

PRG

Photogrammetrie Fernerkundung Geoinformation

Journal for Photogrammetry, Remote Sensing
and Geoinformation Science

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

Jahrgang 2013, Heft 2

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

Redaktionsbeirat (Editorial Board): Clement Atzberger, Andrew Frank,
Christian Heipke, Joachim Hill, Patrick Hostert, Hans-Gerd Maas, Wolfgang
Reinhardt, Camillo Ressel, Jochen Schiewe



E. Schweizerbart'sche Verlagsbuchhandlung
(Nägele u. Obermiller) Stuttgart 2013



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

DAGM

Deutsche Arbeits-
gemeinschaft für
Mustererkennung e.V.



Herausgeber:

© 2013 Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF) e.V.
Präsident: Prof. Dr. Thomas Kolbe, Technische Universität München, Institut für Geodäsie, GIS und Landmanagement, Lehrstuhl für Geoinformatik, Arcisstraße 21, 80333 München, Germany, Tel. +49-89-289-23888
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

All rights reserved including translation into foreign languages. This journal or parts thereof may not be reproduced in any form without permission from the publishers.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in dieser Zeitschrift berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Verantwortlich für den Inhalt der Beiträge sind die Autoren.

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: € 229,- 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 2013, Heft 2 Inhaltsverzeichnis

Originalbeiträge

SCHMIDT, A., ROTTENSTEINER, F. & SÖRGEL, U.: Water-Land-Classification in Coastal Areas with Full Waveform Lidar Data	71
MENZE, M., KLINGER, T., MUHLE, D., METZLER, J. & HEIPKE, C.: A Stereoscopic Approach for the Association of People Tracks in Video Surveillance Systems	83
FITZNER, D., SESTER, M., HABERLANDT, U. & RABIEI, E.: Rainfall Estimation with a Geosensor Network of Cars – Theoretical Considerations and First Results	93
TOSCANI, P., IMMITZER, M. & ATZBERGER, C.: Texturanalyse mittels diskreter Wavelet Transformation für die objektbasierte Klassifikation von Orthophotos	105

Mitteilungen

Berichte von Veranstaltungen	
ISPRS-Kommission VIII, 22. ISPRS Kongress in Melbourne, Australien	123
Racurs 12th International Scientific and Technical Conference, 24.–27. September 2012, Portimão, Portugal.	124
Hochschulnachrichten	
Technische Universität Wien, Dissertation Marcela Doubková	125
Technische Universität Dresden, Habilitation Tobias Bolch	126
Deutsche Geodätische Kommission (DGK): 60 Jahre	127
Persönliches	
Nachruf auf Enrico Clerici	128
Veranstaltungskalender	129
Korporative Mitglieder	130

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



Water-Land-Classification in Coastal Areas with Full Waveform Lidar Data

ALENA SCHMIDT, FRANZ ROTTENSTEINER & UWE SÖRDEL, Hannover

Keywords: lidar, classification, coast, water, conditional random fields

Summary: In this paper, we investigate full waveform lidar data acquired over the German Wadden Sea areas in the south eastern part of the North Sea. We focus especially on classification of the 3D point clouds with the aim to determine water-land-boundaries. This is a first step towards digital terrain model generation in order to analyse the terrain topography in coastal areas and, by comparing different epochs, its dynamics. For the classification of the lidar points, we learn typical class features in a training step and combine local descriptors with context information in a conditional random fields (CRF) framework, a probabilistic supervised classification approach capable of modelling contextual knowledge. We compare the results with those obtained by a fuzzy logic based approach developed specifically for the water-land-classification in Wadden Sea areas. With the latter approach we achieve a correctness rate of more than 82% for water detection. By integrating context, the results can be significantly improved by approximately 10%. Moreover, we investigate the waveform features of the data which reveals unexpected non-linear effects concerning the decomposition of the waveforms.

Zusammenfassung: *Klassifizierung der Küstenlinie in Wattgebieten mit Full-Waveform Lidar-Daten.* In diesem Paper werden Full-Waveform Lidar-Daten in Wattgebieten des südöstlichen Teils der Nordsee untersucht. Zielsetzung ist dabei die Klassifikation der 3D Punktwolke, um Land-Wasser-Grenzen abzuleiten. Dies stellt den ersten Schritt hinsichtlich der Generierung von Digitalen Geländemodellen dar. Hiermit lässt sich die Topographie der Küstenbereiche und im Vergleich unterschiedlicher Epochen deren Dynamik analysieren. Für die Klassifikation der Laserdaten wird die statistische Verteilung typischer Merkmale in einem Trainingsschritt erlernt und lokale Ausprägungen mit Kontextinformation in einem auf Conditional Random Fields beruhenden Ansatz kombiniert. Die Ergebnisse werden mit denen eines speziell für Wattgebiete entwickelten Verfahrens verglichen, welchem die Methodik der Fuzzy Logik zugrunde liegt. Während dieses Verfahren Korrektheitsraten von über 82% für Wasserflächen aufweist, kann mit unserem kontextbasierten Verfahren der Wert um etwa 10% gesteigert werden. Darüber hinaus werden Merkmale der Waveform der Signale untersucht. Hierbei lassen sich unerwartete nicht-lineare Effekte bei der Auswertung der Signalform beobachten.

1 Introduction

1.1 Motivation and Goals

In coastal areas morphological changes are caused by tidal flows, storms, climate change, and human activities. A recurrent monitoring becomes necessary in order to detect undesired changes at early stages, enabling rapid countermeasures to mitigate or minimize potential harm or hazard. The morphology of the terrain can be represented by digital terrain models (DTMs). As the terrain is very

flat, a high accuracy of the DTM is required for tasks such as hydrographic modelling. Airborne lidar (light detection and ranging) has become a standard method for DTM generation in coastal zones. The lidar technique has two main advantages compared to traditional aerial photogrammetry: Firstly, the active laser technique works independently from illumination from the sun, which allows mapping also during night-time. Secondly, the elevation model can be directly inferred from the two-way time-of-flight of the pulse reflected at the ground, whereas stereo techniques rely

on matching of corresponding points in two or more image, which requires sufficient texture. These are the reasons why we use lidar data for our application in coastal regions: the monitoring of Wadden Sea areas.

The German Wadden Sea is a unique habitat in the south eastern part of the North Sea. In 2009 it was inscribed on UNESCO's World Heritage List together with the Dutch part. Due to its biological diversity, a monitoring of the Wadden Sea becomes necessary. A corner stone of such monitoring is the detection of any changes of the terrain geometry. Gap-less DTM modelling usually requires a combination of height data gathered by bathymetry, e. g. ship-based echo sounding, in the sublittoral zone and airborne lidar systems in the eulittoral zone. However, even during low tide, residual water remains in some tidal trenches in the eulittoral zone. Because the near-infrared laser pulses used by standard lidar devices cannot penetrate water, the measured elevation represents the water surface instead of the actual terrain level underneath as would be desired. The generation of a DTM thus requires the detection of water surfaces, which leads to a classification of the lidar point cloud into *land* and *water* areas. If such a classification has been carried out, an additional data source, e. g. sonar, could be used to complete the DTM in the water areas. In the future the problem of wrong height values over water areas could be overcome to some extent by laser bathymetry. Such modern devices operate with a green laser signal that is capable to penetrate the water column (e. g. STEINBACHER et al. 2012). However, since the accessible depth underneath the water surface depends on turbidity, such technique is better suited for clearer waters compared to the Wadden Sea.

In this paper we examine two important aspects of lidar processing in coastal areas. Firstly, we investigate the classification of three-dimensional (3D) point clouds with the aim of extracting water-land-boundaries. For this purpose we implement a supervised classification method based on context: conditional random fields (CRF). We compare the results of our approach with the results obtained by the fuzzy-logic-based method developed by BRZANK et al. (2008), which was investigated especially for the classification of Wad-

den Sea areas. We also investigate the application of full waveform information for the detection of water areas and introduce the echo width as classification feature. Further characteristics derived from the full waveform such as multiple reflections are not considered because we assume to benefit not from them in mudflat areas. Secondly, we illustrate an unexpected effect concerning the signal width in our data. In nadir view, specular reflection from the water surface should cause high amplitude but minimal width of the echo. However, we frequently observe abnormally wide echoes instead. Therefore, we chose not to apply any standard calibration scheme, but to correct the range dependency only.

This paper is organised as follows. In section 2 we discuss related approaches that deal with the classification of lidar data for water detection, the labelling of point clouds by CRFs, and the correction of intensity values of the incoming signal. We describe our classification algorithm in section 3. Section 4 contains some results of our approach including a quantitative evaluation and comparison. We conclude the paper in section 5.

1.2 Dataset

Our test site is located in the south of the island of Spiekeroog in the German Wadden Sea (Fig. 1). It contains several tidal channels of various sizes. The data were acquired by a RIEGL LMS-Q560 lidar system on 19–20 February 2011 at low tide. Over the test site of 0.4 km × 1.2 km, approximately 1.7 million points were acquired with an average point density of about 3.5 points/m². The data were processed by a surveying company with the manufacturer's software. We have no direct access to the original waveform, but only to the following parameters estimated for each detected echo pulse: the 3D coordinates as well as its estimated *amplitude* and *width*. Fig. 2 shows the widths of the laser echoes of the entire flight strip of the test area. In nadir regions, the values are surprisingly high, which corresponds to a large spread of the echo *width*. This observation contradicts the common physical model that the echo *width* increases with the scan angle in flat areas of

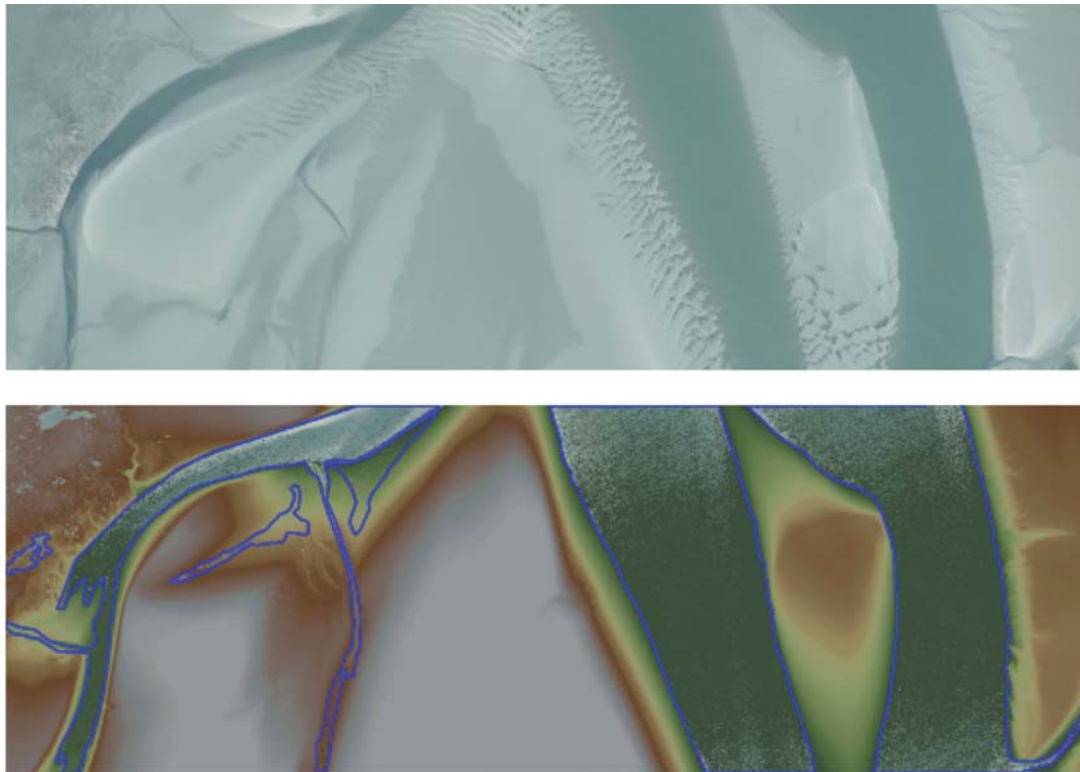


Fig. 1: Orthoimage and lidar point cloud (height span from -1.6 m (green) to 0.4 m (white)) of the test area [0.4 km \times 1.2 km] in the Wadden Sea. The areas outlined in blue are covered by water.

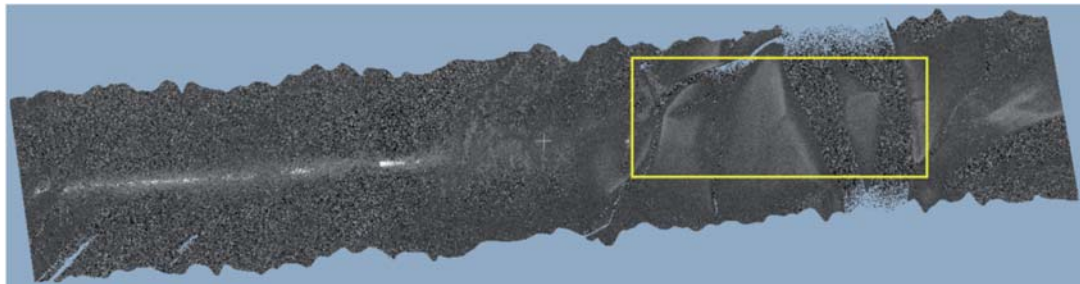


Fig. 2: Distribution of the signal width in the flight strip (increasing from dark to light). The test site for the classification is outlined in yellow.

homogenous land cover. It seems that this behaviour is connected to very large amplitude values. Therefore, we assume that signal saturation in the receive path of the sensor device might cause undesired effects. Such a non-linearity would violate the underlying assumptions of the Gaussian decomposition, i. e., saying that the echo waveform is the result of the convolution of a Gaussian pulse with a sequence of point-like reflectors modelled by their differential laser cross section.

