die Komponente in Fahrtrichtung bzw. senkrecht zur Stereobasis am meisten und ist somit am schlechtesten bestimmt. Die größte Abweichung einer Komponente beträgt 16 cm, die meisten Differenzen pro Komponente liegen jedoch im Bereich bis 5 cm. Die empirische Standardabweichung einer Koordinatendifferenz für die 3D-Position beträgt 9,5 cm.

6 Fazit und Ausblick

Es konnte aufgezeigt werden, dass durch eine konsequente Nutzung der Tiefeninformation aus Dense-Stereo-Matching in der automatisierten Verkehrszeichenkartierung sehr hohe Detektionsraten erreicht werden können. Unter guten Bedingungen können etwa 90 % der Verkehrszeichen mit überwiegend roten, blauen und gelben Farbanteilen im Normal- und Kleinformat in der Schweiz automatisch detektiert und 85 % korrekt klassifiziert werden. Benutzerunterstützt lassen sich diese Werte um weitere 5 % erhöhen. Somit müssen nur noch 5 bis 10 % der Verkehrszeichen interaktiv in den Stereobilddaten oder im Feld digitalisiert werden. Zudem treten durch den benutzerunterstützten Ansatz und diverse Bedingungen nahezu keine ‘False Positives’ auf. Eine vollautomatische Kartierung mit einer 3D-Genauigkeit von unter 10 cm ist möglich.

Der entwickelte Ansatz reagiert robust auf Skalierung, Translation und kleine Rotationen. Obwohl mit nahe gelegenen Verkehrszeichen bessere Ergebnisse zu erwarten sind, kann das Signal im ganzen vordefinierten Distanzbereich detektiert werden. Das Verkehrszeichen kann beliebig im Bild positioniert sein. Kleine Rotationen werden toleriert. Zudem ist es möglich, im gleichen Bild mehrere Verkehrszeichen mit den Formen Kreis, Rechteck, Quadrat, Dreieck und Raute zu detektieren. Eine reduzierte Leistung des Verfahrens ist bei schief stehenden oder stark verdrehten Verkehrszeichen sowie auf sehr stei-

Tab. 2: Praktische Kartierungsgenauigkeit der entwickelten automatischen Verkehrszeichenkartierung für eine Messfahrt in Muttenz bei Basel (mm).

	Δquer	**Δlängs**	**ΔHöhe**	**ΔLage**	**Δ3D**
Mittelwert	-36	23	-36	66	86
Maximum	152	146	157	154	159
m_{Diff}	**46**	**64**	**53**	**79**	**95**

len Straßen zu erwarten, was jedoch in der Praxis noch nicht beobachtet werden konnte.

Neben qualitativ guten Tiefenkarten ist auch eine gute Farbsegmentierung für den Detektionserfolg entscheidend. Dazu müssen passende Schwellwerte vorliegen. Bei den Untersuchungen wurde das Intervall für die jeweilige Komponente relativ groß gewählt. Dies ist aber nur möglich, weil der Suchraum durch die Tiefeninformation markant eingeschränkt wird und fälschlicherweise detektierte Verkehrszeichen durch gewisse Bedingungen verworfen werden.

Da Detektions- und Klassifizierungsgüten von 100 % kaum realistisch sind, besteht die Möglichkeit, die automatisch kartierten Verkehrszeichen in einem georeferenzierten 3D-Video zu überlagern. Die 3D-Videos können mit dem StereoVision-Client (Burkhard et al. 2011) betrachtet, die Ergebnisse somit kontrolliert und die fehlenden Verkehrszeichen spediti v digitalisiert werden. Eine erste Implementierung der Algorithmen für weiße und graue Verkehrszeichen, welche die Tiefeninformation in Kombination mit der Hough-Transformation verwendet, wird in Zukunft weiter verbessert. Geplant sind auch die Detektion von weiteren komplexen Verkehrszeichen und die Erkennung von beliebigem Text (Wu et al. 2005). Eine Erhöhung der geometrischen Qualität und Zuverlässigkeit lässt sich durch Matching in Stereobild-Sequenzen erreichen (Huber et al. 2011). Die Verfolgung der Verkehrszeichen über mehrere Stereobildpaare würde insbesondere eine Steigerung der semantischen Qualität bewirken.

Laufende Arbeiten sollen die Auswirkung der Kameraauflösung auf die Detektions- und Klassifizierungsgenauigkeit aufzeigen. Erste Untersuchungen mit einem Stereosystem bestehend aus höher aufgelösten Industriemesskameras mit je 11 MPixeln deuten auf eine leichte Verbesserung der Ergebnisse hin. Für die Erkennung von Text ist die höhere geometrische Auflösung zwingend notwendig. Laufende Untersuchungen zeigen zudem, dass die Qualität der Tiefenkarten mit den höher aufgelösten Bildsensoren und mit geeigneten radiometrischen Anpassungen nochmals deutlich gesteigert werden kann, was sich wiederum positiv auf die automatisierte Verkehrszeichenkartierung auswirkt.

Referenzen

Baró, X., Escalera, S., Vitrià, J., Pujol, O. & Radeva, P., 2009: Traffic Sign Recognition Using Evolutionary Adaboost Detection and Forest-ECOC Classification. – IEEE Transactions on Intelligent Transportation Systems **10** (1): 113–126.

Barrile, V., Cacciola, M., Meduri, G.M. & Morabito, F.C., 2007: Automatic Recognition of Road Signs by Hough Transform. – **5th** International Symposium on Mobile Mapping Technology, Padua, Italy.

Benesova, W., Lypetskyy, Y., Andreu, J.-P., Paletta, L., Jeitler, A. & Hödl, E., 2007: A Mobile System for Vision Based Road Sign Inventory. – **5th** International Symposium on Mobile Mapping Technology, Padua, Italy.

Burkhard, J., Nebiker, S. & Eugster, H., 2011: Stereobild-basiertes Mobile Mapping: Technologie und Anwendungen. – Geomatik Schweiz **109** (6): 295–298.

Chutatape, O. & Guo, L., 1999: A modified Hough transform for line detection and its performance. – Pattern Recognition **32** (2): 181–192.

Cyganek, B., 2008: Road-Signs Recognition System for Intelligent Vehicles. – Second International Workshop, RobVis **2008:** 219–233, Auckland, New Zealand.

de la Escalera, A., Armingol, J.M. & Mata, M., 2003: Traffic sign recognition and analysis for intelligent vehicles. – Image and Vision Computing **21** (3): 247–258.

de la Escalera, A., Armingol, J.M., Pastor, J.M. & Rodríguez, F.J., 2004: Visual Sign Information Extraction and Identification by Deformable Models for Intelligent Vehicles. – IEEE Transactions on Intelligent Transportation Systems **5** (2): 57–68.

de With, P., Hazelhoff, L., Creusen, I. & Bruinsma, H., 2010: Efficient Road Maintenance. Automatic Detection and Positioning of Traffic Signs. – GEOInformatics **13** (7): 10–12.

Fleyeh, H. & Dougherty, M., 2005: Road and Traffic Sign Detection and Recognition. – **10th** EWGT Meeting and **16th** Mini-EURO Conference: 644–653, Poznań, Poland.

Fleyeh, H., 2006: Shadow And Highlight Invariant Colour Segmentation Algorithm For Traffic Signs. – IEEE Conference on Cybernetics and Intelligent Systems, Bangkok, Thailand.

Fleyeh, H. & Dougherty, M., 2008: Traffic Sign Classification Using Invariant Features and Support Vector Machines. – IEEE Intelligent Vehicles Symposium: 530–535, Eindhoven, The Netherlands.

Fu, M.-Y. & Huang, Y.-S., 2010: A Survey of Traffic Sign Recognition. – International Conference on Wavelet Analysis and Pattern Recognition: 119–124, Qingdao, China.

Gao, X.W., Podladchikova, L., Shaposhnikov, D., Hong, K. & Shevtsova, N., 2006: Recognition of traffic signs based on their colour and shape features extracted using human vision models. – Journal of Visual Communication and Image Representation **17** (4): 675–685.

García-Garrido, M.A., Sotelo, M.A. & Martín-Gorostiza, E., 2006: Fast Traffic Sign Detection and Recognition Under Changing Lighting Conditions. – **9th** International IEEE Conference on Intelligent Transportation Systems: 811–816, Toronto, Canada.

Gómez-Moreno, H., Maldonado-Bascón, S., Gil-Jiménez, P. & Lafuente-Arroyo, S., 2010: Goal Evaluation of Segmentation Algorithms for Traffic Sign Recognition. – IEEE Transactions on Intelligent Transportation Systems **11** (4): 917–930.

Habib, A.F., Uebbing, R. & Novak, K., 1999: Automatic Extraction of Road Signs from Terrestrial Color Imagery. – Photogrammetric Engineering & Remote Sensing **65** (5): 597–601.

Hirschmüller, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. – IEEE Transactions on Pattern Analysis and Machine Intelligence **30** (2): 328–341.

Höferlin, B. & Zimmermann, K., 2009: Towards Reliable Traffic Sign Recognition. – IEEE Intelligent Vehicles Symposium: 324–329, Xi'an, China.

Hossain, M.S., Hasan, M.M., Ali, M.A., Kabir, M.H. & Ali, S., 2010: Automatic Detection and Recognition of Traffic Signs. – IEEE Conference on Robotics Automation and Mechatronics: 286–291, Singapore.

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. – Photogrammetric Image Analysis 2011, Lecture Notes in Computer Science **6952:** 85–95, Springer, Berlin.

Jau, U.L., Teh, C.S. & Ng, G.W., 2008: A comparison of RGB and HSI color segmentation in real-time video images: A preliminary study on road sign detection. – International Symposium on Information Technology: 2576–2581, Kuala Lumpur, Malaysia.

Khan, J.F., Bhuiyan, S.M.A. & Adhami, R.R., 2011: Image Segmentation and Shape Analysis for Road-Sign Detection. – IEEE Transactions on Intelligent Transportation Systems **12** (1): 83–96.

Kim, G.-H., Sohn, H.-G. & Song, Y.-S., 2006: Road Infrastructure Data Acquisition Using a Vehicle-Based Mobile Mapping System. – Computer-Aided Civil and Infrastructure Engineering **21** (5): 346–356.

Liu, H., Liu, D. & Xin, J., 2002: Real-time Recognition of Road Traffic Sign in Motion Image based on Genetic Algorithm. – First International Conference on Machine Learning and Cybernetics: 83–86, Beijing, China.

Lowe, D.G., 2004: Distinctive Image Features from Scale-Invariant Keypoints. – International Journal of Computer Vision **60** (2): 91–110.

Madeira, S.R., Bastos, L.C., Sousa, A.M., Sobral, J.F. & Santos, L.P., 2005: Automatic Traffic Signs Inventory Using a Mobile Mapping System. – International Conference and Exhibition on Geographic Information GIS Planet, Estoril, Portugal.

Maldonado-Bascón, S., Lafuente-Arroyo, S., Gil-Jiménez, P., Gómez-Moreno, H. & López-Ferreras, F., 2007: Road-Sign Detection and Recognition Based on Support Vector Machines. – IEEE Transactions on Intelligent Transportation Systems **8** (2): 264–278.

Maldonado-Bascón, S., Lafuente-Arroyo, S., Siegmann, P., Gómez-Moreno, H. & Acevedo-Rodríguez, J., 2008: Traffic Sign Recognition System for Inventory Purposes. – IEEE Intelligent Vehicles Symposium: 590–595, Eindhoven, The Netherlands.

Malik, R., Khurshid, J. & Ahmad, S.N., 2007: Road Sign Detection and Recognition using Colour Segmentation, Shape Analysis and Template Matching. – **6th** International Conference on Machine Learning and Cybernetics: 3556–3560, Hong Kong.

Nguwi, Y.-Y. & Kouzani, A.Z., 2008: Detection and classification of road signs in natural environments. – Neural Computing and Applications **17** (3): 265–289.

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 (31.1.2012).

Paclík, P., Novovicová, J., Pudil, P. & Somol, P., 2000: Road sign classification using Laplace kernel classifier. – Pattern Recognition Letters **21** (13/14): 1165–1173.

Piccioli, G., De Micheli, E., Parodi, P. & Campani, M., 1996: Robust method for road sign detection and recognition. – Image and Vision Computing **14** (3): 209–223.

Reiterer, A., Hassan, T. & El-Sheimy, N., 2009: Robust Extraction of Traffic Signs from Georeferenced Mobile Mapping Images. – **6th** International Symposium on Mobile Mapping Technology, Presidente Prudente, São Paulo, Brazil.

Ren, F., Huang, J., Jiang, R. & Klette, R., 2009: General Traffic Sign Recognition by Feature Matching. – **24th** International Conference on Image and Vision Computing New Zealand: 409–414, Wellington.

Scharstein, D. & Szeliski, R., 2012: Middlebury Stereo Evaluation – Version 2. – http://vision.middlebury.edu/stereo/eval/ (31.1.2012).

Schweizer Bundesrat, 1979: Signalisationsverordnung. – http://www.admin.ch/ch/d/sr/7/741.21.de.pdf (31.1.2012).

Schweizer Fernsehen, 2008: Überflüssige Schilder: Teuer und gefährlich. – http://www.tagesschau.sf.tv/Nachrichten/Archiv/2008/08/26/Schweiz/Ueberfluessige-Schilder-Teuer-und-gefaehrlich (31.1.2012).

Shi, Y., Shibasaki, R. & Shi, Z.C., 2008: Towards Automatic Road Mapping by Fusing Vehicle-Borne Multi-Sensor Data. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **37** (B5): 867–872.

Trimble, 2009: Signs of Change in Belgium. – Technology & more **2009** (3): 12–13.

Wikipedia, 2012: Bildtafel der Verkehrszeichen in der Schweiz. – http://de.wikipedia.org/wiki/Bildtafel_der_Verkehrszeichen_in_der_Schweiz (31.1.2012).

Wu, W., Chen, X. & Yang, J., 2005: Detection of Text on Road Signs From Video. – IEEE Transactions on Intelligent Transportation Systems **6** (4): 378–390.