2 Related Work

Whereas there are many approaches dealing with the classification of lidar data for the detection of objects such as buildings or vegetation, there are only a few studies on the classification of water surfaces, in particular in Wadden Sea areas. One exception is BRZANK et al. (2008), who present a point-based classification scheme as a first step towards DTM generation in the Wadden Sea. This method is

the baseline for a comparison of our CRF approach. It is a supervised classification technique based on fuzzy logic, where a membership value for the class *water* is determined for each laser point according to the features *height*, *intensity*, and *point density*. The classification into *water* and *land* is performed using a threshold for membership. All parameters of the method are derived automatically from training areas. The method takes into account the influence of height for the separation of water and land points, especially in the areas of transition. HÖFLE et al. (2009) present a segmentation-based method for the detection of water surface detection of rivers. Here, the point cloud is segmented and classified based on different features derived from the heights and the intensities of the points. Finally, the water-land-boundaries are defined by the segment borders. Classification errors occur in the case of multiple reflections, because in general they show lower intensity values for the last echo compared to the first echo, so that low signal intensity is not always caused by water surfaces. BROCKMANN & MANDLBURGER (2001) use a digital surface model (DSM) and a digital model of the water surface (DWM) for the determination of water-land-boundaries. In their method, the DWM is derived by averaging representative river heights from the laser data with regard to the known coordinates of the river axis. The water-land-boundary corresponds to the line of zero height after subtracting the DWM from the DSM. Apart from the detection of water areas, methods for the extraction of water boundaries from lidar data, e. g. river borders, have been developed. An example in the field of Wadden Sea is presented by MASON et al. (2006). They develop a method to extract tidal channel networks from lidar data. Different edge detection operators are applied to a DTM grid. This is followed by a strip detection where two edges from each side of the tidal channel are associated together. Additional high level processing improves the network, e. g. by joining the centrelines, and expanding the channels. By processing lidar data in this way, tidal channel networks are well detected in not markedly complex test areas.

To our knowledge no approach considering context in the water-land-classification pro-

cess exists. In this way, a lidar point can be assigned based on its features as well as on those obtained for all points in a defined neighbourhood. A popular context-based approach is provided by the conditional random field (CRF) framework. For image labelling the use of CRFs was introduced by KUMAR & HEBERT (2006). In comparison to image data, the labelling of point clouds is even more challenging due to the irregular distribution of points in 3D space. Several approaches for the classification of point clouds based on CRFs have been developed in the past. Some of them rely on point cloud segments. For instance, LIM & SUTER (2009) propose a method for the classification of terrestrial laser scanning data. First, they reduce the amount of data by over-segmenting the point-cloud into regions called super-voxels. Resting upon features measured by the scanner system (intensity and colour) as well as features extracted from the points inside the super-voxels, the data are labelled in a CRF framework. The potential of CRFs for airborne laser scanning data was shown by SHAPOVALOV et al. (2010). They propose a method based on segments of points and show the improvement of this non-associative approach in comparison to an associative network for an urban dataset. NIEMEYER et al. (2011) propose a point-wise method for the classification of lidar data, distinguishing three urban object classes. They also compare the performance with respect to a support vector machine, highlighting the improved classification performance of the context-based classifier.

For the labelling of lidar data, features derived from the waveform of the received pulses are of high interest for enhancing land cover classification. However, especially for the amplitude, the true physical meaning is often vague because the manufacturers seem to measure different entities. In addition, often the terms amplitude, intensity, and energy are mixed-up or used interchangeably. Nevertheless, the data should be calibrated in order to compensate for systematic effects and to achieve comparable results (WAGNER et al. 2008). The correction of the intensity eliminates the influence of sensor settings, atmosphere, and the distance to the illuminated surface (HÖFLE & PFEIFER 2007). One approach for Riegl data proposed by JUTZI &

GROSS (2010) is deriving the so-called intensity from the given amplitude and width. In addition, they consider effects depending on the range and on the atmosphere. Assuming a Lambertian reflection model, the intensity can be normalized by the incidence angle, which is the angle between beam direction and the surface normal vector. Since we observe an unexpected behaviour of the signal width in our coastal dataset, we correct the range dependency only instead of applying any standard calibration scheme.

3 Methods

In our approach for the water-land-classification the range dependency of the measured *amplitude* is corrected in a pre-processing step (section 3.1). Then, features of the classification are derived from the point cloud (section 3.2). The point cloud is classified in a supervised approach based on CRFs (section 3.3).

3.1 Pre-processing of the Data

In coastal areas two particularities of the signal must be considered during processing. Firstly, water leads to specular reflection and must not be modelled as a Lambertian surface. Thus, a correction model for the signal intensity following Lambert's cosine law is

prohibitive for water-covered areas and, as we have no a priori information about the land surface type before the classification, for the Wadden Sea areas in general. Secondly, in the course of our investigations we observed an unexpected behaviour of the involving feature *width* in our coastal dataset (section 1.2). Therefore, we decided not to apply any standard calibration scheme embedding this feature, e. g. based on JUTZI & GROSS (2010), WAGNER (2010), and LEHNER & BRIESE (2010). Instead, we use a simplified model to consider the dependence of the amplitude values from the range. Since the laser footprint is small compared to the flat terrain, no signal loss occurs in the path from sensor to object (neglecting the atmosphere). Nevertheless, on the way back only a small fraction of the backscatter is collected by the sensor. This loss can be modelled to be proportional to the squared range. Hence, we correct the amplitude a considering the range R and the distance in nadir R_n (a constant for flat areas). Since we deal here with flat terrain, this dependency can be equivalently expressed using the scan angle α :

$$a_R = a \cdot \left(\frac{R}{R_n} \right)^2 = \frac{a}{\cos^2 \alpha} \quad (1)$$

where a_R is the amplitude received over the range R .

In this way, the commonly low amplitude values for water-covered areas with high inci-

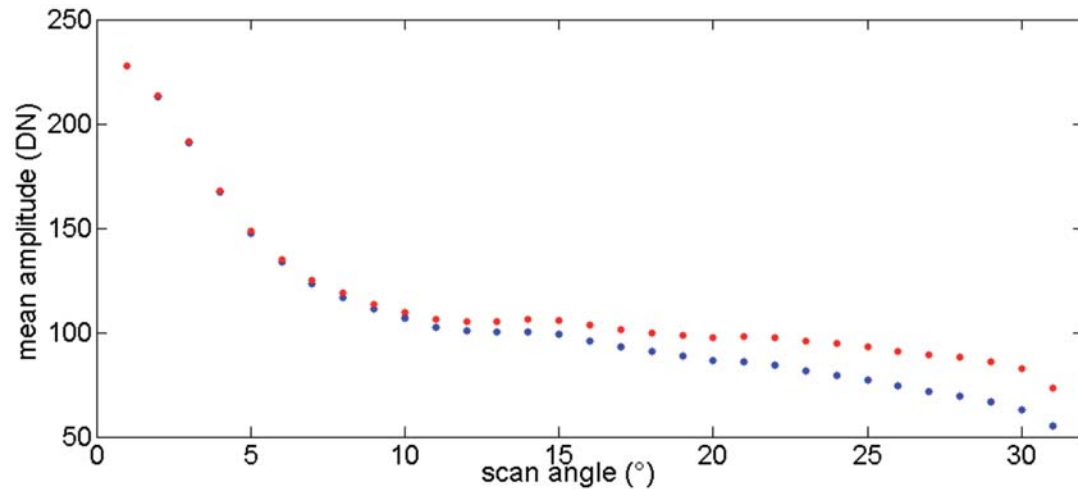


Fig. 3: Plot of mean amplitude values given in digital number (DN) over scan angle: before (blue) and after (red) the correction of the range dependency in the flight strip.

dence angles (and high difference between R and R_n) are increased. Fig. 3 shows the mean amplitude values given in digital numbers in intervals of 1° of the scan angle before and after the correction. Given the flight and sensor parameters of our dataset, these results in their maximum increase by about one third.

3.2 Feature Extraction

For the land-water-classification we use five features: *height*, *amplitude*, *point density*, *signal width*, and *variation of the width*. The first three features have been found to be well-suited for separating water from dry mud areas in the Wadden Sea (BRZANK et al. 2008). The algorithm relies on the hypothesis that water surfaces are lower than land surfaces. This assumption is suitable for water bodies at the same height level, which is usually the case in coastal areas. For laser pulses in the near-infrared, which is the part of the spectrum in which most lidar systems operate, water has a strong attenuation coefficient. This results in lower intensity values for reflections on water surface types (WOLFE & ZISSIS 1993), which can be also observed in our dataset. Fig. 4 shows the distribution of the amplitude values as a function of the scan angle. It can be seen that, firstly, the amplitude of water-covered areas is lower compared to land surfaces, and that the mean amplitude decreases with increasing scan angles, as discussed in

section 3.1. However, the high portion of specular reflection on calm water surfaces is another characteristic already discussed above. Thus, relatively high intensity values can be observed for points near the nadir direction (HÖFLE et al. 2009). The point density is defined as the number of backscattered signals per area. We define a vertical cylinder with the radius r to find adjacent points which are then used to calculate local features. Several radii were tested; however, the impact of r on the resulting features was marginal. We finally set $r = 3$ m. Especially on water surfaces, specular reflections dependent on the incidence angle can cause a significant decrease of the *point density*. The additional full waveform features which influence the investigated detection of the water-land-boundaries include the *signal width* and its *variation* with a standard deviation in a local 2D neighbourhood of lower than $r = 3$ m.

3.3 Classification by Conditional Random Fields

CRFs are a flexible technique for the classification of any kind of 2D or 3D spatial data. Thereby, a class label C_i from a given set of classes \mathbb{C} is assigned to each data site, i. e. pixel or point $i \in [1, \dots, n]$. In our case the sites are the lidar points, which are classified by finding the optimal label configuration that maximises the posterior probability $P(\mathbb{C}|\mathbf{x})$ of

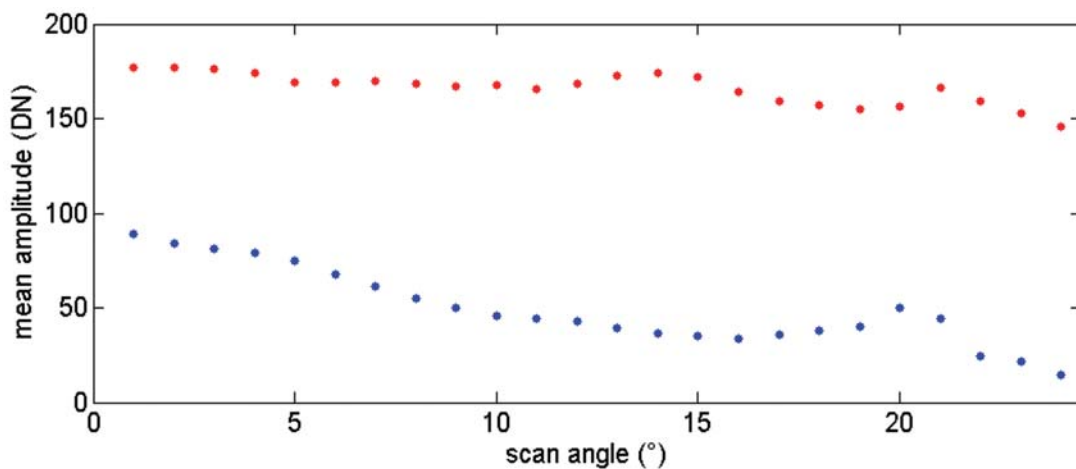


Fig. 4: Average of amplitude values given in digital number (DN) for water (blue) and land (red) surfaces in the test area as a function of the scan angle.

the point labels $\mathbf{C} = [C_1, C_2, \dots, C_n]$ given the observed data $\mathbf{x} = [x_1, x_2, \dots, x_n]$. CRFs belong to the group of graphical models whose nodes correspond to the points and whose edges model the dependencies between labels and/or data of adjacent points. The posterior probability can be modelled by

$$P(\mathbf{C} | \mathbf{x}) = \frac{1}{Z} \exp(E(\mathbf{x}, \mathbf{C})), \quad (2)$$

where $E(\mathbf{x}, \mathbf{C})$ is an energy term and Z is a normalising constant. The energy term can be expressed as the sum of *association potentials* $A(\mathbf{x}, C_i)$ and *interaction potentials* $I(\mathbf{x}, C_i, C_j)$ over the neighbourhood N_i and the dataset S :

$$E(\mathbf{x}, \mathbf{C}) = \sum_{i \in S} A(\mathbf{x}, C_i) + \sum_{i \in S} \sum_{j \in N_i} I(\mathbf{x}, C_i, C_j). \quad (3)$$

The *association potential* $A(\mathbf{x}, C_i)$ indicates the likelihood of a point i belonging to a class C_i given the observations \mathbf{x} . The *interaction potential* $I(\mathbf{x}, C_i, C_j)$ measures how the classes of neighbouring points and data interact. The structure of the graph, the observed features, and the potentials have to be defined for the application.

In our approach, we have the peculiarity of an irregular data structure in the raw laser scanning point cloud. Here, each point is linked to its k nearest neighbours in 2D. A fast access to the nearest neighbours of each lidar point is obtained by indexing the point cloud using a k -d tree. Although we apply 3D data, the reduction to a 2D search is justified by the appearance of the data. In the Wadden Sea there are hardly any objects with a significant extension in height.

A feature vector $\mathbf{h}_i(\mathbf{x})$ which contains the features described in section 3.1 is assigned to each node i . In order to consider context, interaction features modelling the relationship of nodes are introduced for each edge linking the node i and j . Here, we calculated the interaction feature vector $\boldsymbol{\mu}_{ij}(\mathbf{x})$ as the difference of feature vectors of neighbouring nodes i and j

$$\boldsymbol{\mu}_{ij}(\mathbf{x}) = \mathbf{h}_i(\mathbf{x}) - \mathbf{h}_j(\mathbf{x}). \quad (4)$$

The node features $\mathbf{h}_i(\mathbf{x})$ of the class label C_i at site i are linked by the *association potential*

$A(\mathbf{x}, C_i)$. In general, any local discriminative classifier resulting in a probability $P(C_i | \mathbf{h}_i(\mathbf{x}))$ can be used to define the association potential via $A(\mathbf{x}, C_i) = \log P(C_i | \mathbf{h}_i(\mathbf{x}))$. Closely related to KUMAR & HEBERT (2006) we use a generalized linear model for that purpose. Then, $A(\mathbf{x}, C_i)$ can be expressed as

$$A(\mathbf{x}, C_i = l) = \mathbf{w}_l^T \mathbf{h}_i(\mathbf{x}). \quad (5)$$

In (5) vector \mathbf{w}_l contains the weights of node features and is determined by a training step. Such a vector is defined for each class l . The probability that a pair of adjacent nodes i and j has the labels C_i and C_j is described by the *interaction potential* $I(\mathbf{x}, C_i, C_j)$. Analogous to the *association potential* $A(\mathbf{x}, C_i)$ it can be modelled being proportional to $\log P(C_i, C_j | \boldsymbol{\mu}_{ij}(\mathbf{x}))$, obtained again by a generalized linear model:

$$I(\mathbf{x}, C_i = l, C_j = k) = \mathbf{v}_{l,k}^T \boldsymbol{\mu}_{ij}(\mathbf{x}), \quad (6)$$

where $\mathbf{v}_{l,k}$ is the weight vector of the interaction features. Such a vector $\mathbf{v}_{l,k}$ exists for each combination of classes (l, k) .

In the training process the optimal values for the weight vectors are derived from training data. The use of exact probabilistic methods for this is computationally intractable. Thus, they are replaced by approximate solutions. Here, we applied the gradient descent optimisation method of L-BFGS (limited memory Broyden-Fletcher-Goldfarb-Shanno) (LIU & NOCEDAL 1989) for the minimisation of the objective function $f = -\log [P(\theta | \mathbf{x}, \mathbf{C})]$, where θ contains the weight vectors \mathbf{w}_l and $\mathbf{v}_{l,k}$.

The optimal label configuration is determined in an inference step. Thereby, $P(\mathbf{C} | \mathbf{x})$ is maximised for given parameters based on loopy belief propagation (FREY & MACKAY 1998), a standard iterative message passing algorithm for graphs with cycles. The result is one probability value per class for each data point.

4 Experiments

For the evaluation of our approach we use the dataset described in section 1.2. Because lidar data of the Wadden Sea are not readily acces-

sible, we are limited to consider only one dataset. Ground truth is generated by labelling the point cloud manually. As a minimum requirement the connected network of tidal channels have to be found, for example, to fuse the DTM at these areas with bathymetric data. However, isolated local water regions which are often higher than the large tidal channels remain. Of course, the quality of the laser DTM is questionable in such areas, too.

Both methods are supervised approaches and thus require a training step. For training we took about 2% of the points from the entire dataset. Then, the computational costs in the CRF approach depending on the number of features (here: 3–5) and the size of neighbourhood (here: 2) vary between 0.6 and 3.8 minutes. For the algorithm of BRZANK et al. (2008), the computational costs are approximately 0.3 minutes. We processed the tests on a machine with a 2.8 GHz Quad-Core CPU and 8 GB RAM.

In section 4.1 we compare the results of our algorithm with those obtained by the method of BRZANK et al. (2008), wherefore the original software was available for us. In this way, we analyze the influence of contextual information for the land-water-classification. Moreover, we assess the impact of full waveform information on the CRF classification (section 4.2), in particular of the width. For the qualitative evaluation we use the *completeness* and *correctness* rates as well as the *quality* (HEIPKE et al. 1997).

4.1 Comparison of both Approaches

We classify the data using the fuzzy-logic-based method of BRZANK et al. (2008) and our CRF framework. Here, we use the same features (group I: *height, amplitude, point*

density) and training data. Tab. 1 and Fig. 5 show the classification results. It can be seen that both classes have a high rate of correctness and completeness of more than 80% and in most cases of more than 90% in both approaches. By integrating context of the CRF framework the classification rate of *water* rises by approximately 10%. Especially the classification results in the transition zone can be significantly improved. Due to the feature *height*, for which a high weight is derived in the training step of the fuzzy-based approach, misclassifications occur in areas with low local height differences. Thus, the borders of the tidal channels are shifted further onto the land. Moreover, isolated small water areas in different height levels are not detected using the method of BRZANK et al. (2008). These problems can be overcome by integrating contextual knowledge using the CRF approach.

4.2 Comparison of Different Classification Features

We investigate the influence of waveform information on the water detection with our CRF approach in a second test. Here, we use only information about the waveform (group II: *amplitude, width, and variation of width*) as classification features. Fig. 6 illustrates that the variation of the signal width indicates water areas. By a classification using this features group also a high rate of completeness and correctness of more than 85% can be obtained for both classes (Tab. 1). The correctness of *water* decreases due to misclassification of some points with low intensity and high variation of their signal width (both characteristicly for water areas) on some land areas. However, the correctness of *land* as well as the completeness of *water* can be improved. Over-

Tab. 1: Correctness (CR), completeness (CP) and quality (Q) for land (L) and water (W) in the fuzzy logic and the conditional random fields (CRF) framework using different feature combination (I = *height, amplitude, point density*, II = *amplitude, signal width, variation of width*).

	Fuzzy Logic (I)			CRF (I)			CRF (II)			CRF (I + II)		
	CR	CP	Q	CR	CP	Q	CR	CP	Q	CR	CP	Q
L	98.3	94.6	93.1	98.8	97.9	96.7	99.3	94.7	94.1	99.2	97.2	96.4
W	84.3	94.7	80.5	93.5	96.2	90.2	85.2	97.9	83.7	91.5	97.4	89.3

all, it can be seen that even by neglecting geometrical information and the important feature *height* the water-land-classification delivers good results which we could not observe in our previous work in urban areas (SCHMIDT et al. 2011). Combining both features groups, the

correctness and completeness rates are comparable to the good results of the classification with the features *height*, *amplitude* and *point density*. Moreover, the correctness of *land* and the completeness of *water* can be improved by the full waveform features.

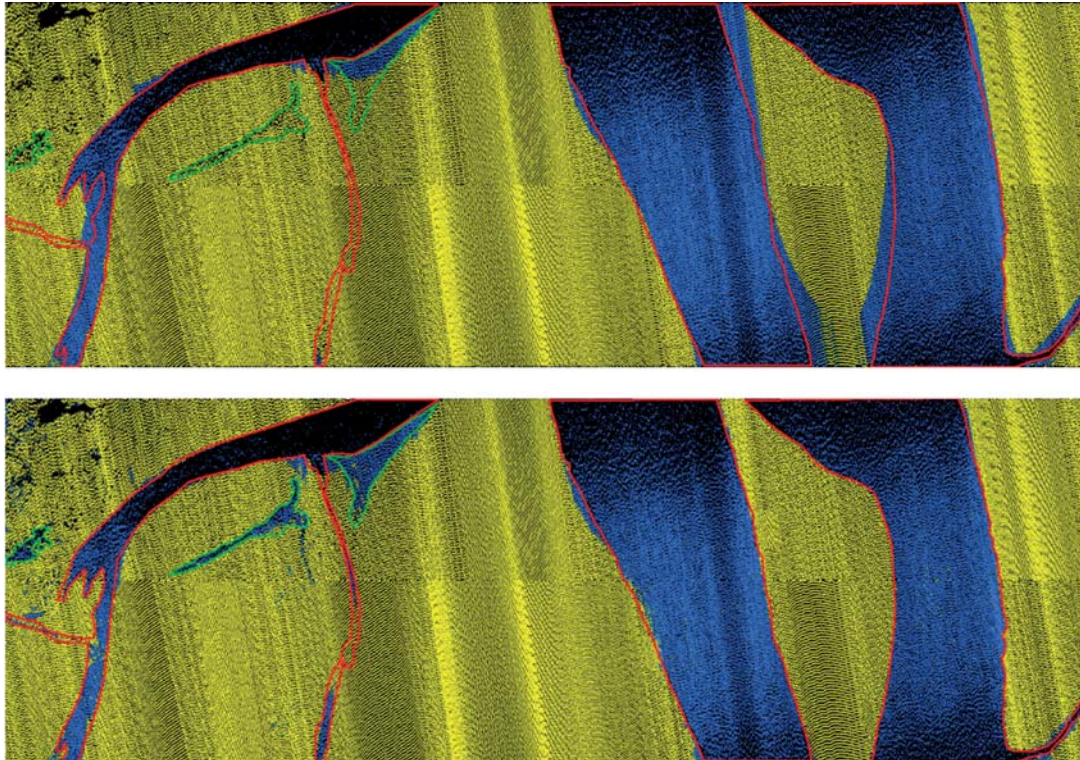


Fig. 5: Classification results of the test area [0.4 km \times 1.2 km] for *water* (blue) and *land* (yellow) using the fuzzy logic (top) and the CRF framework (bottom) with feature group I. Big tidal channels (outlined in red) are well detected by both approaches. Some small water covered areas on a higher height level (outlined in green) are only labelled as *water* by the CRF method.

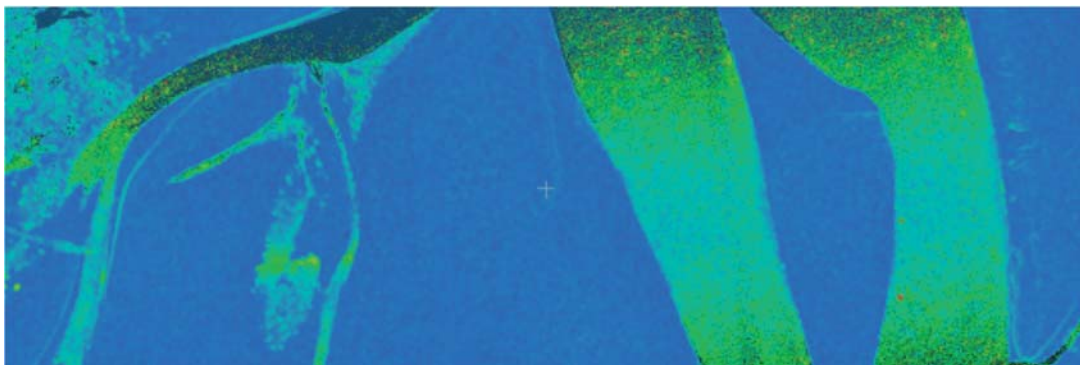


Fig. 6: Distribution of the feature *variation of signal width* coloured from low (blue) to high (green).

5 Conclusion

In this paper we are proposing a method for the water-land-classification of full waveform lidar data in coastal areas. For this task we presented suitable classification features and learnt typical structures of the data in a training step. We integrated contextual knowledge in a supervised classification process based on CRF. As result of the classification process, each point is assigned to one of the two classes *water* and *land*. We compared the results to a non-contextual method and showed that small water bodies on different height levels can be detected by our approach, whereas the algorithm based on the assumption of equal height levels fails. In this way the correctness rate of water can be significantly improved. Moreover, we investigate the full waveform information on water areas. It can be seen that only by using full waveform features (amplitude, signal width and its variation) good results can be obtained. Concerning the width of the signal, we observed an unexpected behaviour in our dataset where high values of the width occur in nadir regions. We assume undesired effects in the receive path of the sensor due to signal saturation.

In the future we intend to further examine this non-linearity because such an effect may occur also for other types of land cover. Moreover, we want to extend the classification to additional land surface types, e. g. by integrating texture features to get detailed information about the structure of the coastal surfaces.