Adresse der Autoren:

Stefan Cavegn & Prof. Dr. Stephan Nebiker, FHNW Fachhochschule Nordwestschweiz, Institut Vermessung und Geoinformation, CH-4132 Muttenz, Tel.: +41-61-467-4336, Fax: +41-61-467-4460, e-mail: {stefan.cavegn}{stephan.nebiker}@fhnw.ch

Manuskript eingereicht: September 2011
Angenommen: März 2012

PFG 2012 / 5, 0647–0652
Stuttgart, October 2012

Report

Aufbau und Charakteristika der Geodateninfrastruktur von Austro Control – Eine GDI für die Zivilluftfahrt im Kontext von INSPIRE

Markus Schaffert, Dessau & Oliver Starka, Wien, Österreich

Schlüsselbegriffe: Geodateninfrastruktur, GDI, INSPIRE, Zivilluftfahrt, statische Geodaten

Zusammenfassung: Die Firma Austro Control Ges.m.b.H. (ACG) ist für die Flugsicherung der zivilen Luftfahrt in Österreich zuständig. Der durch INSPIRE gesetzlich vorgeschriebene Aufbau einer Geodateninfrastruktur (GDI) wurde zum Anlass genommen, auch die internen Geschäftsprozesse zu optimieren. Die für externe und interne Aufgaben entstandene GDI-ACG besteht aus der Basiskomponente Geodaten sowie den Komponenten Metadaten, Geowebserver und Geowebportal. Das Portal ist ein Thin-Client und integriert viele Funktionen, die in den üblichen Geowebportalen nicht enthalten sind. Dazu gehören die Online-Darstellung aller Flugzeugpositionen, die dritte Dimension und meteorologische Informationen.

Summary: *The company Austro Control Ges.m.b.H. (ACG) is responsible for the air traffic control of the civil aviation in Austria*. The demand for the development of a spatial data infrastructure (SDI) based on the INSPIRE directive led also to a reorganization of the internal processes of ACG. The SDI-ACG consists of the base-component Geodata, and of the components Metadata, Geowebserver, and Geowebportal. The portal is a thin-client solution and integrates many functions that are not available in standard web-browsers including the online-position of all aircraft, the third dimension, and meteorological information.

1 Einleitung

„Austro Control sorgt rund um die Uhr dafür, dass sicherheitskritische Informationen wie Radar- und Navigationsdaten sowie Funk-, Wetter-, Flugplan- und Flugberatungsdaten ohne Unterbrechung zur Verfügung stehen und garantiert damit tagtäglich höchste Sicherheitsstandards in der Luftfahrt“ (Austro-Control 2002). Um diese Vorgaben zu erfüllen, bedarf es einer progressiven und robusten IT. Die Geodateninfrastruktur (GDI) der Austro Control ist Teil dieser IT und Thema dieses Artikels.

Die Austro Control – Österreichische Gesellschaft für Zivilluftfahrt m.b.H. (ACG) ist ein privatwirtschaftlich organisiertes Unternehmen, das sich zu 100 Prozent im Besitz der Republik Österreich befindet und das für den sicheren sowie wirtschaftlichen Ablauf des zivilen Flugverkehrs im österreichischen Luftraum verantwortlich ist. Aus dieser Verantwortlichkeit und der zentralen Lage Österreichs in Europa ergibt sich zudem die Bedeutung der Austro Control für den internationalen Luftverkehr.

Für eine Vielzahl von Geschäftsprozessen der Austro Control, wie z.B. die Produktion von aeronautischen Karten, bilden raumbezogene Daten die Grundlage. Diese Daten stammen aus unterschiedlichen unternehmensinternen sowie -externen Quellen und werden in diversen Dateiformaten vorgehalten bzw. weitergegeben. Das Zusammenführen der Daten verursachte deshalb bislang einen nicht zu unterschätzenden Aufwand für die Austro Control und behinderte das zielgruppengerechte Inwertsetzen der Daten durch die Fachabteilungen.

Zur Verbesserung der Geodatenbereitstellung und der nahtlosen Geodatenintegration in bestehende Geschäftsprozesse wurde deshalb der Aufbau einer Geodateninfrastruktur

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0145 1432-8364/12/0145 $ 1.50

durch die Austro Control initiiert und in Zusammenarbeit mit einem Unternehmenskonsortium bestehend aus der Frequentis AG, der Intergraph Ges.m.b.H. sowie der CNS Solutions & Support GmbH umgesetzt.

2 Die GDI-ACG im Kontext von INSPIRE

Bestehende Bestrebungen der Austro Control zur Optimierung interner Geschäftsprozesse durch den Aufbau einer Unternehmens-GDI erhielten durch die INSPIRE-Richtlinie (INSPIRE 2007) und deren Überführung in nationales Recht (GeoDIG 2010) im Jahre 2010 einen obligatorischen Rahmen. INSPIRE verpflichtet die Mitgliedstaaten, in der öffentlichen Verwaltung vorhandene und in digitaler Form vorliegende Geodaten zu definierten Themenbereichen in vorgegebener Struktur und über standardisierte Geowebdienste verfügbar zu machen.

Obwohl privatwirtschaftlich geführt, wurde INSPIRE von Austro Control zum Anlass für die Implementierung einer Geodateninfrastruktur gemäß den in den INSPIRE-Umsetzungsanleitungen definierten technischen Spezifikationen genommen. Durch die auf diese Weise möglich werdende Interoperabilität von Geodaten(-sätzen) und Diensten soll sich die GDI-ACG nahtlos in INSPIRE-getriebene Geodateninfrastrukturknoten auch außerhalb des Unternehmens integrieren lassen. Durch standardkonforme Geowebdienste können bei Bedarf Geodaten von externen Stellen, z.B. die für die Kartenproduktion der Austro Control verwendeten Geobasisdaten des österreichischen Bundesamtes für Eich- und Vermessungswesen, direkt über das Internet in die eigene Infrastruktur eingebunden werden. Umgekehrt lassen sich die bei der Austro Control erfassten bzw. fortgeführten Geofachdaten über das Internet in einheitlichem durch INSPIRE definiertem Standard externen Stellen anbieten.

INSPIRE und in Entsprechung das Geodateninfrastrukturgesetz der Republik Österreich sehen eine schrittweise Umsetzung der durch INSPIRE getroffenen Vorgaben bis ins Jahr 2019 vor. Die Erarbeitung technischer Empfehlungen bzw. Vorschriften durch INSPIRE sind zu einem Teil bereits abgeschlossen; teilweise begleiten sie aber die schrittweise Umsetzung noch oder sie lassen künftige inhaltliche Entwicklungen bei der Versionierung von relevanten Standards erwarten. Diese Veränderlichkeit der technischen Spezifikationen und die fortwährende Notwendigkeit, die implementierte GDI-Software entsprechend anzupassen, um interoperabel zu weiteren Knoten der europäischen GDI-Landschaft zu sein, führte zur Entscheidung, gewartete Softwareprodukte von kommerziellen Anbietern für die GDI-ACG zu verwenden: Der dauerhafte Abgleich der verwendeten Softwareprodukte mit Vorgaben des evolutionären INSPIRE-Prozesses sollte extern erfolgen und die Arbeit der ACG nicht beeinträchtigen.

Die Bedeutung der GDI-ACG für die Umsetzung der INSPIRE-Richtlinie in Österreich zeigt sich darin, dass sie als das erste öffentlich ausgeschriebene Geodateninfrastrukturprojekt in Österreich gilt (Intergraph 2011). Vom Engagement der ACG dürfte deshalb eine Signalwirkung für die INSPIRE-Umsetzung in Österreich ausgehen und die GDI-ACG als Referenz für sich noch entwickelnde GDI-Knoten des Landes dienen.

3 Die GDI-ACG in ihren Komponenten

3.1 *GDI-Basiskomponente „Geodaten“*

Die Luftfahrt unterscheidet zwischen dynamischen und statischen Daten. Sowohl dynamische als auch statische Daten haben regelmäßig einen Raumbezug. Dynamische Daten sind zum Beispiel Flugplan-, NOTAM (Notice to Airmen)- und Wetterdaten. Als NOTAM werden die traditionellen Freitextmitteilungen über temporäre sicherheitskritische Ereignisse bezeichnet, die meist kurzfristig und deshalb innerhalb des 28-tägigen AIRAC-Zyklus an die Luftfahrtteilnehmer geleitet werden müssen. Dies sind bspw. Informationen über aktuelle Sperrungen von Start- bzw. Landebahnen oder Gefahren, wie Fallschirmsprünge oder Vogelschwärme. Dagegen decken statische Daten ein breiteres Spektrum ab. Die für die Arbeit der ACG relevanten statischen

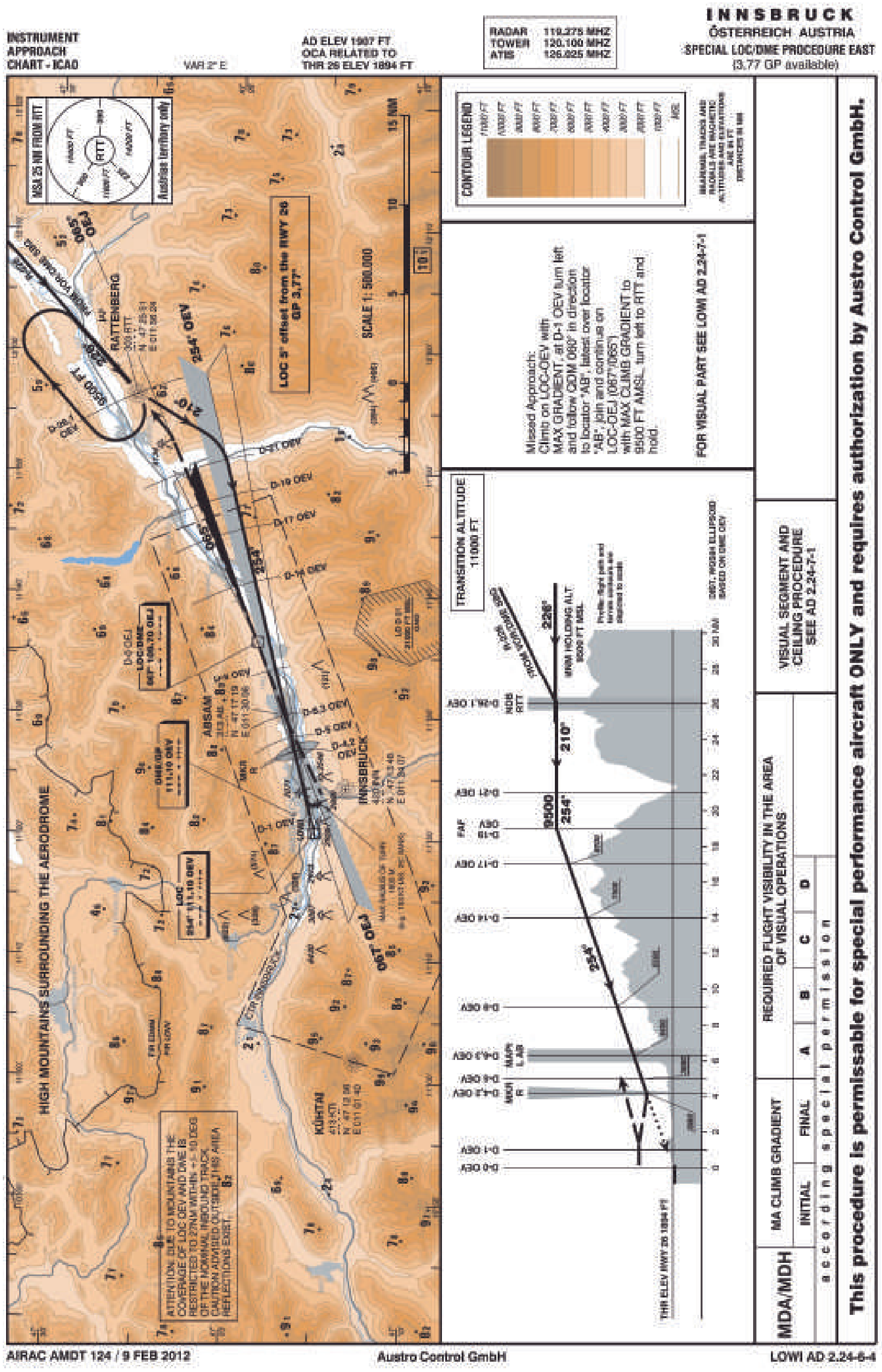


Abb. 1: Beispiel für eine fachthemenbezogene Luftfahrtkarte (© Austro Control Ges.m.b.H.).

Geodaten lassen sich thematisch in Geländedaten (z.B. DHM), Navigationsdaten (z.B. GPS-Tracks), sonstige Geobasisdaten (z.B. Gemeinde- und Landesgrenzen), meteorologische Daten (z.B. Windverhältnisse) sowie aeronautische Daten (z.B. Bauhindernisse, Flugrouten oder Sicherheitszonen) ordnen.

Aus den statischen Geodaten generiert das Unternehmen Luftfahrkarten zu diversen Fachthemen (Abb. 1). Die Luftfahrtkarten werden sowohl einzeln als auch als Bestandteil des österreichischen Luftfahrthandbuchs veröffentlicht. Die Bedeutung gerade der statischen Geodaten für die Arbeit der Austro Control ergibt sich aus der Verpflichtung zur Erstellung und Fortführung des Luftfahrthandbuchs gemäß internationaler Konvention. Seine Publikation folgt einem weltweit einheitlichen Schema, das eine verbindliche chronologische Aktualisierungsfolge (AIRAC-Zyklus (Aeronautical Information Regulation and Control)) vorschreibt und ebenfalls durch die Convention on International Civil Aviation der International Civil Aviation Organization (ICAO) festgelegt ist.

Die statischen Geodaten, die durch Fachabteilungen oder unternehmensferne Quellen wie dem österreichischen Bundesamt für Eich- und Vermessungswesen (BEV) erfasst werden, bilden die Datenbasis der GDI-AGC. Sie liegen bei der Austro Control in vielfältigen Formaten vor: Microstation DGN, AutoCAD DWG und DXF, ESRI SHP, KML, Oracle Spatial, CSV, XLS, ASCII. Die Datenhaltung der GDI-ACG sieht das Zusammenführen all dieser Daten in einer zentralen Datenbank (Oracle 11g R2) vor.