References

- BROCKMANN, H. & MANDLBURGER, G., 2001: Aufbau eines digitalen Geländemodells vom Wasserlauf der Grenzoder. – Tagungsbericht zur 21. Wissenschaftlich-Technischen Jahrestagung der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung: 199–208, Konstanz.
- BRZANK, A., HEIPKE, C., GOEPFERT, J. & SÖRGEL, U., 2008: Aspects of generating precise digital terrain models in the Wadden Sea from lidar – water classification and structure line extraction. – ISPRS Journal of Photogrammetry and Remote Sensing **63** (5): 510–528.
- FREY, B. & MACKAY, D.J.C., 1998: A revolution: Belief propagation in graphs with cycles. – Advances in Neural Information Processing Systems **10**: 479–485.
- HEIPKE, C., MAYER, H., WIEDEMANN, C. & JAMET, O., 1997: Evaluation of automatic road extraction. – International Archives of Photogrammetry and Remote Sensing **32** (3–4, W2): 151–160.
- HÖFLE, B. & PFEIFER, N., 2007: Correction of laser scanning intensity data: data and model-driven approaches. – ISPRS Journal of Photogrammetry and Remote Sensing **62** (6): 415–433.
- HÖFLE, B., VETTER, M., PFEIFER, N., MANDLBURGER, G. & STÖTTER, J., 2009: Water surface mapping from airborne laser scanning using signal intensity and elevation data. – Earth Surface Processes and Landforms **34** (12): 1635–1649.
- JUTZI, B. & GROSS, H., 2010: Investigations on surface reflection models for intensity normalization in airborne laser scanning (ALS) data. – Photogrammetric engineering & remote sensing **76**: 1051–1060.
- KUMAR, S. & HEBERT, M., 2006: Discriminative Random Fields. – International Journal of Computer Vision **68** (2): 179–201.
- LEHNER, H. & BRIESE, C., 2010: Radiometric calibration of Full-Waveform Airborne Laser Scanning Data based on natural surfaces. – 100 Years ISPRS Advancing Remote Sensing Science **38** (7B): 360–365.
- LIM, E.H. & SUTER, D., 2009: 3d Terrestrial Lidar Classifications with Super-Voxels and Multi-Scale Conditional Random Fields. – Computer-Aided Design **41** (10): 701–710.
- LIU, D.C. & NOCEDAL, J., 1989: On the Limited Memory BFGS method for large scale optimization. – Mathematical Programming **45**: 503–528.
- MASON, D.C., SCOTT, T.R. & WANG, H., 2006: Extraction of tidal channel networks from airborne scanning laser altimetry. – ISPRS Journal of Photogrammetry and Remote Sensing **61** (2): 67–83.
- NIEMEYER, J., WEGNER, J., MALLET, C., ROTTENSTEINER, F. & SÖRGEL, U., 2011: Conditional random fields for urban scene classification with full waveform lidar data. – Photogrammetric Image Analysis (PIA), Lecture Notes in Computer Science **6952**: 233–244, Springer, Berlin / Heidelberg.
- SCHMIDT, A., ROTTENSTEINER, F. & SÖRGEL, U., 2011: Detection of Water Surfaces in Full-Waveform Laser Scanning Data. – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **38** (4, W19), Hannover.
- SHAPOVALOV, R., VELIZHEV, A. & BARINOVA, O., 2010: Non-associative Markov Networks for 3d Point Cloud Classification. – International Ar-

- chives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **38** (3A): 103–108.
- STEINBACHER, F., PFENNIGBAUER, M., AUFLEGER, M. & ULLRICH, A., 2012: High resolution airborne shallow water mapping. – *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **39** (B1): 55–60.
- WAGNER, W., HYYPPÄ, J., ULLRICH, A., LEHNER, H., BRIESE, C. & KAASALAINEN, S., 2008: Radiometric calibration of full-waveform small-footprint airborne laser scanners. – *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **37** (B1): 163–168.
- WAGNER, W., 2010: Radiometric calibration of small-footprint airborne laser scanner measurements: Basic physical concepts. – *ISPRS Journal of Photogrammetry and Remote Sensing* **65** (10): 505–513.
- WOLFE, W. & ZISSIS, G.J., 1993: *The infrared handbook*. – 1700 p., The Infrared Information Analysis Center, Environmental Research Institute of Michigan, Detroit, USA.

Address of the Authors:

M.Sc. ALENA SCHMIDT, PD Dr.-Ing. FRANZ ROTTENSTEINER, Prof. Dr.-Ing. UWE SÖRGEL, Leibniz Universität Hannover, Institut für Photogrammetrie und GeoInformation, Nienburger Str. 1, D-30167 Hannover, Tel.: +49-511-762-2482, Fax: +49-511-762-2483, e-mail: {alena.schmidt}{rottensteiner}{soergel}@ipi.uni-hannover.de

Manuskript eingereicht: November 2012
Angenommen: Januar 2013



A Stereoscopic Approach for the Association of People Tracks in Video Surveillance Systems

MORITZ MENZE, TOBIAS KLINGER, DANIEL MUHLE, Hannover, JÜRGEN METZLER, Karlsruhe & CHRISTIAN HEIPKE, Hannover

Keywords: video surveillance, stereo vision, people tracking, camera network

Summary: This article describes the application of stereoscopic analysis to typical image pairs from a surveillance camera network. An approach is presented that establishes correspondences between people detections across adjacent views and derives an estimation of body height for each person in the overlapping parts of the camera views. Dense image matching is applied to short stereoscopic sequences and the results are incorporated in a subsequent monocular tracking to improve the positioning accuracy. The method does not depend on a dedicated stereo setup of the camera network but is applicable to suitable image pairs in addition to monocular people detection and tracking. Based on realistic image sequences, the performance of the proposed approach is evaluated and compared to a current method for appearance-based data association.

Zusammenfassung: Ein stereoskopischer Ansatz zur Zuordnung von Personenpfaden in Multi-Kamera Überwachungssystemen. Dieser Artikel beschreibt die Anwendung eines stereoskopischen Analyseverfahrens auf typische Bildpaare eines Videoüberwachungssystems. Es wird ein Ansatz vorgestellt, der mit Hilfe der dichten Bildzuordnung aus wenigen Stereoansichten von Personen eine Größenschätzung ableitet und eine zuverlässige Übergabe von verfolgten Personen zwischen benachbarten Kameras ermöglicht. Das Verfahren ist nicht auf den Einsatz spezieller Stereosysteme angewiesen, sondern kann in Ergänzung monokularer Methoden in bestehenden Kameranetzen eingesetzt werden. Experimentelle Untersuchungen mit realistischen Bildsequenzen zeigen die Leistungsfähigkeit des vorgestellten Verfahrens verglichen mit denen weiterer Ansätze zur Übergabe von Personenpositionen zwischen verschiedenen Überwachungskameras.

1 Introduction

The automated analysis of surveillance videos is an area of active research, primarily within the Computer Vision community. Given current pan-tilt-zoom (PTZ) cameras equipped with mechanical control as well as on-board computers, research focuses for example on the development of self-organising smart camera networks (BELBACHIR 2010, JÄNEN et al. 2011). Major challenges in terms of image analysis are people detection and tracking as well as the reliable association of trajectories from individual cameras across multiple views. The latter indicates a need for the consistent handling of objects in a common reference frame in order to produce suitable data for wide area

analyses (COLLINS et al. 2001). The more accurate the object coordinates of people in the scene are known, the more detailed analyses can be conducted with respect to motion patterns or interactions between tracked people.

In this paper we present an approach that generates consistent global tracks of people in non-crowded scenarios. The trajectories are calculated in a common reference frame from observations of multiple surveillance cameras. An important step in our work is the estimation of body height as well as a reliable association of tracks across partly overlapping views. For that purpose, stereoscopic analysis is applied to overlapping parts of images which are either generated randomly while scanning wide areas with several PTZ cam-

eras or result from an appropriate reconfiguration of the sensor network in order to focus on areas of special interest, for example as a result of saliency detection. The presented module can be employed in addition to monocular analysis whenever the necessary preconditions are met. To demonstrate the benefit of the proposed approach we process a pair of partially overlapping image sequences according to the described setup and compare our results to manually measured reference data as well as an appearance-based method for people re-identification.

The remainder of the paper is structured as follows. In the next section we summarise previous work. Section 3 contains the strategy of our approach. It is arranged in subsections on basic geometric relations, monocular people detection and tracking, image matching and data association across multiple views. Section 4 shows experimental results derived from a realistic video sequence. Conclusions and an outlook on further work are given in section 5.

2 Related Work

In this section the state-of-the-art in data association within and across camera views is briefly reviewed. It references related work on stereo vision in surveillance applications and discusses approaches to pose estimation in the same domain.

Data association in the context of video surveillance aims at the concatenation of corresponding observations between temporally or spatially adjacent video frames. Features of tracked people that are evaluated to establish correspondences can roughly be categorised into either appearance-based or spatiotemporal features. Appearance-based features like colour histograms, descriptors of interest points or combinations of both are discussed in DORETTO et al. (2011). Especially for wide baselines the appearance of one and the same object may vary significantly in the views of adjacent cameras, so that it is error-prone to use only appearance-based features for the association. Another approach to establishing correspondences is the comparison of spatiotemporal features like position and velocity

(ORWELL et al. 1999). Since either of the methods suffers from individual shortcomings, combined approaches were introduced. CAI & AGGARWAL (1999) hand over targets between adjacent cameras predicting the position of the target in an adjacent view for re-identification. JAVED et al. (2008) model probabilities of people walking certain paths and fuse this geometrical information with appearance cues using a maximum likelihood framework.

Due to the challenging task of consistent tracking, several authors gather additional depth information about the observed scene by applying stereoscopic analysis. Because stereoscopic image matching is often regarded as the central component in surveillance systems, the sensor networks are designed to fulfil the requirements of stereo approaches. The pairwise installation of PTZ cameras (ZHOU et al. 2010) provides image pairs with short baselines. Dedicated stereo devices, as used in DARRELL et al. (2000), HARITAOGU et al. (1998) and many other publications, capture synchronised image pairs that are processed on specialised hardware. Although the advantages of high-frequency depth maps for people detection and tracking is shown (SCHINDLER et al. 2010), a dedicated system design leads to additional costs that, from our point of view, are not necessary, when applying stereoscopic analysis to camera networks.

In contrast, we propose the usage of stereo vision in PTZ sensor networks where overlapping fields of view are not predefined, but selectively available during short periods of handover, i. e. data association between adjacent cameras. This makes the approach applicable to existing systems of spatially distributed cameras. Nevertheless, there is still the additional need for monocular tracking where the stereoscopic approach cannot be applied.

In camera networks conceptually two approaches exist to estimate the object position in a common coordinate reference frame. The monocular position estimation can produce planar coordinates given the ground plane and the orientation parameters, either by directly measuring the intersection of objects with that plane (COLLINS et al. 2001) or by incorporating assumptions about the object height (ZHAO & NEVATIA 2004). Directly observing the point of intersection is often not possible due to occlu-

sions, thus the observation of head tops and the incorporation of default object height is a widely used approximate solution. A second, more robust approach to position estimation is the use of multiple cameras. ESHEL & MOSES (2010) estimate body height calculating correspondences between multiple homographies at discrete height levels. Methods of the second category rely on the simultaneous observation of the objects throughout the entire scene, which is often not feasible due to the high amount of required resources, at least in large camera networks. Since we extract height from short stereo sequences during handover and subsequently apply monocular tracking, this limitation does not hold for the presented approach.

3 Approach

3.1 Overview

Given a reconfigurable network of sparsely distributed PTZ cameras, we apply monocular object tracking most of the time and exploit stereo vision whenever people pass an area viewed by two or more cameras to derive a 3D point cloud of the visible surface of each person. Based on the centroid of each dense point cloud we establish reliable correspondences between individual tracks across the views. In addition, we estimate body height which is used in all subsequent frames to improve the positioning accuracy of trajectories from monocular tracking.

3.2 Geometric Relations

The observations of camera networks have to be transformed to a common reference frame as a prerequisite to the automated interpretation of the whole scene in object space. In video surveillance scenarios there often exists one predominant ground plane. If this is the case, an intuitive definition of a reference frame is to align its X/Y plane to the predominant ground plane in the scene. The axes are rotated so that the Z axis directly represents object height. Image coordinates (x, y) can be projected onto a plane in object space by

applying the collinearity equations (1), if the height Z of the corresponding object is known or if it lies in the ground plane (Z=0).

$$\begin{aligned} X &= X_0 + (Z - Z_0) \frac{r_{11}(x' - x'_0) + r_{12}(y' - y'_0) + r_{13}c'}{r_{31}(x' - x'_0) + r_{32}(y' - y'_0) + r_{33}c'} \\ Y &= Y_0 + (Z - Z_0) \frac{r_{21}(x' - x'_0) + r_{22}(y' - y'_0) + r_{23}c'}{r_{31}(x' - x'_0) + r_{32}(y' - y'_0) + r_{33}c'} \end{aligned} \quad (1)$$

The perspective centre of the camera (X_0, Y_0, Z_0) in the reference coordinate frame and the elements r_{ij} of the rotation between image and reference frame are determined via spatial resection from ground control points within the scene. The image coordinates of the principal point x'_0, y'_0 and the principal distance c' are assumed to be known from camera calibration. Zoom functionality is not taken into account in this paper because we aim at covering wide areas given a sparse network of cameras. Therefore, images are captured at lowest zoom level.

The influence of an erroneous object height on the projected position in object space is illustrated in Fig. 1. In common surveillance scenarios incorrect Z values result in a horizontal bias in the viewing direction of the camera. The partial derivatives of the inverse collinearity equations describe the impact of errors in the parameters. The most important factor is the body height, i. e. Z in (1).

Given a maximum deviation of 0.2 m from the default height (we use 1.72 m, see below) the maximum bias in imaging direction varies from 0.5 m to 1.1 m depending on our setup and imaging distances.

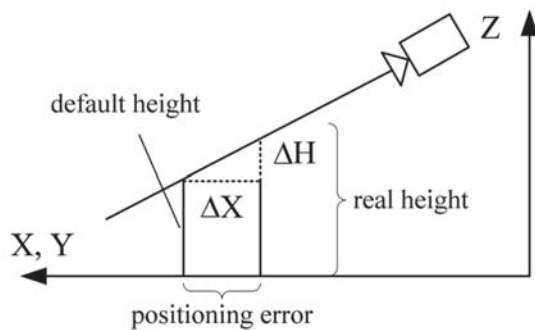


Fig. 1: Influence of height on target localisation from a monocular view.

3.3 Detection and Tracking in Monocular Views

People detection aims at distinguishing people from the image background. Tracking describes the process of establishing temporal correspondences between temporally corresponding people in consecutive frames. In this work we do not concentrate on finding the optimal method for monocular tracking, but apply a combination of state-of-the-art methods that works well enough in non-crowded scenarios.

For detection we follow DALAL & TRIGGS (2005) and classify histograms of oriented gradients within a sliding window in either people or non-people using a support vector machine (SVM). As shown in DOLLAR et al. (2011), false positives in sliding window approaches remain frequent. Therefore, we validate the detections with clues from background subtraction using an improved Mixture of Gaussians approach with shadow detection (KAEWTRAKULPONG & BOWDEN 2001), and select only the detections that have a sufficiently large overlap with any foreground region. Assuming a static background, valid for most surveillance scenarios, we successfully avoid false positive detections with this strategy. However, misplaced detections occur if background structure causes positive classification of a detection window which is also labelled as foreground (Fig. 2). Such incorrect detections are eliminated during data association.

We obtain temporal trajectories by finding the closest match of features between detections in adjacent frames. A greedy approach is applied, combining spatiotemporal and appearance-based features: For each target an appearance model and a motion model



Fig. 2: People detections, resized to unit height. Two correct detections (left) and two incorrect ones (right).

are set up. The appearance of each detection is described in terms of the hue histogram of the corresponding foreground blob. An appropriate similarity measure is given by the Bhattacharyya distance between histograms (BHATTACHARYYA 1946). After successful association the appearance model is updated by incorporating the appearance of the associated detection, which makes the tracker robust against small appearance changes. The detection with the highest similarity is associated to a trajectory, if it also fits the motion model, i. e. considering spatial position and gating the search space.

A spatiotemporal description of each tracked object is realised by a Kalman Filter based on the planar position and velocity on the ground plane. For estimating the position from a monocular view one often cannot rely on the visibility of a clearly defined intersection point of the person with the ground due to occlusions, shadows and the fact, that while walking the feet obviously do not always touch the ground. The upper point of a standing person, in contrast, is likely to be identifiable in such situations. Since in monocular tracking the person's height is not inherently known, a default height is assumed for (1). For our experiments we have chosen a height of 1.72 m which is the average body height of a German adult (STATISTISCHES BUNDESAMT 2009). As shown in section 3.1 height influences target localisation in the imaging direction. This does not necessarily degrade the performance of monocular tracking, since consecutive observations are influenced in the same systematic way. On the contrary, association across views is handicapped because the viewing directions of different cameras are generally not aligned.

The described approach can be transferred to the problem of data association across overlapping views. Only the gating of the search space has to be adapted from a temporal motion model to simultaneous observations.

3.4 Dense Stereo Matching and Feature Extraction

Another approach to data association across multiple views is based on dense image match-

ing. The general idea is to deduce the geometric structure of the scene from a pixel- or region-based comparison of radiometric features in simultaneously captured images. The resulting disparity map then contains a pixel-wise data association between the images.

Since we are primarily interested in pedestrians moving in the scene, the input to the matching algorithm is reduced to the foreground regions of the corresponding images inside the person tracker's bounding boxes, available from the people detection step. This intrinsically eliminates many potential mismatches. For each of the foreground regions dense stereo matching generates a disparity map that can be re-projected to object space as a point cloud on the visible surface of the respective person. As an example Fig. 3 shows two foreground patches from the input data next to the resulting disparity map and the derived point cloud, textured with colour information from the left image.

A prerequisite for successful image matching of moving objects are temporally aligned image pairs. In the absence of externally triggered cameras, we select images based on their timestamp, dropping frames without a corresponding partner. Camera clocks are synchronised using the same Network Time Protocol (NTP) server.

Corresponding images are normalised to epipolar geometry to reduce the search space in the matching procedure. The necessary image orientation is derived in a separate, automatic calibration step.

Consistent disparity maps for noisy and low textured images, as frequently encountered in surveillance footage, can be obtained by introducing smoothness constraints in the matching algorithm. An efficient implementation of such an optimisation is employed in SGM

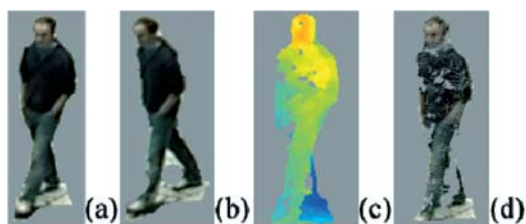


Fig. 3: Input regions for dense image matching (a, b), colour coded disparity map (c) and coloured 3D point cloud (d).

(Semiglobal Matching, HIRSCHMÜLLER 2008) where a global cost function combines local matching costs with two penalty terms for different changes in disparity. Overall matching costs are minimised by means of dynamic programming along several paths performed for each pixel. The matching procedure which uses the Open Source Computer Vision library (OPENCV 2012) is based on SGM: the optimisation of an energy function integrates the described smoothness constraints and pixel-wise local matching costs, the latter being calculated as the sum of absolute differences (SAD) in a 3×3 window. This method is chosen due to its higher computational efficiency compared to dedicated wide baseline approaches like the DAISY descriptor (TOLA et al. 2010) that allows for larger angles of convergence.

Each disparity is validated by a back-matching consistency check as proposed e.g. by HANNAH (1989). Pixels that are matched from left to right have to be confirmed by a corresponding match from right to left.

After calculating the disparity image for each foreground region, pixels inside the bounding boxes of the detector are re-projected to 3D object space and transformed to the common reference frame. From the resulting point clouds, which in our setup consist of 1500 up to 3500 points on the visible surface of each person, the object height is extracted as the mean Z value of the five topmost points inside an axis-aligned square buffer of 0.8 m side length around the median X and Y coordinates corresponding to a specific bounding box.

3.5 Data Association across Multiple Views

Whenever a stereo image pair is processed, matching results are employed to establish correspondences between trajectories in the respective views. Globally tracked objects are mostly instantiated from a single view and associated to corresponding detections in adjacent cameras when the object enters the overlapping regions of the views.

To establish correspondences, the centroids of the three-dimensional point clouds from the matching step are re-projected to the im-

ages. Note that each point cloud corresponds to a detected person. In the image domain it is tested whether the projected point lies inside one of the bounding boxes of the people detector in the corresponding frame of the stereo image pair. If this is the case, the link is stored in a global data-structure representing the tracked object.

Correspondences are only established for unambiguous associations; the results are discarded if the projected centroid falls into multiple boxes in an image due to overlapping detections. Single misses due to such occlusions are tolerable if the overlapping region of the images is large enough to produce several synchronous image pairs showing the person. Each of the pairs is processed individually and is analysed with respect to possible associations. The presented approach conceptually works with only one valid association. However, multiple associations of the same tracks increase reliability.

This procedure relies on simultaneous detections in both views and does not directly support detection by providing clues to the presence of people in the scene. Such integrated procedures are used in multiple view detection approaches, e.g. ZHAO et al. (2005), whereas this paper focuses on handover in short sub-sequences.

4 Experiments

4.1 Overview

In this section, the proposed approach is evaluated on realistic surveillance footage collected in an indoor setting by our prototype system. First, the dataset is described. The remaining subsections evaluate the performance of the proposed approach with respect to data association and the incorporation of the body height estimation.

4.2 Dataset

Although there are several publicly available test datasets addressing people detection and tracking, e. g. performance evaluation of tracking and surveillance (PETS 2012), con-



Fig. 4: Image pair from the test sequence.

text aware vision using image-based active recognition (CAVIAR 2003), and video and image retrieval and analysis tool (VIRAT 2011), none of them provides suitable input to the presented approach. This is due to the fact that, although a few overlapping views exist in some of the datasets, convergence angles and scale differences exceed reasonable limits for stereoscopic analysis as described in this paper.

Therefore, we have decided to acquire additional test data ourselves. To demonstrate handover and tracking, image sequences from two cameras are processed. Fig. 4 shows a synchronous pair of images from the experimental setup. The cameras are mounted approximately 5.7 m above the ground plane with a stereo base of 4.2 m. People pass at distances from 8.5 m to 18.5 m in front of the camera, yielding base-to-distance ratios of 0.49 to 0.23. The image resolution of each camera is 768×576 pixels and images are captured at a frame rate of 15 fps.

4.3 Data Association

Handover is evaluated on a test sequence of 1000 image pairs with 20 people passing the overlapping region. The density of people is relatively low and does not exceed 3 people crossing the region of overlap simultaneously. To be able to process all possible associations, manually labelled detections are used for the experiments in this subsection.

One computationally efficient geometric approach to data association in object space is to establish links based on the distance between detections from different views and a threshold. To compare this strategy (called “mono” in the remainder) to the stereoscopic approach, a simple metric is defined by the ratio of successfully associated vs. all possible

Tab. 1: Absolute number and percentage of correct associations on a per frame basis.

	no. correct	% correct
mono	377	74.8
stereo	466	92.5
mono+appearance	503	99.8

Tab. 2: Absolute number and percentage of correct associations on a per track basis using manually labelled detections.

	no. correct	% correct
mono	18	90
stereo	18	90
appearance	19	95
mono+appearance	20	100

associations. Given manually labelled reference data we can directly compare the performance of the monocular and the stereoscopic approach on a per frame basis. A threshold of 0.3 m was used for the mono cases. Tab. 1 shows that the stereoscopic approach clearly outperforms the simpler mono approach. However, the combination of monocular geometric positioning and appearance-based features, as described in section 3.3, yields almost flawless results for the given dataset.

To evaluate the performance of the presented approach with respect to the state-of-the-art, we processed the described test dataset with a recent appearance-based approach to people re-identification (METZLER 2012) which is based on the mean of covariance descriptors calculated from several images and does not incorporate position information. Since the appearance-based approach builds a descriptor from multiple images of each person, it cannot be compared on a per frame basis. Tab. 2 gives the results in terms of the number and percentage of correctly associated tracks.

The stereoscopic approach works almost as well as the appearance-based method which has the advantage of relatively few candidate matches in the second image. The combined spatiotemporal and appearance-based method (section 3.3) is able to associate all 20 tracks

**Fig. 5:** Heavy occlusion hindering people detection and reconstruction.

correctly. However, the results differ by one or two detected tracks only, when manually detected bounding boxes are used. Thus, it can be concluded that for this experiment, all investigated methods deliver nearly the same quality of results.

Using the results of our detector/tracker instead of manual labelling, the overall performance of the system degrades. Due to partial or complete misses of the detector, the absolute number of possible associations becomes smaller and the reliability of data association decreases as well. For our test dataset only 17 out of 20 people are detected.

Although the stereoscopic approach can deal with a certain amount of occlusion, the reconstruction fails, just as the detector does, if people walk too close together. Fig. 5 depicts a situation where only the person in front is detected and reconstructed. Since the problems depicted in Fig. 5 stem from the detection deficiencies, similar problems occur with the other three association methods. In summary, all four association methods delivered nearly identical results for these detections.

4.4 Incorporation of Body Height

The effect of body height estimation is discussed based on the results of our tracker for a single trajectory. Fig. 6 shows an exemplary tracking result in the reference frame of our test site. The upper right of the figure corresponds to the upper right of the images in Fig. 4. The moving direction of the target that is tracked was from the upper right to the lower left corner in Fig. 6. The rectangle in the middle indicates the area in which the target was observed stereoscopically. Manually labelled ground truth is depicted as a blue

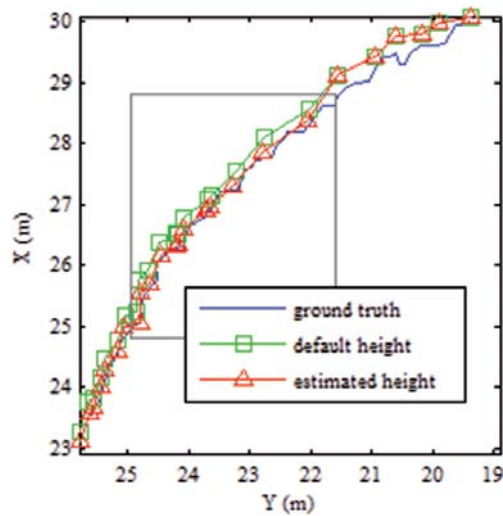


Fig. 6: Exemplary trajectory in a common reference frame. Parts inside the rectangle are processed stereoscopically.

line. The green line is the trajectory after associating the monocular tracking results from the left and right view using the described stereoscopic approach. Detections are projected to the ground plane using the average height. The red line is the trajectory updated with the height estimation from stereo processing as soon as available. None of the trajectories are filtered in the time domain. Initially, both the green and the red trajectory show a systematic offset from what is labelled as ground truth because of the default height they assume. Note that the red trajectory converges towards the reference data from the moment the target reaches the overlapping area, while the green trajectory retains its systematic offset in X direction (imaging direction).

Tab. 3 gives numerical results of a comparison between our measurements and manually labelled reference data, computed as the root-mean-squared difference between corresponding points in the ground plane. The left part gives overall results of the trajectories in Fig. 6 and therefore includes biased re-

sults from default height. The mean difference between tracking results and reference data is reduced significantly from 26 cm to 11 cm in X direction (which is the imaging direction) while differences in Y direction are not influenced. This is the expected improvement, since an error in height directly results in a biased position in imaging direction.

The overall improvement of the target localisation with respect to the reference data demonstrates the benefit of our approach. In the right part of Tab. 3, only those trajectory points with an estimated body height are compared to the respective biased results. While the mean difference in Y direction slightly changes due to the reduced number of samples, the differences in X direction are reduced by an order of one magnitude.

5 Conclusions and Outlook

We presented an approach to the generation of globally consistent trajectories from surveillance videos. Our contribution deals with the handover of tracked objects between different cameras with occasionally overlapping fields of view. By exploiting stereo vision during handover a reliable estimation of body height can be obtained. The major benefit of the approach is the increased geometric positioning accuracy during stereoscopic and subsequent monocular tracking. An improved geometric accuracy of the trajectories enables a more precise description of movements in the scene and a more detailed analysis of interactions.

The approach has been successfully tested on realistic image sequences. Positioning accuracy is improved and data association performs on a level comparable to state-of-the-art methods. The integration of a more sophisticated association procedure could further improve these results. Ambiguous associations of the stereoscopic approach due to overlap-

Tab. 3: Mean difference of tracking results and reference data.

	overall		overlapping	
	ΔX (m)	ΔY (m)	ΔX (m)	ΔY (m)
default height	0.26	0.02	0.22	0.06
estimated height	0.11	0.02	0.02	0.06

ping bounding boxes in the second image could be resolved by incorporating additional appearance-based features in such cases.