Um den Anforderungen der Luftfahrt gerecht zu werden, muss das entwickelte Datenmodell auch die Fortführung der Daten entsprechend dem AIRAC-Zyklus unterstützen und das AIXM 5 (Aeronautical Information Exchange Model) beachten. AIXM basiert auf GML und gibt aeronautischen Informationen, unter anderem auch NOTAM, eine prozessierbare Struktur.

3.2 *Komponente „Metadaten“*

Um eine webbasierte, INSPIRE-konforme Erfassung von Metadaten zu gewährleisten, verwendet die Austro Control die Intergraph Software Catalogue Editor Web (CEW). Der CEW bietet eine Erfassungsoberfläche, die die Anforderungen des INSPIRE-Metadatenprofils abbildet. Darüber hinaus lassen sich einmal erfasste Metadaten durch die Software auf ihre INSPIRE-Konformität hin validieren.

Neben dem INSPIRE-Metadatenprofil wurde für interne Zwecke und zur Erfüllung der regulativen Anforderungen der Austro Control zudem ein spezifisches Metadatenprofil erarbeitet, welches eine Erweiterung des INSPIRE-Profils darstellt. Dieses wird in der Software als alternatives Metadatenerfassungsprofil angeboten, aus dem sich automatisiert das INSPIRE-Profil ableiten lässt.

Das Zusammenspiel der einzelnen GDI-Komponenten ermöglicht, dass sich dieser CSW (Catalogue Service for the Web) sowie weitere standardkonforme CSWs in das Geowebportal der ACG-GDI einbinden lassen. Dort kann – neben dem Bedienen von portalspezifischen Funktionen – auch nach Metadaten recherchiert und die gefundenen Dienste direkt im Kartenviewer dargestellt bzw. bearbeitet werden. Dadurch werden zum einen parallele Softwarelösungen für die Metadatenrecherche obsolet. Zum anderen setzt dieser Ansatz das Publish-Find-Bind-Muster einer serviceorientierten Softwarearchitektur (SOA-Paradigma) um.

3.3 *Komponente „Geowebserver“*

Der „Geowebserver“ ist diejenige Infrastrukturkomponente der GDI-ACG, welche die Geodaten und die dazugehörigen Metadaten mittels Geowebdiensten bereitstellt. Die für die Geowebserverimplementierung verwendete Software der Firma Intergraph, namentlich GeoMedia WebMap und GeoMedia SDI Pro, wird entsprechend den Empfehlungen von INSPIRE entwickelt.

3.4 *Komponente „Geowebportal“*

Das Geowebportal bildet die zentrale Nutzerschnittstelle der GDI-ACG. Folglich muss es besondere Anforderungen in Hinblick auf Bedienbarkeit und Performance erfüllen. Darüber hinaus ist es notwendig, dass sich wesentliche Geowebdienste in ein solches Portal einbinden und gemeinsam darstellen lassen. Das Geowebportal der Austro Control, welches mit der Software GeoMedia SDI Portal von Intergraph umgesetzt wurde, gewährleistet dies für praktisch alle vom OGC verabschiedeten Web-Dienstestandards.

Über die von der Austro Control geforderten Standards hinaus sind zudem auch nicht standardisierte, aber verbreitete Dienste wie Google Maps, MS Bing, Intergraph Map Publisher Service und das ERDAS APOLLO ECWP streaming in die Portalsoftware einbindbar. Das Portal erlaubt zudem die integrierte Suche nach Metadaten von Datensätzen, Datensatzreihen und Diensten innerhalb der Portaloberfläche. Die Ergebnisbereiche (Bounding Box) gefundener Metadaten können dabei ebenso wie die Inhalte gefundener Dienste per Mausklick im Portal-Kartenviewer dargestellt werden.

Das Dienstespektrum des Geowebportals der GDI-ACG lässt sich durch die Möglichkeit zum rechtegesteuerten Dateidownload in verschiedenen Dateiformaten wie KML, DGN oder CSV direkt aus der Portaloberfläche heraus ergänzen.

Das implementierte Portal stellt eine Thin-Client-Lösung dar, die in allen gängigen Internetbrowsern lauffähig ist und keine clientseitige Installation erfordert. Dennoch bietet es GIS-Basisfunktionalitäten, die deutlich über Möglichkeiten zur gemeinsamen Darstellung und Symbolisierung von Dienste-Layern oder zur Anzeige von Attributen einzelner Geoobjekte im Kartenviewer hinausgehen. So lassen sich in der Portaloberfläche – basierend auf den jeweiligen Diensten und ihren spezifizierten Vorgaben – auch Analysen durchführen (WFS, Web Feature Service) und Objekte erzeugen bzw. verändern (WFS-T, WFS-Transactional). Der Client bietet darüber hinaus auch Druck- und Redlining-Funktionalitäten sowie die Möglichkeit, Kartenansichten bestimmter aeronautischer Themen in chronologischer Reihenfolge mittels Schieberegler anzuzeigen. Dadurch lassen sich, basierend auf Geowebdiensten sowie unter Beachtung von AIXM-Daten und Nutzerberechtigungen, auch AIRAC-Zyklen abbilden und vom Anwender nachvollziehen (Abb. 2).

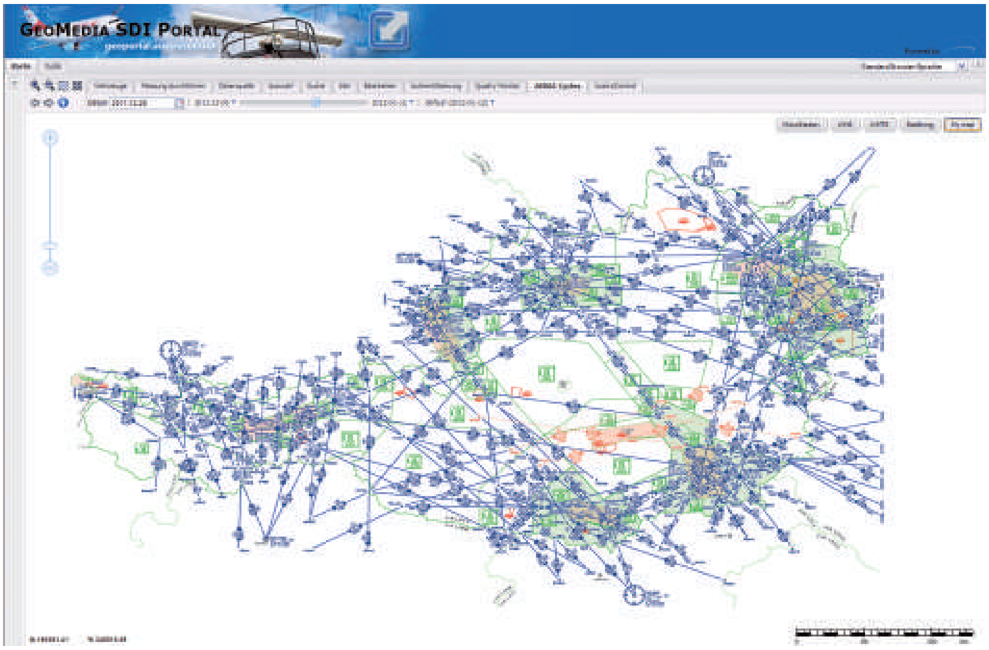


Abb. 2: Das Geowebportal der Austro Control (© Austro Control Ges.m.b.H.).

Durch die Integration von GPS-Tracks im GPX-Format sind zudem Flugrouten im Portal darstellbar. Desweiteren erhält der Bearbeiter für die jeweils aktuelle Mausposition im Kartenviewer nicht nur x- und y-Wert für mehrere Koordinatensysteme, sondern auch die Höhe in Metern – eine für die Luftsicherheit elementare, in Geowebportalen „out-of-the-box“ aber regelmäßig nicht implementierte Anforderung.

4 Fazit

Mittel- und langfristig bieten Geodateninfrastrukturen Möglichkeiten, die über heute bereits greifbare Vorteile hinausgehen dürften. Für die GDI-ACG könnte ein solches Zukunftspotential im verbesserten Austausch von Geodaten zwischen den Fachabteilungen der Flugsicherung und den Luftfahrzeugen selbst liegen: So werden heute bei jedem Flug noch vielfältige gedruckten Flugkarten durch die Piloten im Cockpit mitgeführt – die Zukunft sollte jedoch auch hier der digitalen Karte gehören (Grothe 2005), so dass dann die Luftfahrzeuge selbst als integrierter Teil der GDI anzusehen sind (Robin et al. 2011). Umgekehrt würden sich spätestens zu diesem Zeitpunkt auch weiterführende Anforderungen an die Integration von dynamischen, in den Fahrzeugen erfassten Geodaten in eine umfassende Flugsicherungs-GDI stellen und sich durch die Verarbeitung dieser Daten weiterer Mehrwert für die Sicherheit und Wirtschaftlichkeit der Zivilluftfahrt ergeben.

Literatur

AustroControl 2002: Geschäftsbericht 2001 der Austro Control Ges.m.b.H.

GeoDIG, 2010: Bundesgesetzblatt für die Republik Österreich. Ausgegeben am 1. März 2010 Teil I: 14. Bundesgesetz über eine umweltrelevante Geodateninfrastruktur des Bundes (Geodateninfrastrukturgesetz – GeoDIG).

Grothe, C., 2005: Updates für Onboard-Datenbanken. – Masterarbeit am Fachbereich Informatik der Technischen Universität Darmstadt, http://www.dvs.tu-darmstadt.de/publications/MScs/Christian.Grothe_May05.pdf (28.4.2012).

INSPIRE, 2007: Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). – http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:en:PDF (28.4.2012).

Intergraph, 2011: GDI-Lösungen von Intergraph unterstützen raumbezogene Datendienste der österreichischen Zivilluftfahrtbehörde Austro Control, Pressemitteilung 07/2011. – SG&I Deutschland GmbH http://www.intergraph.com/global/de/pressreleases/2011/sgi28062011.aspx (28.4.2012).

Robin, C., Kopeinig, R., Leitner, M., Gäbler, K. & Garcia-Barrios, V.M., 2011: Aerodrome Mapping Database – Entwurf und Modellierung eines Datenbanksystems für Flughäfen. – Konferenzbeiträge der AGIT **2011:** 924–934, Salzburg.

Anschriften der Autoren:

Dr.-Ing. Dipl.-Geogr. Markus Schaffert, Institut für angewandte Geoinformatik und Raumanalysen (AGIRA) e.V., Bauhausstraße 8, D-06846 Dessau, Tel.: +49-340-5197-1620, e-mail: schaffert@institut-agira.de

Ing. Mag. Oliver Starka, Intergraph Ges.m.b.H. Österreich, Margaretenstr. 70/I/1, A-1050 Wien, Tel.: +43-1-9610567-4672, Fax: +43-1-9610567-4601, e-mail: oliver.starka@intergraph.com

Manuskript eingereicht: März 2012
Angenommen: Juni 2012

Berichte von Veranstaltungen

Interexpo Geo-Siberia 2012, April 17–19, 2012, Novosibirsk, Russia

Russian educational institutions in geodesy and geoinformatics have a long tradition. MIIGAiK, the Moscow State University of Geodesy and Cartography, was established 1779 in Moscow under Catherina the Great. With Russia as the largest country in the world, it played a vital role in the surveying and mapping task since 1917, with the map coverage completed in 1955. In 1935 the educational capacity in geodesy was augmented by the establishment of NIIGAiK, the Novosibirsk State University of Geodesy and Cartography, supplying the Russian manpower in geodetic engineering. NIIGAiK has now been renamed in Siberian State Academy for Geodesy (SSGA). The Russian geodetic educational system of these two institutions was introduced at Wuhan in China in 1956 for the Wuhan Technical University for Surveying and Mapping (WTUSM), which is now part of Wuhan University.

SSGA started to organize annual conferences and exhibitions in geodesy and geoinformatics in 2005 at Novosibirsk, known as “Geo Sibir”. In 2012 the events were renamed to “Interexpo Geo-Siberia”. The event not only has a strong national component, but it also includes a large international participation with cooperating countries, such as Germany, Kazakhstan, Mongolia and China, and with international organizations such as FIG, ISPRS and ICA. Rector prof. Aleksander Karpik and vice rector prof. Vladimir Seredovich summarized the event: In 2012 there were 600 participants (35 of them from abroad) and 140 international exhibitors at the new Expo Center in Novosibirsk.

The technical program was presented in 22 sessions with over 700 presentations. 655 papers were published in the proceedings. Special international sessions with simultaneous translation were organized on Disaster Management by ICA (Milan Konecny, Czech Republic) and GNSS engineering surveys by FIG (Rek, Schmitt, Jaeger, Schwieger, Germany, and Milev, Bulgaria). Special attention was given to the presentation of a geoportal by the Novosibirsk Oblast, which manages geodata for the cadastre and the environment. The Russian operational use of internationally acquired satellite imagery by Scanex was likewise very impressive.

SSGA will host a workshop in September 2012 at Almaty, Kazakhstan, in cooperation with local Universities, as it has done in 2011 in Ulanbaatar, Mongolia, and in 2010 in Ust-Kamenogorsk, Kazakhstan.

Gottfried Konecny, Hannover

Participants at closing of the conference Interexpo Geo-Sibiria.

DOI: 10.1127/1432-8364/2012/0111 1432-8364/12/0111 $ 4.00

Geospatial World Forum, April 23–27, 2012, Amsterdam, The Netherlands

GIS Development, now renamed "Geospatial World", calls itself a Geospatial Communications Network, which provides a platform for the geospatial community in the world to exchange and to promote ideas on how industry, government and academia may contribute towards a sustainable development of the world. The seat of this development is Delhi in India. It communicates via publications and conferences. Beginning with conferences in India, other meetings were held regularly in the Middle East and South East Asia since 2005, and more recently also in Latin America and Africa. In 2007 GIS Development began to organize the first World Forum in Hyderabad in India. This magnificent effort brought together geospatial technology leaders of the globe. It was repeated in 2007 and 2009 in India's cradle of IT.

In 2012 the Geospatial World Forum was held in Europe for the first time, at the Amsterdam Expo Center R.A.I. At this event, there were 1054 delegates, 415 speakers and 51 exhibitors from 78 countries. 4 plenary sessions were held. There were 4 Exchange Forums (*Earth Observations, Land Administration, Capacity Building, Public Private Partnership*); also 6 symposia (*Energy, Water, Mining, Construction, Defence, Enterprises*) and 9 seminars (*Cloud Computing, Real-Time, Image Information, European Data, Mobile Mapping, 5D Modelling, Standards, Open Source, Environmental Management*); 6 further sessions (*Cartography, Governance, WebGIS, Cloud Computing, Surveying, Photogrammetry*). Many of the events were held in parallel so that it was only possible to comment on the plenary sessions and on some other highlights. Most of the presentations are downloadable from the Geospatial World Forum 2012 Website. This is a wealth of materials, characterizing the present state of development in geoinformatics, which was estimated as a 2 billion $ per annum business.

The Forum was officially opened on Tuesday, April 24, 2012 by Dorine Burmanje, president of the Dutch cadastre, which was the local organizer of the Forum. She welcomed Abdul Kalam, the 11th president of India. Cadastre and Mapping have been forerunners of geospatial activity in the Netherlands. Four of the six national key registers are now geospatial. Steven Berglund, president of Trimble, noted, that technology has been the geospatial enabler since 25 years. Geospatial activities now become central for energy, oil and gas, mining, telecommunications, infrastructure, transportation, water, agriculture, local government and forestry. We note, that three technologies have merged, sensors, software, and wireless communications. K.K. Singh, chairman and director of Rolta, considered the geospatial technology mature to transform the world from mapping toward spatial analysis, noting that the mind is visually oriented. Melanie Schultz van Haergen-Maas Geesteranus, the Dutch Minister for Infrastructure and Environment, confirmed that geoinformation is crucial. In the Netherlands the country is flown with 10 cm GSD images every year to be able to map the country with an accuracy of 20 cm. There was a national effort toward standardization of addresses and the establishment of a national satellite image data base. Netherlands will have free and open data in 2015. Her statement was: "*the government procures the data, the community provides the apps*" concluding that information is a source for learning, unless its organization is a burden. Former president Abdul Kalam of India delivered an address on "*geospatial technology for sustainable development*". Primarily for 3 billion people in rural areas India has launched 200 satellites. ISRO data are shared by 70 countries. Satellite data have provided a ground water map of India, showing that 46 million ha are still wasteland, which need to be made productive. Fishing has benefitted from twice weekly measurements of ocean temperature and chlorophyll. Children have been engaged with GPS-GIS to map the socio-economic environment, to find water bodies and to maintain a clean neighbourhood.

The plenary sessions started with a panel discussion on Geospatial Industry chaired by Ingrid van den Berghe, president of EuroGeographics, Belgium. Jack Dangermond, president of Esri, USA, noted that the world changes rapidly (climate, population, politics, energy, water, biodiversity). But the geospatial community is growing, permitting crowd sourcing. A cloud GIS can be utilized for sharing data including utilization of sensor networks. Olla Rollen, president of Hexagon, Sweden, compared the development of geospatial technology with the growth of the music industry. Traditional governmental high accuracy data with few updates have been changed to low accuracy data with faster updates by Google. This may lead to low accuracy real time updates when the information cloud is unlocked. A tendency in surveying, photogrammetry, remote sensing and CAD towards 3D can be stated. Also, a geospatial mixing of imagery and the point cloud can be expected. Greg Bentley, CEO of Bentley Systems, USA, stressed the importance of infrastructure for billions of people. In 166 countries a 14 trillion $ is served to 78% by Bentley products. In the 1980's it was by 2D CAD. Since 2000 it was in 3D. Since 2010 it was by semantic web. The initiation of a 3D cadastre in the Netherlands has disclosed that there is an underground utility chaos. Paul Cheung, director of the UN Statistics Division, New York, noted that industry is in front of the interest at the Geospatial World Forum, and government in the back. But the role of government is not small. Governments are pushed by Google and Microsoft, but industry cannot be the sole producer of geospatial data, even at the time when budgets are cut. Governments need a collective voice but the geospatial institutions are not represented in national frameworks. Therefore, Global Geospatial Information Management (GGIM) promoted by the UN Secretariat was established with four major tasks: (1) for countries to believe in GGIM, (2) to generate an inclusive platform for national and global data sharing, (3) to include an ethics commitment, and (4) to develop mutual beneficial relationships between governments and industry. Governments need to realize a vision: accuracy, authoritativeness, reliable supply of information to support national, regional and global development.

The panel discussion on "*Enabling a Spatial Culture*" was chaired by Vanessa Lawrence, director general of the Ordnance Survey, UK. Steven Hagan, vice president of Oracle, USA, outlined the challenges with geospatial data like big data volumes, cloud computing (unstructured data), and spatial data bases (semantic web). Alain de Taeye, board member of TomTom reported that a navigational coverage, accurate and up to date, for 36 million km of roadways has been created. This includes 156 million address points. Navigational systems have become dynamic maps including traffic information. Bem Semmes, Pitney Bowes, concluded that there is an increasing demand for answers based on smarter customer experience (reservations, restaurants). David Schell, OGC founder, started by saying: "*I came into this business to make progress, not money*". There is a convergence of applications. The geosciences are the pillar. There are major social issues, e.g. water, and there is a need for better education, otherwise we cannot communicate; the cloud needs a service bureau, since there is a controversy between belief and facts. Sir Stewart Peach, Joint Forces Command, pointed out that technology is a bridge to enable a geospatial culture. But vision is required, and teamwork is a critical factor to avoid silos. Creativity cannot be stopped by government regulations.

The Exchange Forum "*Earth Observation System for Nation Building*" started with the Keynote of Gottfried Konecny, Hannover. He stated that technology, such as GNSS, satellites, digital imaging and wireless communication were the drivers in earth observation. But it needed cooperative programs to make a global impact. Such programs were GEO on the global level and GMES for Europe. An especially effective program was the International Charter for Disaster Relief in which the space agencies of the globe cooperated on a voluntary basis. Ger Nieuwport of the Netherland's Space Office discussed national perspectives for earth observations. There is a technology push also nationally, but there is no full earth observation market, so that technology is underutilized. The need is for information, not for data. Global issues like GMES depend on complex systems. Scientific data from satellites become available, e.g. the Sentinels, and they are available in an open data policy. For that purpose a Dutch satellite data portal has been established. Gunter Schreier, DLR, Germany, discussed 3 challenges for earth observation: Urbanization, natural disasters, and civil security. Geoff Sawyer, EARSC (European Association of Remote Sensing Companies) established in 1999 as a non-profit organization with 65 members in 22 countries, estimated an 800 million to 1 billion € business for these companies in 2006. GMES is a European key program and should make data available free and open (Sentinels) or under modest cost licensing agreements. Finland, Netherlands, Spain, UK, Australia aim for a minimum cost model. After the failure of Landsat commercialization, Landsat and CBERS data are free since 2008. JOHN HORNSBY, MDA, Canada, noted the growth of the number of satellite operators: in 1997 there were 69 satellites by 17 operators, in 2007 151 satellites by 28 operators. Emphasis has shifted from research (1972) to natural resources and environment uses (1990) to time critical uses (2010). TONY FRAZEER, Vice President of GeoEye, USA, pointed out that, with satellite images, governments can do more with less; they can concentrate on underserved markets; they can provide acceptable solutions faster and cheaper, enabling leapfrogging. Examples are public health improvements (malaria), public safety issues (tribal conflicts) in Africa, land ownership issues in Russia where already 49 million parcels have been mapped from satellite images, urban issues in India, creating 3D city models to augment urban planning. TODD DOUGLAS, Trimble, Canada, stated that the developed world has long established cadastral mapping, yet major challenges are fast updates, cost saving and restriction to use of aerial images only. The developing world has poor or no cadastral maps. Still lots of challenges remain, in particular establishing a cadastral system with innovative automated approaches. CHRISTIAN HEIPKE, Leibniz University Hannover, discussed earth observation uses for quality control and the update of geospatial databases, showing examples carried out for BKG (German Federal Government) and other regions of the world (Africa, South America). Semi automatic procedures applied to images were 3 times more effective than by human operators.

The Seminar "*GIS-Ready Information from Imagery*" chaired by CHRISTIAN HEIPKE, covered some special activities: GOTTFRIED KONECNY reported on the ongoing ISPRS-UNGGIM project of a UN member country survey of existing map coverage and its updating at different scales. Contacts to international data providers (Google, Microsoft, TomTom, Navteq) have also been established with a request to them to communicate their status of data coverage and the update cycles. DIETER FRITSCH, University of Stuttgart, discussed progress in 3D city models: Several cities (Hamburg, Berlin, New York) have included LOD2 city models on Google Earth. Requirements should be for LOD3 (windows, stores, balconies) or in some cases for LOD4 (inside buildings).

In a session on "*Photogrammetry*" ALEXANDER WIECHERT of Microsoft Vexcel gave a presentation on the development of the UltraCam digital camera. By 2006, 47 cameras were sold, by 2008, 101, and by 2012, 216. The new model is the "Eagle" with 20,000 pixels per array. The software for their use was OPC in 2004, Ultramap 1 in 2011 and Ultramap 2 in 2012 for Bingmaps. Ultramap 3 is before its

Discussions at the Geospatial World Forum on Earth Observation.

release for fully automatic DSM ortho's ("true orthophotos"). The collection efficiency of the UltraCam is considerably better than lidar. Only for corridor mapping lidar has an advantage. The UltraCam is also utilized for Microsoft's Global Ortho with 30 cm GSD in the US and in Europe. Lorraine Tighe of Intermap, Canada, presented the availability of Intermap's 30 m posting global elevation model, which is based on a combination of Intermap's NextMap radar DSM/DTM coverages (North America, Western Europe and SE Asia), NASA's SRTM (DSM globally at 90 m posting), Japan's ASTER Stereo DSM at 30 m posting and NASA's IceSat altimeter at 30 m posting. A test site established by airborne lidar revealed ASTER accuracies of +/-10 m, the 90 m SRTM of +/-3 m and the World Intermap 30 m DEM at +/-3 m. NextMap with a coverage of 18 million km^2 had an accuracy of +/-1 m. While 90 m SRTM and ASTER data are free of charge, DTM's from aerial photos are in the range of 40 \$/km^2, from airborne lidar between 100 and 150 \$/km^2 and InSAR NextMap data at 30 \$/km^2. The same pricing is valid for InSAR TanDEM-X data. InSAR COSMO-SkyMed data may cost about 20 \$/km^2. Gerry Mitchell, President of PhotoSat, USA, reported a project in Eritrea, where geophysical boreholes have been targeted for imaging with optical high resolution satellites in stereo (WorldView-1 and GeoEye). The project area was 400 km^2 and the accuracy achieved was better than 50 cm.

At the closing several participant expressed their appreciation for a most interesting and successful event. Sanjay Kumar concluded the meeting with the announcement, that the next Geospatial World Forum 2013 will take place in Amsterdam again.

Gottfried Konecny, Hannover

32nd EARSeL Symposium, May 21–25, 2012, Mykonos, Greece

The European Association of Remote Sensing Laboratories EARSeL is composed of about 250 member laboratories in 39 countries (mostly European). It is a regional member of the 102 year old ISPRS. EARSeL was founded in 1976 at the Danish Technical University in Lyngby with the support of the European Space Agency ESA and the Council of Europe in Strasbourg, where the first General Assembly was held in 1977. Since then, General Assemblies were organized every year in different European countries. Starting with the 1981 meeting in Voss in Norway the General Assembly was combined with a scientific and technical symposium. The 32nd EARSeL Symposium was held in Mykonos, Greece. It was organized in the hotel Royal Mykonian, with a splendid coastal view. The local organizer was the University of Thessaly in Volos.

EARSeL Chairman 2009-2012 Rainer Reuter from the University of Oldenburg outlined in his introductory speech, that remote sensing, with the upcoming global energy demand, faces new challenges. These are in finding optimal locations for the generation of photovoltaic solar energy, but also for the optimal utilization of wind energy. Orhan Altan, the president of ISPRS, described the context into which remote sensing is embedded within the International Council of Scientific Unions ICSU. He called for joint action between geospatially oriented international societies. Maurice Brigeaud, representing ESA, outlined the European Space Agency's earth observation program and its perspective: ESA continues to operate meteorological satellites (Meteosat 1 to 7, MSG 1 to 3 and MTG) in cooperation with EUMETSAT. The Earth Watch missions, after the recent termination of Envisat's operation, will be continued by the Sentinels 1 to 3, 4/5 and 5 PC to provide GMES data in cooperation with the European Commission. The scientific objectives are to be met by the currently operating 3 Earth Explorers (SMOS, GOCE and CryoSat). Further 5 earth explorers are scheduled or planned.

Among the many other paper presentations some of the highlights were: Barbara Kosztra from FÖMI, Budapest, and George Büttner from the European Environmental Agency, Copenhagen, reported on the documentation of land cover changes in Europe by the CORINE program, in which land cover data were derived for 5.8 million km^2 in 39 countries in 3 epochs from satellite images (1988-1989, 2001-2005, 2007-2011). A fourth epoch (2013-2014) is in preparation. In con-

Opening session of the EARSeL Symposium in Greece.

Coffee break with (from left to right): Chen Jun, Karsten Jacobsen, Orhan Altan, Gottfried Konecny, Mario Hernandez, Chengli Wang.

trast to the homogenized European effort, Ursula Gessner of the German Aerospace Center DLR compared global land cover maps produced from different sources for West Africa. In the WASCAL Project data of the Global Land Cover 2000, the data of the University of Maryland 1992/93, Globe cover 2005 for Africa and MODIS satellite images of 2001 were compared; this showed many discrepancies, indicating that a regional rather than a global approach is necessary for land cover mapping. Chen Jun of the National Geomatics Center of China (NGCC) presented global land cover mapping efforts made in China for ICSU's "Future Earth" Program. NGCC uses archived USGS Landsat 1 to 8 data. 10,200 scenes are available for 2000 and 9900 scenes for 2012 covering 70% of the global land area. In China the datasets are supplemented by Chinese HJ satellite data with 30 m GSD. The aim is to generate a web based information system for global land cover by 2015. Aurelia Sand of CNES, France, showed the first Pleiades high resolution images with 0.7 m GSD in pan and 2.8 m GSD in multispectral resolution. A very interesting presentation was given by Steven Achal, ITRES, Canada. The company developed the thermal imager TABI-1800 and put it to operational use to combat forest wild fires in Northern Alberta. When flown at a height of 1000 m, a swath of 720 m can be imaged with 1800 pixels at a thermal resolution of 0.008 °C. In Northern Alberta up to 40,000 ha may be flown during one night with data delivery in the morning.