Focussing single people of interest at higher resolution will yield more detailed point clouds. Those could be analysed with respect to full body motion and action recognition. More extensive tests will be conducted in future works by integrating the approach into a wider network of self-organising smart cameras.

Our method is designed for the application in reconfigurable sensor networks with limited resources that use spatially distributed PTZ cameras to cover wide areas. In such a setup, stereo vision is only applied during the short periods of handover. The improvement in target localisation achieved by incorporating the correct body height makes our approach also suitable for systems that aim at the optimisation of global trajectories in post-processing. So far, only online applications in self-organising smart camera networks are considered that require tracking results on the fly. Applying corrections to past observations after a body height was determined may be useful for the backward tracking of people in recorded sequences.

Acknowledgements

This research was funded by the German Federal Ministry of Education and Research (BMBF), 13N10809 – 13N10814. The support is gratefully acknowledged.

References

- BELBACHIR, A. (ed.), 2010: *Smart Cameras*. – 1. edition, 404 p., Springer, New York.
- BHATTACHARYYA, A., 1946: On a Measure of Divergence between Two Multinomial Populations. – *Sankhya: The Indian Journal of Statistics* **7**: 401–406.
- CAI, Q. & AGGARWAL, J., 1999: Tracking human motion in structured environments using a distributed-camera system. – *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21** (11): 1241–1247.
- CAVIAR, 2003: <http://homepages.inf.ed.ac.uk/rbf/CAVIAR> (7.1.2013).
- COLLINS, R., LIPTON, A., FUJIYOSHI, H. & KANADE, T., 2001: Algorithms for Cooperative Multisensor Surveillance. – *IEEE* **89**: 1456–1477.
- DALAL, N. & TRIGGS, B., 2005: Histograms of oriented gradients for human detection. – *IEEE Conference on Computer Vision and Pattern Recognition* **1**: 886–893, San Diego, USA.
- DARREL, T., GORDON, G., HARVILLE, M. & WOODFILL, J., 2000: Integrated Person Tracking Using Stereo, Color, and Pattern Detection. – *International Journal of Computer Vision* **37** (2): 175–185.
- DOLLAR, P., WOJEK, C., SCHIELE, B. & PERONA, P., 2011: Pedestrian Detection: An Evaluation of the State of the Art. – *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34** (4): 743–761.
- DORRTO, G., SEBASTIAN, T., TU, P. & RITTSCHER, J., 2011: Appearance-based person reidentification in camera networks: problem overview and current approaches. – *Journal of Ambient Intelligence and Humanized Computing* **2** (2): 127–151.
- ESHEL, R. & MOSES, Y., 2010: Tracking in a Dense Crowd Using Multiple Cameras. – *International Journal of Computer Vision* **88** (1): 129–143.
- HANNAH, M., 1989: A System for Digital Stereo Image Matching. – *Photogrammetric Engineering and Remote Sensing* **55** (12): 1765–1770.
- HARITAAGLU, I., HARWOOD, D. & DAVIS, L., 1998: W⁴S: A Real-Time System for Detecting and Tracking People in 2 ½ D. – *Lecture Notes in Computer Science* **1406**: 877–892.
- HIRSCHMÜLLER, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. – *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30** (3): 328–341.
- JÄNEN, U., HUY, M., GRENZ, C., HOFFMANN, M. & HÄHNER, J., 2011: Distributed Three-Dimensional Camera Alignment in Highly-Dynamical Prioritized Observation Areas. – *IEEE International Conference on Distributed Smart Cameras*: 38–43.
- JAVED, O., SHAFIQUE, K., RASHEED, Z. & SHAH, M., 2008: Modeling inter-camera space-time and appearance relationships for tracking across non-overlapping views. – *Computer Vision and Image Understanding* **109** (2): 146–162.
- KAEWTRAKULPONG, P. & BOWDEN, R., 2001: An improved adaptive background mixture model for real-time tracking with shadow detection. – **2nd** European Workshop on Advanced Video Based Surveillance Systems: 1–5, London, UK.
- METZLER, J., 2012: Appearance-based Re-Identification of Humans in Low-Resolution Videos using Means of Covariance Descriptors. – *IEEE International Conference on Advanced Video*

- and Signal-Based Surveillance: 191–196, Beijing, China.
- OPENCV, 2012: Open Source Computer Vision library, <http://opencv.willowgarage.com> (1.1.2013).
- ORWELL, J., MASSEY, S., REMAGNINO, P., GREENHILL, D. & JONES, G., 1999: A Multi-agent Framework for Visual Surveillance. – International Conference on Image Analysis and Processing: 1104–1107, Venice, Italy.
- PETS, 2012: <http://www.pets2012.net> (7.1.2013).
- SCHINDLER, K., ESS, A., LEIBE, B. & VAN GOOL, L., 2010: Automatic detection and tracking of pedestrians from a moving stereo rig. – ISPRS Journal of Photogrammetry and Remote Sensing **65** (6): 523–537.
- STATISTISCHES BUNDESAMT, 2009: www.destatis.de (1.1.2013).
- TOLA, E., LEPETIT, V. & FUA, P., 2010: DAISY: An Efficient Dense Descriptor Applied to Wide-Baseline Stereo. – IEEE Transactions on Pattern Analysis and Machine Intelligence **32**: 815–830.
- VIRAT, 2011: <http://www.viratdata.org> (7.1.2013).
- ZHAO, T. & NEVATIA, R., 2004: Tracking multiple humans in complex situations. – IEEE Transactions on Pattern Analysis and Machine Intelligence **26**: 1208–1221.
- ZHAO, T., AGGARWAL, M., KUMAR, R. & SAWHNEY, H., 2005: Real-time wide area multi-camera stereo tracking. – IEEE Conference on Computer Vision and Pattern Recognition **1**: 976–983, San Diego, USA.
- ZHOU, J., WAN, D. & WU, Y., 2010: The Chameleon-Like Vision System, – IEEE Signal Processing Magazine **27** (5): 91–101.

Addresses of the Authors:

MORITZ MENZE, TOBIAS KLINGER, Dr.-Ing. DANIEL MUHLE, Prof. Dr.-Ing. habil. CHRISTIAN HEIPKE, Leibniz Universität Hannover, Institut für Photogrammetrie und GeoInformation, Nienburger Str. 1, 30167 Hannover, Germany, Tel.: +49-511-762-17488, Fax: +49-511-762-2483, e-mail: surname@ipi.uni-hannover.de

JÜRGEN METZLER, Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung, Abteilung Videoauswertesysteme, Fraunhoferstraße 1, 76131 Karlsruhe, Germany, Tel.: +49-721-6091-453, e-mail: juergen.metzler@iosb.fraunhofer.de

Manuskript eingereicht: November 2012

Angenommen: Januar 2013



Rainfall Estimation with a Geosensor Network of Cars – Theoretical Considerations and First Results

DANIEL FITZNER, MONIKA SESTER, UWE HABERLANDT & EHSAN RABIEI, Hannover

Keywords: geosensor network, decentralized processing, precipitation, rainfall, estimation, calibration

Summary: Areal rainfall information is one of the most important inputs to hydrological models. This paper presents some theoretical considerations and initial results on the idea of using a geosensor network of cars as a data source for areal rainfall estimations. The types of car sensors and different calibration schemes for the rainfall estimation functions (W-R functions) in the cars are presented. A special focus is given to the decentralized online calibration of these functions in the network by communicating measurements between measuring units. This would allow the dynamic adaptation of the functions to different situations such as different drivers, the current car environment or the current wind speed and direction. Then, results from laboratory and field experiments are presented.

Zusammenfassung: *Niederschlagsschätzungen mit einem Geosensornetz von Autos – Konzepte und erste Ergebnisse.* Schätzungen der räumlichen Niederschlagsverteilung sind eine wichtige Datengrundlage für hydrologische Modelle. Dieser Beitrag beschreibt grundlegende Konzepte der Regenschätzung mit einem Geosensornetz von Autos. Es werden sowohl die verwendeten Sensoren beschrieben, als auch die Möglichkeiten ihrer Kalibrierung. Dabei wird ein besonderer Fokus auf die dezentrale Online-Kalibrierung gelegt, die es ermöglicht, die Schätzfunktionen dynamisch anzupassen und so beispielsweise Faktoren wie Fahrer, Autoumgebung oder Windgeschwindigkeit und -richtung zu berücksichtigen. Im Anschluss daran werden erste Ergebnisse aus Labor- und Feldexperimenten zur Bestimmung der Schätzfunktionen präsentiert.

1 Introduction and Overview

Areal rainfall, representing a good estimation of the spatial variability and of the mean value over specific areas, is one of the most important inputs to hydrological models. Especially models used for reanalysis and forecasting of highly dynamic processes like floods and erosion have high requirements regarding the rainfall input. Furthermore, if predictions are required for small catchments or urban areas, the processes are very fast. However, estimating areal rainfall, especially for short time steps, is still a challenging task: In general, the density of recording rain gauges is low; further, weather radar suffers from large space-time biases and the general problem that rainfall is not measured, but estimated from the measured reflectivity values.

The idea of estimating rainfall with cars originates from HABERLANDT & SESTER (2010). There, computer simulations for a river catchment supported the assumption that areal rainfall estimations with cars might be superior compared to an existing network of stationary rain gauges. HABERLANDT & SESTER (2010) concluded that estimating areal rainfall with cars is theoretically feasible and that the accuracy depends on the number of cars equipped with sensors.

A possible application of this idea would be operational flood forecasting where real-time measurement and prediction of precipitation is required. Another possible application in a different field is car navigation, where real-time predictions of rainfall are needed for the online speed estimation for shortest path calculation (THAKURIAH & TILAHUN 2013). As cars

cannot measure rainfall directly, their original raw measurements, e. g. wiper frequency, have to be calibrated. This calibration has to take additional factors into account, such as speed and the local environment the car drives through. A realization of a system would require that the calibration has to be performed for each car, using data from each of its measurements. In a centralized calculation, all cars would communicate their measurements to a central server. If many cars participate in such a system, this leads to scalability problems, as well as a high amount of data to be communicated to the central server. Therefore, the idea is to use the measurement and processing capabilities of the car itself to conduct the calibration in a decentralized way, i. e. take the cars as a distributed geosensor network (DUCKHAM 2012). Such an approach not only solves the scalability problem, but also enables analyses beyond the calibration: the cars can collaborate in order to determine the rainfall pattern and detect homogeneities or heterogeneities; furthermore, the locally determined rainfall estimates can be used for other purposes such as risk warnings (aquaplaning) or real-time speed estimation.

SCHULZE et al. (2010) investigated this idea of calibrating the rainfall estimation functions in the cars while driving, subsequently called online calibration. Both earlier works relied entirely on computer simulations of either car measurements, rainfall or both. In FITZNER et al. (2012), online calibration of the estimation functions was investigated using real data from a car equipped with a wiper frequency sensor. The work analysed different online calibration schemes and sensor communication ranges. It concluded that, if the communication range (distance) is adequately short, the online calibration could improve rainfall estimation even under significant a priori training. However, not all influencing parameters such as car speed have been modelled yet in FITZNER et al. (2012) and no sound statistical model for the online calibration has been provided. Therefore, additional investigations are required, which can be summarized as follows:

a) Robust models for estimating rainfall from the sensor readings have to be derived from experiments.

- b) A statistically sound model for the online calibration has to be developed.
- c) Both, the feasibility as well as the advantages of online calibration have to be proven by experimental data and computer simulations.

This paper gives initial results for a) and discusses requirements for b). After a brief review of related works, the types of car sensors are introduced in section 3. Then, in section 4, the calibration of the car sensors in laboratory and field are discussed with a special focus on the decentralized online calibration. Section 5 gives first empirical results of the laboratory and field experiments and section 6 concludes.

2 Related Work

Information about rainfall is essential for hydrological predictions and water resources management (BEVEN 2001, CHOW et al. 1988). Due to its high variability in space and time, areal estimation of rainfall is still a challenging task. There are several methods for measuring rainfall: non-recording rain gauges are available in a high density, however they do only provide aggregates of the amount of precipitation over a whole day. Recording rain gauges would be required, however, even in Germany, the network density is only approximately one station per 1,800 km². Weather radar also is an indicator for rainfall, however, it does not measure rainfall directly, but reflections, which have to be transferred to rainfall using calibration – for which a sufficiently dense point precipitation network is needed (SMITH et al. 2007, KRAJEWSKI & SMITH 2002). Other special and innovative methods for rainfall observation use satellites (GRIMES & DIOP 2003, WARDAH et al. 2008), microwave links (LEIJNSE et al. 2007, MESSER et al. 2006) or rain gauges aboard moving ships to measure rainfall at sea (HASSE et al. 1998, YUTER & PARKER 2001). Utilising rainfall information from different sources together and applying sophisticated interpolation or merging methods can further improve precipitation estimation for hydrological applications (GOUDENHOOFDT & DELOBBE 2009, CHIANG et al. 2007, GOOVAERTS 2000, HABERLANDT 2007, EHRET et al. 2008). In

a recent study the inclusion of distributed low cost and low accuracy measurement devices for the improvement of radar rainfall was investigated (HILL & FARZAN 2012).

3 Car Sensors

Different types of sensors for estimating rainfall with cars have been used in the experiments. These include sensors for measuring the wiper frequency of a car as well as optical sensors that are typically installed in cars with automated wiper control.

A wiper frequency sensor has been developed in the course of the project which is based on a microcontroller connected to a GPS-receiver and a magnetic sensor. The magnetic sensor is placed behind the windscreen and triggers each time a magnet attached to the windscreen wipers passes, which occurs two times for each single wipe. Then an NMEA String, i. e. position, time and additional information, is recorded in an ASCII text file on an SD-card. The wiper frequency sensors are installed in cars with both manual and automatic wiper control. Cars with manual wiper control have discrete wiper frequency classes to be manually adjusted by the driver. In addition to the manual option, some cars have an automatic wiper system. When automatic wiping is switched on, an optical sensor attached to the windscreen detects raindrops on the windscreen surface and triggers the wipers. Also in this case, however, the wiper frequency is not determined completely automatic but depends on a manually controlled sensitivity setting.