During the Council Meeting of EARSeL, a new Bureau was elected for the next 2 years with Joannis Manakos, Aristoteles University of Thessaloniki, as chairman and Lena Halounova, Technical University of Prague, as vice chair. The next EARSeL Symposium will take place from May 13 to 18, 2013, in Matera, Italy.

Gottfried Konecny, Hannover

8. GIS-Ausbildungstagung, 14.–15. Juni 2012, am GFZ Potsdam

Zum achten Mal fand in Potsdam die Veranstaltung zur GIS-Aus- und Weiterbildung statt. Mit über 70 Teilnehmern hat sich die Veranstaltung damit im deutschsprachigen Bereich als Plattform für den Erfahrungsaustausch zu aktuellen Themen wie z.B. „*GIS an Schulen*" oder neuen Methoden in der GIS-Hochschulausbildung etabliert. Auch in diesem Jahr fand zuvor eine Esri-Fortbildung mit der Überschrift „*Web-GIS in Forschung, Lehre und Verwaltung*" statt. Diese war mit 20 Teilnehmern fast schon überbucht und nach einhelliger Meinung der Teilnehmer zu kurz, um einen soliden Einblick in die vielen neuen Technologien zu bekommen.

Wie schon in den Vorjahren dominierte die Gruppe der Hochschullehrenden mit über 50%. In diesem Jahr war ein gewisser Kreis der Teilnehmer aus dem behördlichen Bereich zu verzeichnen, welches auch auf den interessanten Keynotevortrag von Prof. Hans-Jörg Kutterer (Bundesamt für Kartographie und Geodäsie, BKG) zurückzuführen war. Das interessant gestaltete Programm ließ viel Zeit für Diskussionen. Zudem sorgten, wie schon in den vergangenen Jahren, das nette Ambiente auf dem Telegrafenberg, das hervorragende Wetter und das abendliche Buffet dafür, dass eine angenehme Atmosphäre entstehen konnte. Der große Anteil von Teilnehmern, für die diese Tagung zu den gesetzten Terminen im Jahr gehört, unterstreicht diese Aussagen. Ein Fünftel der Teilnehmer war zum ersten Mal bei der Veranstaltung.

Wie auch schon in den vergangenen Jahren wurden Aus- und Weiterbildungsthemen im Umfeld der Geoinformatik fächerübergreifend diskutiert und Erfahrungen ausgetauscht, wobei sich die insgesamt 19 Vorträge um die Schwerpunkte „*Schulen*“, „*Ausbildung*“, „*Virtuelle Aus- und Weiterbildung*“ und „*Hochschulausbildung*“ gruppierten. Hans-Jörg Kutterer verdeutlichte in seinem Keynotevortrag, dass seine Behörde in den kommenden Jahren einen großen Bedarf an qualifiziertem Personal mit GIS-Kenntnissen haben wird. Ursache ist in erster Linie die gegebene Altersstruktur, die vor allem im gehobenen Dienst in den nächsten Jahren viele Stellenausschreibungen erwarten lässt. Ob diese Stellen vorwiegend mit Geodäten und GIS-Spezialisten besetzt werden können, hängt natürlich auch vom Angebot an qualifizierten Bewerbern ab. Das Thema „*GIS an Schulen*“ wurde in zwei Blöcken behandelt. Rainer Uphues stellte dabei gleich die Gretchenfrage: Warum setzen nur ca. 15% der Geographielehrer GIS ein? Dieses gilt weltweit! Anhand empirischer Daten wurden die 21 wichtigsten Hemmnisse identifiziert und gruppiert. Demzufolge sind die technische und didaktische Komplexität Haupthemmnisse bei der Implementation von GIS in Schulen. Interessanterweise wurden diese Erkenntnisse von anderen Kollegen bestätigt. Bei einer Untersuchung zur Verwendung verschiedener Programme zur GIS-Ausbildung an der Schule zeigten sich auch Schwächen beim Einsatz von Geowebbrowsern und bei der Satellitenbild-Lesekompetenz. Einen gleichwohl einfachen, wie lehrreichen Ansatz hat Hans-Jörg Stark aus Basel mit dem Projekt „see you“ vorgestellt. Dabei wurden Schülern die GIS-Fähigkeiten ihrer Smartphones vor Augen geführt. Durch kontinuierliches Tracken der Position können individuelle und gruppenbezogene Bewegungsmuster und Karten erstellt werden. Den Schülern werden dabei die Vor- und Nachteile moderner Technologie gezeigt und somit die Medienkompetenz erhöht. Dass mit großem persönlichem Einsatz auch komplexe GIS-Projekte mit Schülern durchgeführt werden können, zeigte der Beitrag von Gerhard Ortmann.

In einem Themenblock „*Ausbildung*“ lieferten Kollegen vom GFZ einen Werkstattbericht über die laufenden guten Erfahrungen mit der Geomatiker-Ausbildung. Dass vielleicht trotz der guten Ausbildung keiner der Lehrlinge direkt am GFZ als Fachangestellter bleiben wird, sondern alle studieren werden, ist eigentlich ein gutes Zeichen, dass diese Ausbildung zu mehr befähigt. Thomas Witke erläuterte angesichts des offensichtlichen Fachkräftemangels in Baden-Württemberg die vielfältigen Aktivitäten des Ministeriums

Hans-Jörg Kutterer (BKG), Jochen Schiewe (HCU Hamburg), Ralf Bill (Uni Rostock), Gerd König (TU Berlin), Matthias Möller (Beuth Hochschule Berlin), Görres Grenzdörffer (Uni Rostock) und Bernhard Harzer (Harzer Verlag Karlsruhe) bei der GIS-Ausbildungstagung in Potsdam.

für Ländlichen Raum und Verbraucherschutz und verschiedenen anderen Akteuren im GIS- und Vermessungsumfeld. Dabei geht es nicht nur um die Ausbildung oder die Hochschule, sondern auch um die Fort- und Weiterbildung sowie die Ausbilder und die Werbung an Schulen. Im Bereich „*Hochschulausbildung*“ standen in den zwei Vortragsblöcken verschiedene Themen zur Hochschulausbildung, wie die Vorstellung neuer Studiengänge, die Mühen, GIS in ein geotechnisches Curriculum zu integrieren, und der fortwährende Bolognaprozess im Vordergrund. Z.B. wurde die Gesamtsituation der Entwicklung der Studentenzahlen in der Geodäsie in dem Vortrag von Andreas Eichhorn (TU Darmstadt) beleuchtet. Dabei wurde deutlich, dass die Anzahl der Absolventen in der Geodäsie kaum ausreichen wird, um den kommenden Bedarf an Geodäten in den vielen Berufsfeldern zu decken. Gleichzeitig mahnte Eichhorn an, dass in den letzten Jahren eine fortwährende Zersplitterung der Namensgebung der geodätischen Studiengänge zu beobachten ist, die nicht unbedingt für ein eindeutiges Markenzeichen „Geodäsie“ förderlich ist. Robert Seuss (FH Frankfurt) stellte den Master „Geoinformation und Kommunaltechnik“ an der FH Frankfurt vor und betonte auch hier, dass die Chancen für die Absolventen sehr gut seien. Sven Fuhrmann von der Texas State University zeigte den Teilnehmern, wie in den USA die internationale Zusammenarbeit für die studentische Ausbildung in Form von Study Abroad Programmen organisiert ist. Dabei stellte er auch klar, dass solche Programme für die Studenten nicht gerade billig sind und ein großer organisatorischer Aufwand dahinter steht.

Die Beiträge standen, wie auch schon in den Vorjahren auf einer CD gesammelt, zur Konferenz bereit. Weitere Informationen und die Präsentationen sind in dem Programm verlinkt und unter http://gis.gfz-potsdam.de/index.php?id=245 abrufbar. Die Veranstalter setzen diese Veranstaltung mit alten und neuen Schwerpunkten im kommenden Juni 2013 fort. Der genaue Termin wird in Kürze auf der Webseite bekannt gegeben.

Görres Grenzdörffer, Rostock

WorldView Global Alliance User Conference, June 20 – 21, 2012, Munich

The WorldView Global Alliance User Conference was held June 20th and 21st 2012 in Munich, Germany, and attracted more than 150 people to participate. WorldView-1 and WorldView-2 satellite sensors belong to the most advanced sensors available at the moment, recording remote sensing imagery with a super high spatial resolution up to 0.5 m. The annual user group meeting brings together the operators and the community working with the image data.

The first day of the event started with a welcome address of Adrian Zevenbergen, head of EUSI and the announcement of the “GMES master” program. The WorldView satellites are designed for a flexible data acquisition; both platforms can be shifted and rotated into a special pointing direction for an image acquisition just in seconds. Depending on the actual cloud coverage derived from weather satellite information, e.g. Meteosat, the programming of a target area can be performed just before the image recording starts. This new feature was presented by the satellite operators and leads to less cloudy image data. Especially when combining WorldView together with QuickBird image data the results are impressive: almost cloud-free, up-to-date images can be produced out of several image takes by the merging algorithm. Day two was reserved for best practice presentations of the user community from scientific research and from the application side. WorldView-2 coastal band is of great benefit when mapping the seafloor in shallow water; bathymetric maps could be produced, differenciating between several underwater cover types. The yellow, red edge and the two infrared bands are useful for the analysis of crop yields. The land use resp. land cover of built-up areas was detected using image differenciating tools and special designed indices. The German Aerospace Center (DLR) with its subdivision Center for Crisis Information (ZKI) uses super high resolution satellite image data and smart analysis tools for almost real time detection of changes in cases of natural and man-made hazards. There is still a lack of world-wide elevation data with a reliable spatial resolution and height information, because SRTM and ASTER GDEM provide a 90 m and 30 m resolution only. Digital elevation models (DEMs) derived from stereo image data of the WorldView satellites with a 2.5 m spatial resolution have been tested for their accuracy with reliable results over the area of Istanbul, Turkey. A number of (DEMs) calculated from triple stereo WorldView image data have been used for the production of the impressive, animated 3D mountain model for Himalayan K2.

Special thanks are directed to European SpaceImaging (EUSI) staff based in Munich. EUSI organized a very balanced social frame program, an evening in typical Bavarian beer environment. The satellite challenge was related to the European Soccer Championship: a number of different European stadiums had to be located. The program of the meeting was very well in balance, profound presentations about current and future sensor design were mixed with scientific and practical application presentations of the users.

Matthias Moeller, Berlin und Salzburg, Austria.

11. Internationales 3D-Forum Lindau, 20. – 21. März 2012

Am 20. und 21. März 2012 fand in der Inselhalle Lindau das 11. Internationale 3D-Forum Lindau mit einem neuen Teilnehmerrekord von 190 Besuchern und 21 ausstellenden Firmen statt. Mit den Schwerpunktthemen *3D in Tourismus und Marketing*, *Doppik und Mobile Mapping sowie Gebäudeextraktion aus Laser- und Photogrammetriedaten* boten die Veranstalter Dipl.-Ing. Claus Bihl (Stadt Lindau) und Dr.-Ing. Achim Hellmeier (Ingenieurbüro Real.IT, Aalen) dem Fachpublikum aus Wirtschaft, Wissenschaft und Verwaltung einmal mehr ein hochinteressantes und aktuelles Programm. Neben diesen Schwerpunktthemen standen wieder Anwendungen von und mit 3D Stadtmodellen im Vordergrund der Veranstaltung.

Partner der Veranstaltung waren der Deutsche Verein für Vermessungswesen (DVW), die Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF), das Virtual Dimension Center (VDC) und die Stadtwerke Lindau. Hauptsponsor der Veranstaltung war die Firma Esri Deutschland GmbH.

Die Veranstaltung wurde wie in den Vorjahren in einen Vortragsteil am ersten Tag und in Vertiefungsthemen und Workshops am zweiten Tag aufgeteilt.

Die Begrüßung und Eröffnung am ersten Tag übernahm in Vertretung der Oberbürgermeisterin der Stadt Lindau der neue Leiter des Stadtplanungsamtes Christian Herrling. Interessant und positiv motivierend für die Teilnehmer war dabei sicherlich, wie Herrling als Architekt und Stadtplaner ausführte, dass er 3D Stadtmodelle mehr und mehr als unverzichtbares Werkzeug im städtebaulichen Planungsprozess ansieht. Anschließend folgte der Eröffnungsvortrag von Alexander Zipf vom Geographischen Institut der Universität Heidelberg mit dem Beitrag *3D Stadtmodelle im Wandel – von nice to have zum praxisnahen Werkzeug*. Zipf gab darin nicht nur einen guten Überblick über die Entwicklung von 3D Stadtmodellen, er zeigte vielmehr auf, welche Entwicklungsmöglichkeiten noch in dieser Thematik stecken. Der nächste Beitrag folgte von Nicolas Heyer von der Firma Google Deutschland. Heyer zeigte mit seiner überaus lebendigen und nicht rein technischen Präsentation auf, welche Interessen ein Unternehmen wie Google an dem Thema 3D Geo-/Gebäudedaten hat.

Im dritten Beitrag am Vormittag stellte Robert Roschlaub vom Landesamt für Vermessung und Geoinformation in München die Strategie und Vorgehensweise des Freistaates Bayern bei der landesweiten Erschließung der dritten Dimension vor. Hervorgehoben sei hier das Projekt LoD2 für ganz Bayern, welches vor kurzem begonnen wurde und 2015 abgeschlossen sein soll. Im letzten Vortrag vor der Mittagspause stellte der Architekt Jan Halatsch von der ETH Zürich ein neues 3D-Werkzeug für Stadtplaner vor. Interessant dabei war u.a., mit welcher Einfachheit damit im Stadtplanungsprozess Variantenplanungen von Straßen oder Gebäudekomplexen simuliert und visualisiert werden können.

Bevor dann die Mittagspause begann, stellten die Veranstalter die ausstellenden Firmen in Kurzporträts vor.

Den zweiten Teil des Vortragsprogramms eröffnete Michael Pospiech von der Stadt Nürnberg, in dem er über das neue 3D Stadtmodell von Nürnberg berichtete. Pospiech machte deutlich, dass nach der technischen Realisierung die Vermarktung des 3D Stadtmodells sowohl innerhalb der Stadtverwaltung als auch extern bei Architekten, Planungsbüros etc. eine nicht zu unterschätzende und aufwendige Arbeit ist. Auch der nächste Beitrag *3D-Geodaten für Tirol* befasste sich primär mit der Vermarktung von 3D Modellen im Bereich Tourismus und Standortmarketing, vorgetragen von Patrick Fritzmann von der Landesregierung Tirol und Florian Siegert (3D Reality Maps).

Im anschließenden, sehr praxisbezogenen Beitrag zeigte Joachim Ebert von der Stadt Aalen auf, wie Mobile-Mapping-Daten in der städtischen Verwaltung genutzt werden können. Besonderes Interesse bei den Teilnehmern fand dabei das Thema Doppik (Doppelte Buchführung in Konten), welches derzeit bei vielen Kommunen ein überaus aktuelles Thema ist.

Den vortraglichen Abschluss des ersten Tages machte Benjamin Sattes (Planungsbüro Z&M 3D WELT GmbH) mit dem interessanten Beitrag *Standortauswahl von Windkraftanlagen auf der Basis von 3D Geomodellen*. Diese Thematik wird aus aktuellem Anlass sicherlich auch in den folgenden Jahren immer wieder auf der Tagesordnung stehen.

Der erste Veranstaltungstag klang dann am Abend in gemütlicher Atmosphäre zur Starkbierzeit im mittelalterlichen Gasthaus Sünfzen auf der Insel Lindau aus, wo neben Speis und Trank alte Bekanntschaften gepflegt und neue Kontakte geknüpft wurden.

Am zweiten Tag standen am Vormittag drei Vertiefungsthemen auf dem Programm: Beginnend mit dem Thema *Internetvisualisierung von 3D Stadtmodellen,* wo über Stand und neueste Trends sowie über Formate und Standards wie CityGML und X3DOM referiert wurde, folgte der zweite Themenblock mit der Überschrift *Fortführung und Verdichtung von 3D Stadtmodellen*. Hier wurde u.a. vertiefend auf die Integration von Mobile-Mapping-Daten in 3D Stadtmodelle eingegangen. Den Abschluss bildete der sehr praxisnahe Beitrag *Vom Foto zur 3D-Animation*, worin den Teilnehmern ganz konkrete Werkzeuge vorgestellt wurden, die sie selbst für die Erstellung von 3D-Animationen nutzen können.

Die Vertiefungsthemen wurden von Gerhard Juen (Fachhochschule Gelsenkirchen), Volker Coors (Hochschule für Technik, Stuttgart) und Günter Pomaska (Fachhochschule Bielefeld) präsentiert.

Am Nachmittag fanden zwei Firmen- und ein CityGML-Workshop statt. Neben der Firma Esri, die ihre neuen Aktivitäten im 3D-Bereich mit der CityEngine vorstellte, war die Firma UVM Systems präsent, die u.a. über die automatische Texturierung von Gebäudemodellen berichtete.

Der CityGML Workshop mit Experten aus der SIG3D und der Standard Working Group des OGC (Open Geospatial Consortium) fand bereits zum zweiten Mal in Lindau statt und

Veranstaltungsort Inselhalle in Lindau.

Präsentation von Softwarelösungen beim 3D-Forum in Lindau.

beschäftigte sich neben neuen Modulen mit den Neuerungen des OGC Standards CityGML 1.1.

Die Veranstaltung ging gegen 17:00 Uhr zu Ende. Sie kann nicht nur wegen des neuen Teilnehmerrekordes als voller Erfolg gewertet werden, sondern auch wegen der guten Resonanz. Das nächste Internationale 3D-Forum Lindau findet am 19.–20. März 2013 statt (3d-forum.li).

Andreas Lindenmüller, Lindau
Achim Hellmeier, Aalen

Nachruf

Professor Heinz Draheim verstorben

„Eigentlich bin ich ja Photogrammeter, denn ich war schon in den 60er Jahren im Vorstand der Deutschen Gesellschaft für Photogrammmetrie….“ Das betonte Heinz Draheim bei meiner Vereidigung an der Universität Karlsruhe im Jahre 1983. Er beendete damals seinen aktiven Dienst und seine Tätigkeit als Rektor. 14 Jahre lang hatte er die „Fridericiana“ geführt, und statt „Magnifizenz“ durfte man ihn daher auch mit „seine Permanenz“ anreden oder gar mit „seine Mumifizens“.

Am 5. November 1915 in Pommern geboren, prägte ihn doch Berlin, wo er nach dem Abitur an der Technischen Hochschule Vermessungswesen studierte und 1939 mit dem Diplom abschloss. Nach Krieg und Gefangenschaft zog es ihn zurück an die TH im völlig zerstörten Berlin. Er arbeitete als wissenschaftlicher Assistent und Oberingenieur und an seiner Promotion (1952), legte auch die Große Staatsprüfung für den Höheren Technischen Verwaltungsdienst ab und habilitierte sich 1958.

Soweit vollständig ausgestattet mit solidem akademischen Rüstzeug, wurde Draheim bereits ein Jahr später als Professor an die Technische Hochschule Karlsruhe berufen mit einem Bewerbungsvortrag über Transformation ungleichartiger Koordinaten nach C. Strinz. Am Karlsruher Geodätischen Institut hatte er das große Arbeitsfeld „Ausgleichungsrechnung und Landesvermessung“ zu vertreten. Es war daher auch konsequent, dass Professor Draheim sehr bald die Schriftleitung der Zeitschrift Allgemeine Vermessungs-Nachrichten (AVN) übernahm (1962 bis 1993), zumal der Wichmann-Verlag in Karlsruhe wurzelte. Das „blaue Wunder“ pflegte er die AVN zu nennen wegen ihres blauen Einbandes und ihrer wunderbaren Inhalte und grenzte sie damit ab zur „roten Fahne“ der Zeitschrift für Vermessungswesen (ZfV), anspielend auf die bekannte Dresdner Brücke einerseits und das Organ der Kommunistischen Partei andererseits.

Die eigentliche, herausragende Phase des beruflichen Lebens von Professor Draheim begann jedoch mit seiner Wahl zum Rektor der Universität Karlsruhe im Jahre 1968. Er hat sie durch gesellschaftspolitisch sehr schwierige Zeiten navigiert und brachte es mit Präsenz und persönlichem Einsatz zustande, die divergierenden Gruppierungen vom gemeinsamen Ziel hochklassiger Lehre und Forschung zu überzeugen. Studentische Unruhen wie an anderen Universitäten gab es in Karlsruhe nicht, man sprach hier von der „Pax Draheim".

Ein erster sichtbarer Schritt zur Modernisierung der Karlsruher Hochschule bestand in der Erarbeitung und Verabschiedung einer neuen Grundordnung. Rektor Draheim gelang es „…die Mitwirkung aller Mitglieder nahezu optimal zu gestalten, ohne die Funktionsfähigkeit der Universität zu gefährden", wie es sein Weggefährte Professor H.G. Kahle zu seinem 65. Geburtstag formulierte. Die Grundordnung war eine sehr praktische, liberale Verfassung ohne hinderliche Bürokratie. Diese war Rektor Draheim so sehr verhasst, dass er, wenn er es für richtig hielt, Erlasse des Ministeriums auch einfach ignorierte. Das offene Wort und Mut im Sinne von Zivilcourage hat Draheim jederzeit ausgezeichnet. Es ist in dieser Ausprägung wohl auch ein Berliner Import, welcher allerdings in Baden bisweilen auch missverstanden werden kann.

Neben „Hochschulpolitik" – Rektor Draheim vertrat auch von 1971 bis 1977 die Universitäten im Wissenschaftsrat – war es die internationale Arbeit, die er besonders beförderte. Der Kooperationsvertrag mit der TU Budapest, abgeschlossen bereits 1970 in der Eiszeit des Kalten Krieges und noch vor Aufnahme diplomatischer Beziehungen, setzte Maßstäbe für die Überwindung des Eisernen Vorhangs, womit die Wissenschaft also der Politik voranging. Professor Draheim wurde unter anderem auch dafür von seiner Partneruniversität mit dem Ehrendoktor ausgezeichnet. Er war Präsident und später Ehrenpräsident der Fédération Internationale des Géomètres (FIG), Ehrenmitglied der Royal Institution of Chartered Surveyors und Commandeur dans l´Ordre des Palmes Académiques, um nur einige Ehrungen aus seinem internationalen Arbeitsfeld zu nennen.

Es hieße jedoch den Menschen Heinz Draheim vollkommen zu verkennen, ließe man seine Freude an geistreicher Diskussion und sein Engagement für das Musische unerwähnt: Vorstand der Gesellschaft der Freunde des Badischen Staatstheaters, Mitglied im Rundfunkrat des Süddeutschen Rundfunks, praktizierender Cellist. Beim Musizieren allerdings überließ er freiwillig seiner Frau, einer ausgebildeten Konzertpianistin, den Vortritt. Zusammen mit seinem späteren Rektor-Nachfolger Professor Heinz Kunle trat diese auch öffentlich auf unter dem gemeinsamen Namen „Kunheim". Umblätterer beim Vierhändigspiel war übrigens Heinz Draheim.

Professor Heinz Draheim verstarb am 25. April 2012 an seinem Wohnsitz bei Karlsruhe im Alter von 96 Jahren „nach einem langen, glücklichen, der Wissenschaft, der Kunst und der Musik gewidmeten Leben…", wie es in der Traueranzeige heißt.

Mit ihm verlieren wir eine große Persönlichkeit, welche die Bedeutung und die Faszination unseres geodätischen Berufsfeldes weitervermittelt hat, und einen geradlinigen, liebenswerten Menschen.

Hans-Peter Bähr, Karlsruhe

Neue Mitglieder der DGPF

Als neue persönliche Mitglieder begrüßt die DGPF

Prof. Dr. Lars Bernard, Dresden
B.Sc. Katja Dörnhöfer, Kiel
Christopher Ebert, Wulkenzin OT Neuendorf
Dipl.-Ing. Robert Eckardt, Jena
Dipl.-Geogr. Hendrik Herold, Dresden
Prof. Dr.-Ing. Stefan Hinz, Karlsruhe
PhD. Miroslav Kardoš, Zvolen, Slowakei
Dipl.-Geoökologe Stefan Kärgel, Potsdam
Nadja Käuper, Münster
M.Eng. Tomasz Kogut, Sianów, Polen
Dipl.-Geogr. Lukas Lehnert, Marburg
Dipl.-Ing. Frank Liebold, Dresden
B.Eng. Tobias Lochmann, Wesenberg
Dipl.-Geogr. Thomas Lotz, Marburg
B.Sc. Stefan Menze, Bochum
Dr.-Ing. Mehdi Ravanbakhsh, Melbourne, Australien
Dipl.-Geol. Gottfried M. Reimann, Herne
M.Sc. Sergej Rudsinski, Bochum
Benjamin Sackel, Neubrandenburg
M.Sc. Lukas Schack, Hannover
B.Eng. Martin Scherpinski, Rostock
Marcel Schwieder, Berlin
Brenner Silva, Marburg
B.A. Roman Skorzus, Erfurt
Dr. Boris Thies, Marburg
Dipl.-Ing. (FH) Michaela Trum, Unnau
Dipl.-Ing. Sebastian Tuttas, Eichenau
Dipl.-Ing. Detlev Wagner, Kirchheim
Dipl.-Ing. Patrick Westfeld, Dresden

Als neue korporative Mitglieder begrüßt die DGPF die Firmen

M.O.S.S. Computer Grafik Systeme GmbH, Taufkirchen
topometric GmbH, Göppingen
IABG mbH, Ottobrunn

Hochschulnachrichten

Universität Stuttgart, Dissertation Alexander Fietz

Herr Dipl.-Ing. Alexander Fietz promovierte am 8.3.2012 an der Fakultät für Luft- und Raumfahrttechnik und Geodäsie der Universität Stuttgart mit der Arbeit *„Ableitung von Bewegungsstrategien zur automatisierten, vollständigen Vermessung von Innenraumszenen auf autonom navigierenden Plattformen“* zum Dr.-Ing.

Referent: Prof. Dr.-Ing. habil. Dieter Fritsch, Stuttgart; Korreferent: Prof. Dr. rer. nat. habil. Paul Levi.

Kurzfassung: Die Selbstlokalisierung eines autonomen mobilen Roboters in der Umgebung ist eine seiner grundlegenden Aufgaben. Eine präzise Lokalisierung ist für viele Applikationen notwendig. Eine Voraussetzung dafür ist das Vorliegen einer genauen und vollständigen Karte. Während die durch den Roboter erzeugte Karte bei typischen Robotikanwendungen ein notwendiges Nebenprodukt darstellt, rückt diese bei einer Betrachtung von vermessungstechnischer Seite unmittelbar in den Fokus. Es stellt sich die Frage, inwiefern sich die gesammelten Raumdaten bezüglich der Genauigkeit, der Vollständigkeit und des Detaillierungsgrades für eine Umgebungskartierung eignen, die den Anforderungen eines menschlichen Nutzers genügen. Im Alltag benötigt der Nutzer exakte Modelle von Innenräumen, deren Beschaffung häufig mit großem Aufwand verbunden ist. Eine vollautomatische Generierung dieser Innenraummodelle wäre für ihn daher wünschenswert. Die Arbeit soll hierzu einen Beitrag leisten und versucht die Verfahren und Mittel zusammenzustellen, die zur Modellierung einer Innenraumszene quasi „per Knopfdruck“ notwendig sind.

Konzipiert wurde ein mobiles Messsystem, durch das sich Innenraumumgebungen vollständig und mit möglichst hoher Genauigkeit in 2D als auch in 3D einmessen lassen. Der Aufbau des Messsystems erfolgte in zwei Stufen. In einer Basisstufe wurde eine mobile Plattform mit einem low-cost Laserscanner ausgestattet, um eine 2D-Exploration von Innenraumszenen zu ermöglichen. In einer Ausbaustufe erlaubt die Anbringung einer Digitalkamera – basierend auf der Anwendung photogrammetrischer Methoden – eine zusätzliche 3D-Rekonstruktion.