Fig. 1: Xanonex (left) and Hydroeon (right) (XANONEX 2013, RAINSENSORS 2013).

Two other sensors are used, which measure rainfall in an optical way via transmitted and sensed infrared (Xanonex and Hydroeon) (Fig. 1).

4 Sensor Calibration – Derivation of the W-R-Relationship

In order to estimate rainfall with the sensors, a functional relationship between the sensor readings and rainfall needs to be established, termed Wiper-Rainfall (W-R) relationship or W-R function. The nature of such a relationship ranges from a simple linear regression with a single predictor variable, e. g. wiper frequency, to more complex non-linear models with multiple parameters. Calibration can either be performed in an off-line fashion in a controlled environment in a laboratory or also “in the field”, by equipping cars on the road with sensors. The latter can also be organized in an online way, taking advantage of the fact that cars can continuously measure wiper frequencies and compare it with given rainfall measurements of stations and other cars in their vicinity.

In the following the different calibration strategies are described. Whereas the major

Tab. 1: Influencing factors and their consideration in the current experiments.

Factor	Lab Experiments		Field Experiments
	Wiper Frequency	Optical Sensor	
Car speed	No	Yes	Yes
Windscreen angle	Yes	Yes	Yes
Car environment	No	No	Yes
Wind speed / wind direction	No	No	No
Drivers	No	No	Yes
Road type (and road surface), spray	No	No	Yes

parameters for the determination of the rainfall are the sensor readings, there are additional factors which can influence the calculation of the rainfall, such as car speed and preferences of the driver (Tab. 1). Some of the factors influence the rainfall as such, e. g. wind, car environment such as buildings, trees, others influence the measurement of the rainfall, e. g. speed, driver. The details are explained in section 4.1.

The calibration data for the sensors is collected in the laboratory or field and used for estimating the W-R function coefficients. This can be done e. g. by executing a least squares algorithm such as ordinary least squares (OLS). Once the coefficients are determined, they can be implemented in the cars to be used for calculating rainfall during subsequent rainfall events. The offline calibration in a laboratory and field setting is summarized in 4.1. The online calibration within the geosensor network of cars is described in 4.2. First calibration results are given in section 5.

4.1 Derivation of a Base Calibration in Lab and Field

A basic W-R relationship can be derived in laboratory and field experiments. For the laboratory experiments, a sprinkler system has been designed and built (Fig. 2).

In order to produce different rainfall intensities, the system allows different combination of nozzles as well as different pressures on the

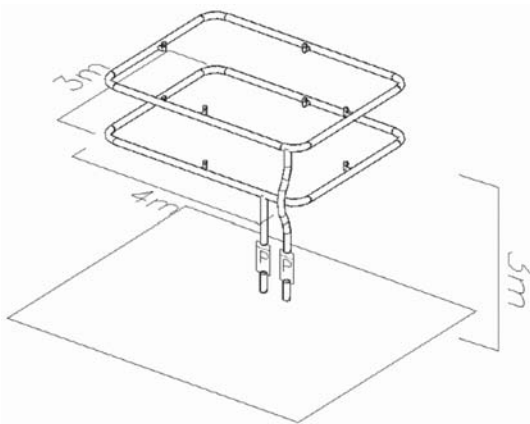


Fig. 2: Sprinkler system for the laboratory experiments.

two layers. The size of the simulator is sufficient for creating a homogeneous rainfield for a single car and all sensors under consideration at the same time. Rainfall intensities between 9 mm/h and 100 mm/h can be produced. As a reference for the car sensors, i. e. “ground truth”, both, a recording rain gauge with a tipping bucket sensor and a disdrometer are used. Tipping bucket sensors are widely used to provide point-rainfall measurements. Disdrometers give valuable additional information of the rainfall characteristics such as the rain droplet distribution or its falling speed. In order to investigate the effect of car speed on the optical devices, the two optical sensors are placed on a rotating machine. The Xanonex sensor is placed with an angle of 45° in rotating direction, in order to simulate an average windscreen angle. Due to minor variations in the rain intensity for each individual run during the experiments, each of the dynamic sensors is compared with a static one of the same type within a particular run. Currently, speeds of up to 45 km/h can be generated.

W-R functions can also be established in a real-world setting by using cars equipped with sensors for recording wiper activity, position and time with a particular sampling rate (Fig. 3). In this way, it is possible to investigate influencing factors that are not easily testable in the laboratory, such as wind speed or local environment. The main problem in deriving accurate W-R relationships from the field experiments is the availability of suitable and correct “ground truth”-data, i. e. reference rainfall at the car positions. Currently, data from the 11 stationary rain gauges in the Hannover area are used as reference, with an inverse distance weighted (IDW) interpolation to derive rainfall estimations at the car position. As the stationary rain gauges provide measurement every minute, the sensor readings of the cars (and the car positions) have been averaged over a minute as well. Currently, 10 cars are equipped with the wiper frequency sensor, among them 6 from a taxi fleet.

4.2 Geosensor Network of Cars

Instead of storing the sensor readings and evaluating them in a postprocessing or at a

central server, a W-R function could be directly applied to estimate rainfall while the car is driving, with the advantages as described in section 1. In addition, taking advantage of technologies of wireless sensor networks for vehicular networks, there is the possibility for the cars to exchange information with each other (Car2Car) or with stationary rain gauges equipped with communication facilities (Car2Station). Besides the general benefit of scalability described in section 1, this information exchange has two main advantages:

- a) Each car iteratively calibrates a W-R relationship that takes car specific factors into account that cannot be calibrated in advance, e. g. a particular windscreen angle or driver type.
- b) When exterior situations change, which are not (yet) respected in the W-R function, e. g. the current local wind speed and direction, the deviations from the estimation can be corrected on-the-fly and the model can be adjusted to the current situation.

As soon as a car enters the communication range of a stationary rain gauge (Car 1 in Fig. 3), a data exchange is established and the stationary rain gauge transmits its current rainfall measurement to the car. This results in a (*sensor readings, rainfall*)-sample that can

be used to update the current W-R function. When the car enters the communication range of another car (measurement 7 of Car 1), both cars exchange measurements (Car2Car) and improve their W-R function. The accuracy of the update samples is a function of:

- a) the original accuracy of the measurement or estimation at the location of the sending unit. Stationary rain gauges can be expected to provide higher accuracies compared to cars,
- b) the spatial distance of the sending and receiving unit and the temporal lag between the time, the receiver measured the predictor variables, e. g. the wiper frequency,
- c) the spatio-temporal variability of the current rainfall field.

While it is possible to determine a) and b), the estimation of c) is more difficult and subject to future work. For example radar estimations or the geosensor network of cars could be employed. The update accuracies can be taken into account; e. g. by methods such as weighted least squares (RAWLINGS et al. 1998).

The factors influencing W-R function estimations from Tab. 1 can be categorized as shown in Tab. 2. Here, the term “dynamic” re-

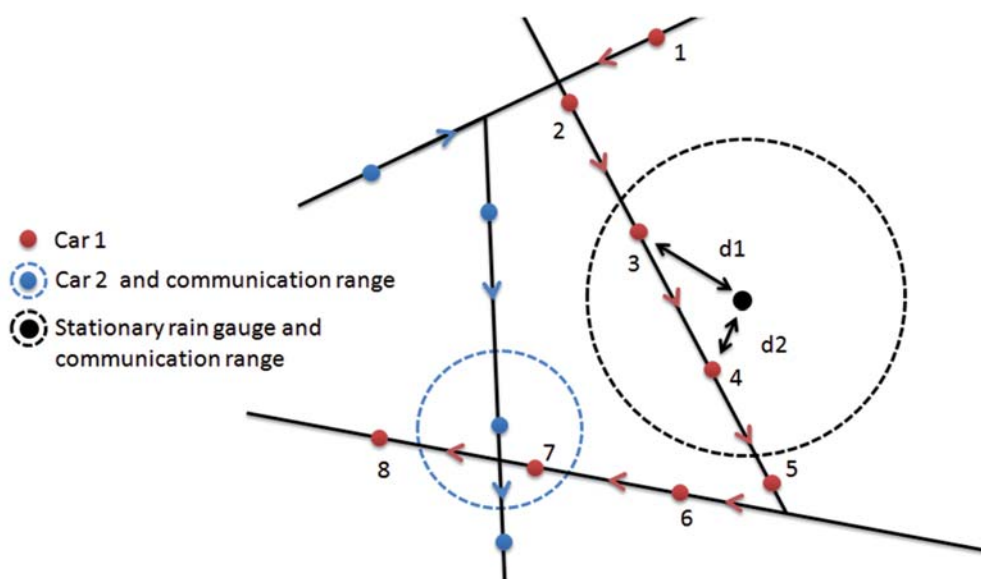


Fig. 3: Car2Car and Car2Station communication.

Tab. 2: Categorisation of influencing factors.

	Constant	Dynamic
Measured	-	Car speed, wiper frequency
Not (yet) Measured	Sensor system, windscreen angle (driver)	Wind direction and speed, local car environment, e. g. landuse

fers to factors that can change for a particular car while driving.

Factors that are measured by a car can function as parameters of the W-R function and hence their influence, i. e. coefficients, can be calibrated. This means, when the car receives a rainfall estimation from a neighbouring sensor (other car or station), it has typically a value for each measured parameter such as car speed or wiper frequency available. Therefore, an update results e. g. in a (*wiper frequency, car speed, received rainfall*)-sample that can be used to calibrate the model.

Some other factors, such as a specific sensor type or windscreen angle, remain constant for a particular car. This means, the calibration samples (*sensor readings, rainfall*) do not contain a reading for that particular factor, e. g. a particular value for a windscreen angle. However, with the online updates, the model is still implicitly adjusted to account for such constant factors. With a sufficiently large set of calibration samples collected over time, it can be expected that a specific car will generate a model that will work best on average for its car specific constant factors. Some influencing factors, such as the driver, change less frequently or sometimes even remain constant for a particular car and hence, cannot easily be categorized into constant or dynamic.

The calibration and application of a W-R function for online measurements requires that all factors are either:

- a) **dynamic** and **measured** by a car, a model parameter, e. g. car speed, or
- b) **static** for a particular car, e. g. windscreen angle, and **not necessarily measured**

Factors that are neither static nor measurable cannot be calibrated. This means, a particular

car cannot learn the influence of a particular dynamic factor that it cannot measure. Since there will most likely always be unknown factors, the online estimations of rainfall by the W-R functions will most likely always deviate from the true rainfall.

Even a W-R function perfectly calibrated for a set of predictors does not imply accurate individual rainfall estimations, it just estimates best on average given the available predictors. For example, in an area where the car is covered by obstacles preventing the rain at least partly to reach the car sensor, it is likely that any W-R function will underestimate rainfall. E. g. a wiper frequency of 30 wipes/min measured under normal conditions in open space will deliver a quite accurate rainfall estimation, i. e. the estimated rainfall for that wiper frequency will be close to the true rainfall. The same model and the same wiper frequency will lead to a significant underestimation of rainfall when the car is placed in an area such as a dense forest, where obstacles, e. g. trees, prevent the rain from reaching the windscreen.

These dynamic and not measureable factors include the car environment, the wind speed and the direction or the erratic wiper operation by the driver. As these factors are often spatially, e. g. car environment, temporally, e. g. current driver, or spatio-temporally correlated, e. g. wind speed and direction, they result in autocorrelated deviations of the W-R function estimations from the true (but unobservable) rainfall values. These deviations e. g. appear as autocorrelated residuals in the analyses (Fig. 7a). For example, if a car is travelling through a dense forest and hence underestimates the rainfall, it is quite likely that at a subsequent measurement, the car is still travelling under similar conditions and still underestimates. This can also be considered as the problem of dynamic model coefficients. Model fitting approaches that take this into account are e. g. rolling regression models or the Kalman filter (KALMAN 1960).

5 Experimental Results

In the following the results of the lab and field experiments will be presented.

5.1 Laboratory Calibration Results

The red points in Fig. 4 show the relationships between wiper frequency (x-axis) and reference rainfall measured by a tipping bucket (y-axis) in the laboratory, averaged over a minute, and a linear function fitted with OLS. Fig. 4a shows the results for the W-R relationship of a car with automatic wiper option. Fig. 4b shows the result for the wipers of the same car being adjusted completely manually. This means, a single human operator initiated a single wipe each time, his visibility was impaired. The dashed lines in both figures illustrate the 95% prediction interval.

The figures demonstrate that there is a significant correlation between wiper frequency and rainfall. It can be observed that the manually operated wipers lead to higher coefficients of determination than the automatic wipers. The conclusion drawn from this is that wiper actions triggered by the desire for a clear visibility indeed is an indication for the rainfall. Thus, the better a person or an automatic system fulfils this desire, the better the correlation with rainfall is.

In this way, the laboratory experiments support the projects' underlying assumption that there is, indeed, a relationship between visibility and rainfall intensity. Additional experiments have been conducted to investigate the influence of speed. There are different factors effecting the general overestimation of the rainfall when the sensor is moving:

- the sensor speed,
- the shape and angle of the optical sensor,

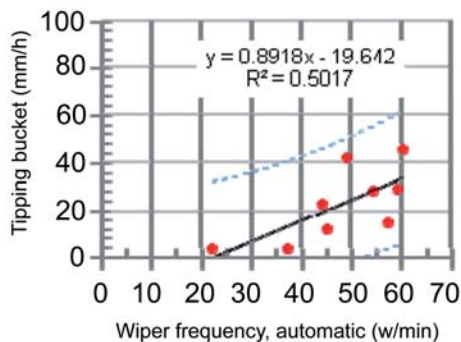
- the rainfall type, e.g. heavy rain, light rain, etc., corresponding to droplet falling speed.