In der Arbeit werden Positionierungsstrategien für die 2D als auch die 3D Vermessung vorgestellt, deren primäres Ziel eine Genauigkeitsmaximierung der resultierenden Raumdaten ist. Im ersten Teil der Arbeit wird ein 2D-Messsystem vorgestellt, das in der Lage ist, unbekannte Innenraumszenen zu erkunden und exakte Grundrisspläne von diesen zu erstellen. Dabei fährt das System iterativ Messposen an, die durch eine Positionierungsstrategie bestimmt werden. Die an den einzelnen Messpositionen aufgezeichneten 2D-Punktwolken werden über ein Scan-Matching Verfahren in einem gemeinsamen Koordinatensystem registriert. Die Positionierungsstrategie beruht auf einer globalen Betrachtung der Umwelt als Verkettung von Liniensegmenten. Da die Enden dieser Segmente auf Datenlücken hinweisen, wird über diese die Exploration einer Szene bis zu deren vollständiger Erfassung vorangetrieben. Der zweite Teil der Arbeit stellt eine Positionierungsstrategie vor, durch die sich ein Bildverband für eine photogrammetrische Rekonstruktion aufnehmen lässt. Bereits vor der eigentlichen Aufnahme werden mögliche Posenkonfigurationen über eine Genauigkeitsabschätzung eruiert. Die Annahme, sich in ebenen Umgebungen zu bewegen, ermöglicht eine Einschränkung der Wahl möglicher Kameraposen auf ein 2D Suchproblem. Ausgangsinformation der Posebestimmung ist die 2D Umgebungskarte, die sich durch das vorgestellte 2D Messsystem erstellen lässt. Aus dieser werden iterativ pseudozufällige Posenkonstellationen für definierte Umgebungsbereiche abgeleitet und anhand einer Kostenfunktion miteinander verglichen. Die Kostenfunktion versucht die im Bündelausgleich entstehenden Varianzen der Objektpunkte abzuschätzen. Dies wird durch eine Reduzierung des funktionalen Modells des Bündelausgleichs auf eine 2D Betrachtung möglich, wobei ein 2D Richtungsnetz entsteht. Form und Größe der resultierenden Fehlerellipsen lassen Rückschlüsse auf die Güte möglicher Kameraposen zu und erlauben eine vergleichende Evaluierung.

Ein wesentlicher Teil der Arbeit beschäftigt sich mit der empirischen Evaluierung der Systeme, um deren Leistungsvermögen und die Güte der resultierenden Raumdaten zu erörtern. Anhand von Versuchen in realen Umgebungen wird die praktische Anwendbarkeit der entwickelten Messverfahren belegt. Im Falle der scannenden Vermessung belegen die Experimente, dass das entwickelte Messsystem auch komplexe Innenraumszenarien einzumessen und zu explorieren vermag. Eine Betrachtung der erstellten Punktwolke zeigt, dass deren Genauigkeit vielen vermessungstechnischen Ansprüchen genügt und das Verfahren diesbezüglich herkömmlichen Messmitteln überlegen ist. Bei einer anschließenden Modellierung werden jedoch feinere Umgebungsstrukturen fälschlich abgebildet oder gehen ganz verloren. Auch die 3D-Messstrategie ist existierenden Strategien nachweislich überlegen. Das rein passive Verfahren führt jedoch zu Punktwolken, die nicht dicht genug sind, um mit entsprechender Software detaillierte Umgebungsmodelle erzeugen zu können.

Die Dissertation ist in der Reihe C der Deutschen Geodätischen Kommission bei der Bayerischen Akademie der Wissenschaften (ISBN 978-3-7696-5097-6) unter der Nr. 685 online veröffentlicht (http://dgk.badw.de/index.php?id=12). Gleichzeitig erscheint die Arbeit in elektronischer Form bei der Bibliothek der Universität Stuttgart (http://elib.uni-stuttgart.de/opus/doku/e-diss.php).

Technische Universität Clausthal, Dissertation Michael Schäfer

Herr Dipl.-Geol. Michael Schäfer wurde am 27.2.2012 an der Fakultät für Energie- und Wirtschaftswissenschaften der Technischen Universität Clausthal mit der Arbeit *„Atmosphäre als Phasenbestandteil der differentiel-*

len Radarinterferometrie und ihr Einfluss auf die Messung von Höhenänderungen“ zum Dr.-Ing. promoviert.

Referent: Prof. Dr.-Ing. Wolfgang Busch, Technische Universität Clausthal; Korreferentin: Prof. Dr. rer. nat. Christiane Schmullius, Friedrich-Schiller-Universität Jena; Vorsitzender der Promotionskommission: Prof. Dr.-Ing. Norbert Meyer, Technische Universität Clausthal.

Kurzfassung: Die differentielle Interferometrie auf Grundlage von satellitengestützten Radarfernerkundungssensoren mit synthetischer Apertur (dInSAR) hat sich für die Erfassung von Verformungen der Erdoberfläche bewährt. Das Verfahren ermöglicht prinzipiell Messgenauigkeiten im Millimeterbereich, wodurch es neben vielen geowissenschaftlichen Fragestellungen auch für geodätische Messungen interessant wird, beispielsweise zur Ermittlung bergbaubedingter Höhenänderungen. Für das Erreichen hoher Messgenauigkeiten müssen jedoch Fehlereinflüsse in den Daten bekannt sein und möglichst vermieden werden. Diese Arbeit befasst sich vornehmlich mit dem Einfluss der Erdatmosphäre auf die Ergebnisse der Radarinterferometrie.

Die Erdatmosphäre besitzt einen bedeutenden Einfluss auf die Ausbreitung elektromagnetischer Wellen, welche von satellitengestützten Radarsensoren ausgesendet werden. Dabei spielen insbesondere kleinräumige Inhomogenitäten in der Troposphäre, vor allem bedingt durch das Wetter, sowie unterschiedliche Elektronendichten in der Ionosphäre eine Rolle. In Interferogrammen äußert sich eine Atmosphärenänderung zwischen den Aufnahmezeitpunkten als zusätzlicher Phasenbestandteil in Überlagerung mit den gesuchten Phaseninformationen, welche die Deformation und Topographie der Geländeoberfläche widerspiegeln.

Vorgestellt wird ein Verfahren, welches als Novum die Ermittlung der atmosphärischen Phasenverschiebung zu einem einzelnen Zeitpunkt ermöglicht. Die Berechnung erfolgt auf der Grundlage einer Zeitreihe von Radaraufnahmen. Sämtliche Interferogrammkombinationen bezogen auf einen Aufnahmezeitpunkt beinhalten auch die Atmosphäre dieses Zeitpunkts, was eine statistische Ermittlung erlaubt. Als Besonderheit des gezeigten Verfahrens ist keine vorherige Lösung von Phasenmehrdeutigkeiten nötig. Nebenbei ermöglicht die Vorgehensweise zudem die Detektion von Deformationsgebieten sowie die Ermittlung von einzelnen Pixeln, die eine zeitliche Phasenstabilität aufweisen.

Gezeigt werden Ergebnisse basierend auf sechs Zeitreihen dreier Radarsensoren an Bord der Erdbeobachtungssatelliten TerraSAR-X, Envisat und ALOS. Diese Sensoren weisen verschiedene Radarwellenlängen im X-, C- und L-Band auf, was die Untersuchung frequenzabhängiger Einflüsse ermöglicht. Die mit dem neuen Verfahren aus Radardaten gewonnen Informationen werden mit Wetterdaten verglichen. Vor allem optische Fernerkundungsdaten von Meteosat MSG und Envisat MERIS zeigen eine gute Übereinstimmung. Auch die zukünftige Nutzung der erzeugten Ergebnisse zur räumlich hochaufgelösten Gewinnung von meteorologischen Kenngrößen der Atmosphäre ist vorstellbar. Darüber hinaus werden diverse Parameter entwickelt, welche eine Charakterisierung und den Vergleich von atmosphärischen Zuständen zu verschiedenen Zeitpunkten ermöglichen. Mit ihrer Hilfe können deutliche jahreszeitliche Unterschiede in der Ausprägung der Atmosphäre festgestellt werden, wobei Szenen im Sommer eine stärkere Beeinflussung durch Atmosphäre zeigen als im Winter. Auch Vergleiche zwischen den Sensoren werden angestellt und diskutiert, wobei vor allem die Sonneneinstrahlung je nach Tageszeit der Aufnahmen die Stärke der Atmosphäre in den Radardaten beeinflusst. Ferner bewirkt die Ionosphäre bei Radarwellenlängen ab dem Dezimeterbereich, dass Aufnahmen, z. B. von L-Band Radar auf ALOS, ein deutlich abweichendes jahreszeitliches Verhalten gegenüber den kürzeren Wellenlängen von Envisat und TerraSAR-X zeigen. Abschließend werden Betrachtungen zur theoretischen Erfassbarkeit von Höhenänderungen mittels differentieller Radarinterferometrie angestellt. Unter dem Aspekt des Auftretens von Atmosphäre in den Radardaten werden zudem die Auswirkungen von atmosphärischen Phasenverschiebungen auf verschiedene differentiell-interferometrische Auswerteverfahren behandelt, wie beispielsweise die rasterbasierte Stapelung von Interferogrammen oder die punktbasierte Persistent Scatterer Interferometrie (PSI).

Die Dissertation ist in der Schriftenreihe „Geotechnik und Markscheidewesen" des Instituts für Geotechnik und Markscheidewesen der Technischen Universität Clausthal als Heft Nr. 22 (ISBN 3-938924-15-2) erschienen. Zudem ist die Arbeit elektronisch veröffentlicht beim Gemeinsamen Bibliotheksverbund (GBV) unter http://www.gbv.de/dms/clausthal/E_DISS/2012/db110817.pdf.

Technische Universität Clausthal, Dissertation Diana Walter

Frau Dipl.-Ing. Diana Walter wurde am 28.9.2011 an der Fakultät für Energie- und Wirtschaftswissenschaften der TU Clausthal mit der Arbeit *„Systematische Einflüsse digitaler Höhenmodelle auf die Qualität radarinterferometrischer Bodenbewegungsmessungen"* zum Dr.-Ing. promoviert.

Referent: Prof. Dr.-Ing. Wolfgang Busch, TU Clausthal; Korreferent: Prof. Dr.-Ing. Uwe Sörgel, Leibniz Universität Hannover; Vorsitz der Promotionskommission: Prof. Dr.-Ing. Norbert Meyer, TU Clausthal.

Kurzfassung: Auf dem Gebiet der Bodenbewegungsüberwachung ist das Interesse an der Radarinterferometrie als Messmethode insbesondere durch die neuen hochauflösenden SAR-Satelliten stark gestiegen. Die Qualität der Interferogramme hat sich erheblich verbessert, was vor allem auf die höheren Wiederholraten, aber auch auf die geringeren Schwankungen der Satellitenorbits zurückzuführen ist. Durch die höheren Bodenauflösungen wurde auch das Anwendungsspektrum speziell im Bereich der Einzelobjektüberwachung erweitert. Anhand verschiedener Anwendungsbeispiele konnte mehrfach nachgewiesen werden, dass Messgenauigkeiten von wenigen Millimetern pro Jahr erzielbar sind. Die Anerkennung der Radarinterferometrie als Messmethode im klassisch geodätisch-markscheiderischen Sinne und die damit verbundene Überführung in die Praxis macht eine genaue und umfassende Beschreibung und Quantifizierung der Messfehler erforderlich. Die Messfehler resultieren überwiegend aus räumlichen und zeitlichen Einflussfaktoren. Systematische Fehler werden durch die Aufnahmegeometrie hervorgerufen, hauptsächlich durch den Abstand zwischen den Satellitenaufnahmepositionen. Die Erfassung von Bodenbewegungen der Erdoberfläche mittels Radarsensoren setzt wiederholte Aufnahmen voraus. Aus den gemessenen Phasen der zurückgestreuten Signale werden Interferogramme generiert und im Allgemeinen unter Verwendung von vorhandenen Höhenmodellen topographiebedingte Einflüsse korrigiert. Fehler in den Höhenmodellen werden dabei direkt, aber auch indirekt, in die Bewegungsmessung übertragen und führen zu Messfehlern. Der Einfluss wird in der praktischen Anwendung häufig unterschätzt. Speziell in bergbaulich beeinflussten Gebieten oder auch dicht bebauten städtischen Bereichen treten verstärkt topographiebedingte Messfehler auf.

Im Rahmen der Dissertation wird der systematische Einfluss von Höhenmodellfehlern innerhalb des interferometrischen Auswerteprozesses ausführlich untersucht sowie die Auswirkungen auf die Genauigkeit der abgeleiteten Bodenbewegungen quantifiziert. Neben der funktionalen Beschreibung der Zusammenhänge erfolgt exemplarisch an einem bergbaulich beeinflussten Gebiet die Analyse der vertikalen Messfehler in Abhängigkeit unterschiedlicher Höhendaten und Satellitensensoren (TerraSAR-X, Envisat ASAR, ALOS PALSAR) sowie unterschiedlicher Aufnahmegeometrien. Dabei wird differenziert auf Auswirkungen von Höhenfehlern in Bereichen mit stetigen topographischen Veränderungen sowie in Stadt- und Vegetationsgebieten eingegangen. Spezielle Untersuchungen erfolgen hinsichtlich der Genauigkeit bei der Überführung von Höhenmodelldaten ins Range-Doppler-Koordinatensystem (Geokodierung), die Voraussetzung für die topographische Korrektur ist.

Die Validierung mittels terrestrischer Daten zeigt, dass die realen Messfehler häufig kleiner sind als die theoretisch berechneten Werte. Bei X-Band Sensoren bewirken bereits kleine Höhenfehler große Phasenfehler, die jedoch im Vergleich zu L-Band Ergebnissen deutlich kleinere Messfehler verursachen. Auf

Grundlage der durchgeführten Untersuchungen werden in der Arbeit Empfehlungen gegeben, mit deren Hilfe eine Reduktion von topographiebedingten Messfehlern in Interferogrammen unter Einhaltung sensorspezifischer Aufnahmeparameter sowie Verwendung geeigneter Höhenmodelle möglich ist. Vertiefende Kenntnisse zur Identifikation von Höhenfehlern in Interferogrammen sollen Anwendern als Interpretations- und Bewertungshilfe zur quantitativen Einschätzung möglicher Messfehler dienen.

Die Dissertation ist in der Schriftreihe des Instituts für Geotechnik und Markscheidewesen der TU Clausthal als Heft Nr. 21 erschienen.

VDV-Preis (Verband Deutscher Vermessungsingenieure)

Der diesjährige VDV-Preis für die bundesweit beste Masterarbeit im Fachgebiet Geodäsie und Geoinformatik wurde am 1. Juni 2012 an der Jade Hochschule in Oldenburg verliehen. Herr M.Sc. FOLKMAR BETHMANN erhält den VDV-Preis für seine Masterarbeit „*Entwicklung einer Monte-Carlo-Simulation zur Optimierung der 6DOF-Navigation mit einem mikroskopintegrierten Einkamerasystem*“. BETHMANN hat in seiner Masterarbeit ein Simulationsprogramm zur Ermittlung der mit dem System von mikroskopintegrierten Einkamerasystemen erreichbaren Genauigkeiten entwickelt. Die Genauigkeit ist gerade in der Neurochirurgie und der HNO-Chirurgie bei beengten Verhältnisse von großer Bedeutung. Der VDV-Preis wird einmal jährlich bundesweit verliehen für je eine Diplom-, Bachelor- und Masterarbeit. Der VDV ist mit ca. 6.500 Mitgliedern die berufspolitische Vertretung der Vermessungs- und Geoinformationsingenieure in Deutschland.

BURKHARD KREUTER, VDV, Wuppertal

Buchbesprechung

MÖSER/HOFFMEISTER/MÜLLER/SCHLEMMER/STAIGER/WANNINGER: Handbuch Ingenieurgeodäsie – Grundlagen, vierte, völlig neu bearbeitete Auflage

Das vorliegende Buch hat seine Ursprünge in dem bekannten Fachbuch „Ingenieurgeodäsie“ von HENECKE/WERNER/MÜLLER aus dem Jahr 1982, dessen zweite Auflage ich als Student gern und oft in die Hand genommen habe, weil viel Wissen komprimiert und verständlich weitergegeben wurde. Man erhielt zunächst einen kurzen Abriss der jeweils behandelten ingenieurgeodätischen Fragestellung und darauf aufbauend den passenden Überblick über die zur Verfügung stehenden Instrumente und Verfahren sowie der physikalischen Grundlagen der eingesetzten Sensorik.

Nachdem in den letzten dreißig Jahren ein immenser technischer Wandel stattgefunden hat, bietet der vorliegende Band der vierten Auflage in neun Kapiteln nun diesen umfassenden Überblick in die notwendigen „Grundlagen“ der Ingenieurgeodäsie. Nach einem Überblick des Themenkomplexes „Ingenieurgeodäsie“ stellt das zweite Kapitel einen siebzehnseitigen Exkurs in das „öffentliche Vermessungswesen“ dar und gibt eine konzentrierte Übersicht über die Organisation des behördlichen Vermessungswesens in Deutschland und die baurechtlichen Grundlagen, die bei Planung und Bauvorbereitung von Interesse sind. Kapitel 3 befasst sich eingehend mit der „Auswertung von Messungen“, allgemein gesagt: von den Grundbegriffen der „Fehler-lehre und Statistik“ bis zu Koordinatentransformationen. Von den altbekannten geodätischen Fachtermini wie „Standardabweichungen“ über den „GUM“ („Guide to the Expression of Uncertainty in Measurement“) wird der Bogen zu den im Bauwesen üblichen „Toleranzen“ gespannt. Einzelne statistische Testverfahren, das „Totale Differential“ oder sogar die Kovarianzfortpflanzung für korrelierte Größen werden formelmäßig detailliert beschrieben und sind für den Leser sicherlich sofort nachvollziehbar. Das Fachgebiet „Ausgleichungsrechnung“ wird auf den fünfzig Seiten dieses Kapitels verständlicherweise nur kurz angerissen, so dass der vorliegende Band das Studium weiterer Fachliteratur nicht ersetzen kann. Vertieft wird das Wissen zur Ausgleichungsrechnung später im siebten Kapitel „Grundlagennetze“. Diese „Zweiteilung“ erfährt auch die Thematik „Bezugssysteme und Koordinatensysteme“, die zunächst im eigenständigen Kapitel 4 und später erneut in „7 Grundlagennetze“ mehr anwendungsbezogen diskutiert werden.

Kernstück dieses Bandes sind sicherlich die Kapitel „5 Grundlagen der Sensorik“ und „6 Messverfahren“ sowie die Kapitel zur „Absteckung“ (8 und 9). Im Kapitel 5 werden umfassend und verständlich die physikalischen Grundlagen der Optik sowie der Elektronik dargestellt und der Weg aufgezeigt, wie aus einer zu messenden Größe zunächst das (elektrische) Signal eines Aufnehmers und daraus wiederum der eigentliche (digitale) Messwert wird. Daran anschließend wird im sechsten Kapitel eine umfassende Auswahl ingenieurgeodätischer Instrumente, die in der Regel eine Kombination mehrerer Einzelsensoren sind, im Zusammenhang mit ihren Hauptanwendungen betrachtet. Für viele ingenieurgeodätische Problemstellungen werden typische Lösungsansätze von der Auswahl des Instrumentariums und des eigentlichen Messverfahrens mit Abschätzung des Fehlerbudgets bis zur Auswertung beschrieben, teilweise unter Angabe aller benötigten Formeln. Im achten Kapitel kommt man von den allgemeineren Verfahren zum Thema „Absteckung“, das unter „9 Bogenabsteckung“ nochmal spezialisiert wird.

Bei einer solch umfassenden Darstellung der Sensorik, des Instrumentariums und der Verfahren kann man sich sicherlich vortrefflich darüber streiten, ob einzelne Themenpunkte in das eine oder lieber in das andere Kapitel gehört hätten. Beispielsweise wird der „unbedarfte“ Leser das klassische Messband wohl zunächst vergeblich unter den physikalischen Grundlagen bzw. den Messverfahren suchen, bis er es unter der Thematik Absteckung (an zwei Stellen) prägnant behandelt findet. Auch wenn das sehr gute Stichwortverzeichnis hervorragende Dienste leistet, sollte man nicht einzelne Begriffe lexikonartig nachschlagen, sondern sich die Muße nehmen, das Buch als Ganzes zu lesen. Denn das ist es auch wert.

Das Buch wird dem Anspruch gerecht, eine systematische Darstellung der Ingenieurgeodäsie sowohl für Studierende als auch für praktizierende Ingenieure der Geodäsie und anderer Fachdisziplinen zu geben. Auch diesen Band werde ich, wie schon seinen Ursprungsband von 1982, sicherlich immer wieder gern in die Hand nehmen und kann ihn der angesprochenen Personengruppe unumwunden empfehlen.

Karl Foppe, Neubrandenburg

Veranstaltungskalender

2012

4.–5. Oktober: 1. Gemeinsames Arbeitskreis-Treffen des **AK „Fernerkundung" der DGfG** und des **AK „Interpretation von Fernerkundungsdaten" der DGPF** in **Bochum**. ak-fernerkundung.de

7.–13. Oktober: **ECCV 2012**, European Conference on Computer Vision, in **Florenz**, Italien. eccv2012.unifi.it

9.–11. Oktober: **INTERGEO 2012** und **60. Deutscher Kartographentag** und **3. Europäischer Kongress der CLGE** (Comité de Liaison des Géomètres Européens) in **Hannover**. intergeo.de

16.–18. Oktober: 9th **Symposium on Location-Based Services** in **München**. www.lbs2012.tum.de

6.–7. Dezember: **3D-NordOst 2012**, 15. Anwendungsbezogener Workshop zur Erfassung, Modellierung, Verarbeitung und Auswertung von 3D-Daten am GFaI in Berlin-Adlershof. 3d-nordost.de.

2013

13.–14. Februar: 12. **Oldenburger** 3D-Tage. jade-hs.de/3dtage

17.–23. Februar: 17. **Internationale Geodätische Woche** in **Obergurgl**, Österreich. uibk.ac.at/vermessung/veranstaltung/obergurgl.html

30. April – 2. Mai: 8th **International Symposium on Mobile Mapping Technology** 2013 in **Tainan**, Taiwan. conf.ncku.edu.tw/mmt2013/

25.–30. August: 26th **International Cartographic Conference (ICC)** in **Dresden**. icc2013.org/

2.–6. September: XXIVth **CIPA Heritage Documentation Symposium** in **Straßburg**, Frankreich. cipa.icomos.org

9.–13. September: 54. **Photogrammetrische Woche** in **Stuttgart**. www.ifp.uni-stuttgart.de/phowo

8.–15. Dezember: **ICCV 2013**, International Conference on Computer Vision, **Sydney**, Australien. iccv2013.org

Korporative Mitglieder

Firmen

AEROWEST GmbH
AICON 3D Systems GmbH
aphos Leipzig AG
Becker GeoInfo GmbH
Bernhard Harzer Verlag GmbH
Blom Deutschland GmbH
Brockmann Consult GmbH
bsf swissphoto GmbH
Büro Immekus
CGI Systems GmbH
con terra GmbH
DB Netz AG
DELPHI IMM GmbH
Deutsches Bergbau-Museum
EFTAS Fernerkundung Technologietransfer GmbH
ESG Elektroniksystem- und Logistik-GmbH
Esri Deutschland GmbH
EUROPEAN SPACE IMAGING
Eurosense GmbH
fokus GmbH
g.on experience gmbh
GAF GmbH
GeoCart Herten GmbH
GeoContent GmbH
Geoinform. & Photogr. Engin. Dr. Kruck & Co. GbR
geoplana Ingenieurgesellschaft mbH
GEOSYSTEMS GmbH
GGS - Büro für Geotechnik, Geoinformatik, Service
Hansa Luftbild AG
IGI - Ingenieur-Gesellschaft für Interfaces mbH
ILV Ingenieurbüro für Luftbildauswertung und Vermessung
Imetric 3D GmbH
Infoterra GmbH
INVERS - Industrievermessung & Systeme
ITT Visual Information Solutions Germany
J. Linsinger ZT-GmbH
Jena-Optronik GmbH
Leica Geosystems GmbH
Luftbilddatenbank-Würzburg
Messbildstelle GmbH
Microsoft Photogrammetry
MILAN Geoservice GmbH
M.O.S.S. Computer Grafik Systeme GmbH
PHOENICS GmbH
PMS - Photo Mess Systeme AG
RIEGL Laser Measurement Systems GmbH
RWE Power AG, Geobasisdaten/Markscheidewesen
technet GmbH
TRIGIS Vermessung + Geoinformatik GmbH
Trimble Germany GmbH
trimetric 3D Service GmbH
Wichmann, VDE Verlag GmbH
Z/I Imaging Ltd.

Behörden

Amt für Geoinformationswesen der Bundeswehr
Bayerische Landesanstalt für Wald und Forstwirtschaft
Bundesamt für Kartographie und Geodäsie
Bundesmininisterium für Ernährung, Landwirtschaft und Verbraucherschutz
Hessisches LA für Bodenmanagement und Geoinformation
Innenministerium NRW, Gruppe Vermessungswesen
Institut für Umwelt- und Zukunftsforschung
LA für Geoinformation und Landentwicklung, BW
LA für Vermessung und Geoinformation, Bayern
LB Geoinformation und Vermessung, Hamburg
LB für Küstenschutz, Nationalpark und Meeresschutz, SH
Landesvermessung und Geobasisinformation Niedersachsen
Märkischer Kreis, Vermessungs- und Katasteramt
Regierungspräsident Tübingen, Abt. 8 Forstdirektion
Regionalverband Ruhr
Staatsbetrieb Sachsenforst Pirna
Stadt Bocholt, Fachbereich 31
Stadt Düsseldorf, Vermessungs- und Katasteramt
Stadt Köln, Amt für Liegenschaften, Vermessung und Kataster
Stadt Wuppertal, Vermessung, Katasteramt und Geodaten
Thüringer LA für Vermessung und Geoinformation

Hochschulen

BTU Cottbus, Lehrstuhl für Vermessungskunde
FH Frankfurt a.M., FB 1, Studiengang Geoinformation
FH Mainz, Institut für Raumbezogene Informations- und Messtechnik
Jade Hochschule, Institut für Angewandte Photogrammetrie und Geoinformatik
HCU HafenCity Universität Hamburg, Geomatik
HfT Stuttgart, Vermessung und Geoinformatik
HS Bochum, FB Vermessung und Geoinformatik
HS Karlsruhe, Fakultät für Geomatik
HTW Dresden, FB Vermessungswesen/Kartographie
LUH Hannover, Institut für Kartographie und Geoinformatik
LUH Hannover, Institut für Photogrammetrie und Geoinformation
MLU Halle, FG Geofernerkundung
Ruhr-Uni Bochum, Geographisches Institut
RWTH Aachen, Geodätisches Institut
TU Bergak. Freiberg, Institut für Markscheidewesen und Geodäsie
TU Berlin, Computer Vision & Remote Sensing
TU Berlin, Institut für Geodäsie und Geoinformationstechnik
TU Braunschweig, Institut für Geodäsie und Photogr.
TU Clausthal, Institut für Geotechnik und Markscheidewesen
TU Darmstadt, Institut für Photogrammetrie und Kartographie
TU Dresden, Institut für Photogrammetrie und Fernerkundung
TU München, FG Photogrammetrie und Fernerkundung
TU Wien, Institut für Photogrammetrie und Fernerkundung
Uni Bonn, Institut für Photogrammetrie
Uni Göttingen, Institut für Waldinventur und Waldwachstum
Uni Heidelberg, IWR Interdisziplinäres Zentrum für Wissenschaftliches Rechnen
Uni Kassel, FB Ökologische Agrarwissenschaften
Uni Kiel, Geographisches Institut
Uni Stuttgart, Institut für Photogrammetrie
Uni Würzburg, Geographisches Institut
Uni zu Köln, Geographisches Institut