Theoretically, there is a positive linear relationship between the velocity of an object in rain and the water mass hitting the object (BOCCI 2012). The slope of the relationship is a function of the rainfall type and windscreen angle. The red and blue lines in Fig. 5 illustrate the theoretical relationships according to BOCCI (2012) for two droplet falling speeds of 2 m/s and 5 m/s, respectively and a flat sensor placed 45° towards the direction of movement.

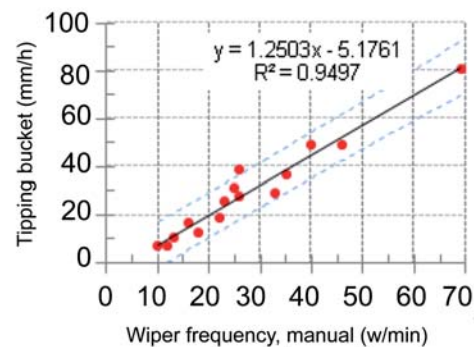
The green line in Fig. 5 illustrates the empirical result. It shows the ratio, dynamic divided by static sensor readings, between a dynamic sensor and a static one of the same type (Xanonex), averaged over different rain intensities. In contrast to the theoretical curves, the result is not a linear increment but becomes static at approximately 20 km/h. Possible reasons are a) the amount of drops remaining on the sensor surface which changes with increasing speed and b) the centrifugal force on the drops, originating from the rotation of the sensor.

5.2 Field Calibration Results

A set of 6 cars with manually operated wipers has been equipped with frequency measurement sensors. In total, around 36 hours of car data with substantial wiping activity has been collected (~2200 wiper frequency measure-



(a) W-R relationship for wiper frequency of automatic wipers



(b) W-R relationship for completely manual initiation of single wipes

Fig. 4: W-R (wiper-rainfall) functions determined in laboratory experiments.

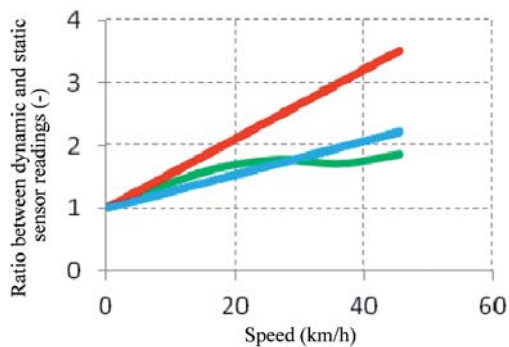


Fig. 5: Influence of speed on optical sensor readings.

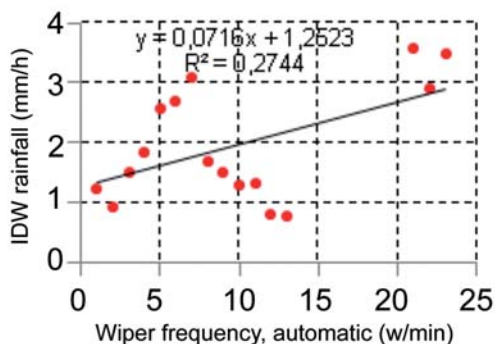
ments with a sampling rate of 1 min). 4 cars with automatically controlled wipers have been equipped with frequency sensors, resulting in 30 hours of wiper frequency measurements with one minute resolution (~1800 min). This sample has been preprocessed in order to produce specific selection sets, which allow for a separate investigation of influencing factors. Further, only car measurements have been selected, where the distance to a rain gauge is lower than a given maximum. For the first experiments, the distance has been set to 4000 m, since it seemed to provide a good trade-off between the number of samples and the interpolation accuracy. In order to further reduce errors in the car measurements, e. g. due to windscreen cleaning, only those car trajectories have been evaluated with a substantial and long wiping activity.

Whereas this dataset allows for first analyses presented in the following, the number of

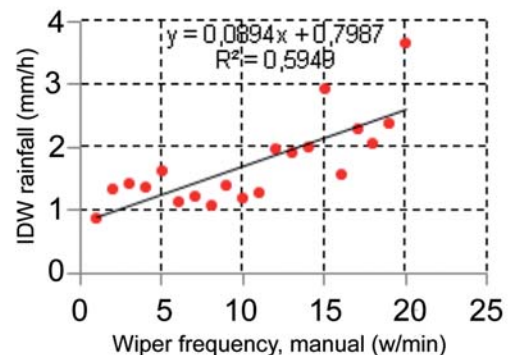
data points is too low to perform reliable fine-grained analysis of additional factors such as the influence of different car installations (car types, drivers) on the W-R function.

For the analyses, the mean IDW rainfall for each measured wiper frequency is calculated, i.e. the rainfall conditioned on the wiper frequency. In this way, the set of samples for model fitting is reduced such that there is a single sample for each wiper frequency, represented by the individual red points in Fig. 6. A linear regression is fitted to the data, representing the W-R function for automatically (a) and manually controlled wipers (b). Here, manually refers to the manual selection and adjustment of the discrete wiper classes (e. g. slow, medium, fast) in cars without a rain sensor controlling the wiper frequency. As the conditional mean values are plotted, no confidence or prediction intervals are provided. In contrast to the laboratory experiments, the R^2 values are lower – and they are even significantly lower when all samples are plotted. The major reason for this is that influencing factors are contributing to the function, which are not yet respected. Furthermore, the quality of the interpolated rainfall is not yet known. Although the distance to the stations is limited to 4000 m, this distance still might be too large.

The analysed car data shows a clear indication of multicollinearity, i.e. a linear relationship between the predictor variables wiper frequency and speed (RAWLINGS et al. 1998). This can be explained by the way, in which the wiper frequency adjustment in cars works. For periodic wiping, i.e. wiping with intervals



(a) Conditional mean values and fitted linear W-R relationship for a car with automatically operated wipers



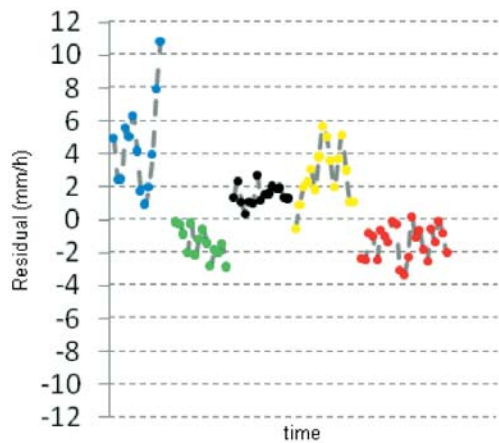
(b) Conditional mean values and fitted linear W-R relationship for a car with manually operated wipers

Fig. 6: W-R functions determined in field experiments.

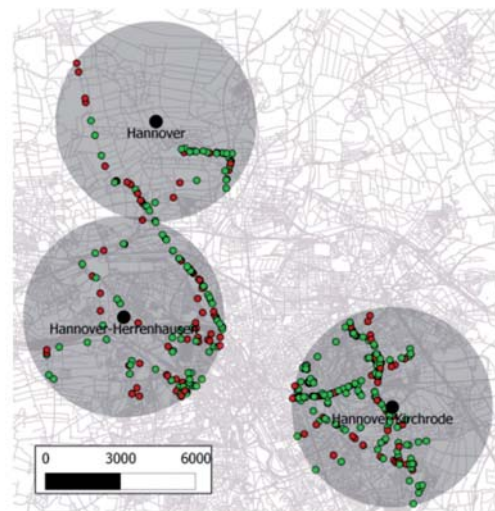
of no wiping in between the single wipes, the wiper frequency increases with increasing car speed, i.e. the interval in between single wipes get shorter. Therefore, certain wiper frequencies occur preferably or even only at certain car speeds. This makes a fine-grained analysis of the correlated predictor variables difficult. In addition, it seriously affects the reliability of the ordinary least squares estimates. This has to be investigated further. This dependence between car speed and wiper frequency also provides the explanation why in Fig. 6b are no discrete frequency classes, e.g. fixed

frequencies corresponding to slow, medium and fast.

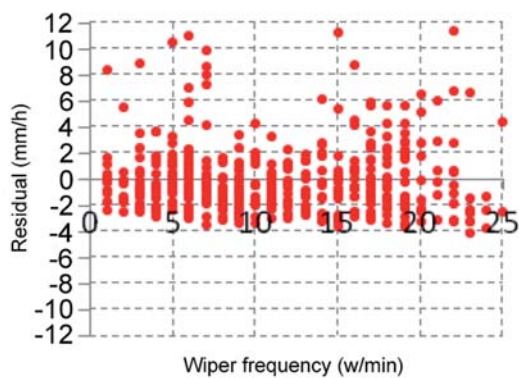
Fig. 7a shows the residuals, assumed true rainfall – estimated rainfall, against the time for a single car with manually operated wipers and a linear model fitted to all samples using OLS with the two predictor variables wiper frequency and car speed. Different colours indicate different time periods. Residuals within a time period are separated by 1 min. Residuals of different time periods are separated by an arbitrary time, e.g. hours, days or weeks. The residual autocorrelation can be clear-



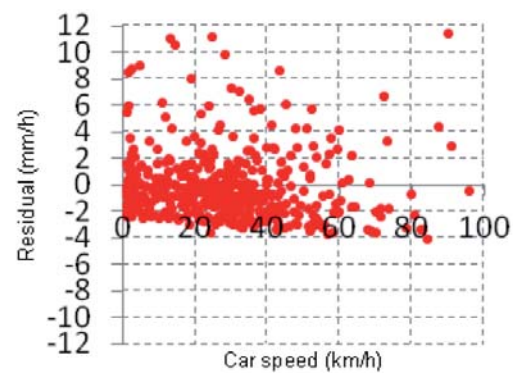
(a) Time ordered (left to right) residuals for a single car with manually controlled wipers and a fitted linear model with predictors speed and wiper frequency. Serial correlation in the residuals that are separated by 1 minute (connected by grey lines) can be recognized.



(b) Positive residuals (red), corresponding to a rainfall underestimation by the W-R functions, and negative residuals (green), corresponding to a rainfall overestimation by the W-R functions. The data of all cars served as a basis with W-R functions established for each car individually. No spatial pattern can be recognized.



(c) Residuals vs. wiper frequency for all cars. No pattern can be recognized.



(d) Residuals vs. car speed for all cars. No pattern can be recognized.

Fig. 7: Analyses of data collected in the field.

ly recognized and is confirmed by a Durbin Watson test (DURBIN & WATSON 1951). As described in section 4.2, identifying the reasons for this residual autocorrelation is difficult. They can be related to the car environment, the current wiper operation by the driver or dynamic factors such as wind speed and direction. In the case, that the car environment is the reason, it can be expected that the deviations will highly correlate with the spatial location where a particular car collected the sample. Fig. 7b shows over- and underestimations of the W-R functions of all cars distributed over the area. However, in this area (city environment) no spatial pattern can be recognized, thus static spatial factors, such as a particular car environment are unlikely to be the reason for the deviations.

Another possible source for autocorrelated residuals are wrongly specified models, i.e. wrong functional forms of the estimation functions (RAWLINGS et al. 1998). This can be analyzed in plots of the relationship between the residuals and the predictor variables. Figs. 7c and 7d show the relationship between residuals and wiper frequency, respectively the car speed. As no clear pattern can be recognized, the linear model is currently considered as suitable. However, further experiments are required, especially since the residuals also show a clear indication of deviation from normality.

6 Summary and Outlook

This paper presents investigations on rainfall estimation with cars. Laboratory and field experiments are presented and different possibilities for calibration are discussed. Initial calibration results of the estimation functions are given.

There are several issues which need further research. Currently, the W-R functions are assumed to be linear. Especially with the introduction of new predictor variables into the W-R functions, it might turn out that this assumption will be not valid anymore and that other functional forms, e. g. non-linear models, will be required. In addition to the analyses presented here, experiments with different function estimation techniques will be car-

ried out. For example, different weights could be chosen for the samples, depending on the proximity to the next stationary rain gauge. In addition, other interpolators and the quality of the interpolations will be investigated.

Further, computer simulations including also real car trajectory data will have to prove that the online calibration indeed provides improved rainfall estimations by the cars. For these experiments, different models, such as autoregressive schemes or the Kalman filter, will be investigated and a rigorous model for the data will be developed. Further, it might turn out that an interpolated rainfield is not sufficient for assessing the benefit of the online adjustment, due to the smoothing effect. A different ground truth, such as radar data, might be considered.

Another topic is the determination of suitable communication ranges between the sensors. The communication ranges are determined by the devices, however, the inclusion of a measured value has to be restricted by the homogeneity of the current rainfield. An idea is to quantify this homogeneity and in general to detect patterns in the precipitation fields in the geosensor network itself (SESTER 2009).

Acknowledgements

This research was funded by the German Research Foundation (DFG). The authors wish to thank BASTIAN HEINRICH for programming and producing the circuits of the optical sensors.

References

- BEVEN, K., 2001: Rainfall-Runoff Modelling: The Primer. – 360 p., John Wiley & Sons.
- BOCCI, F., 2012: Whether or not to run in the rain. – *European Journal of Physics* **33**: 1321–1332.
- CHIANG, Y.-M., HSU, K.-L., CHANG, F.-J., HONG, Y. & SOROOSHIAN, S., 2007: Merging multiple precipitation sources for flash flood forecasting. – *Journal of Hydrology* **340**: 183–196.
- CHOW, V.T., MAIDMENT, D.R. & MAYS, L.W., 1988: *Applied Hydrology*. – McGraw-Hill.
- DUCKHAM, M., 2012: Decentralized spatial computing: Foundations of geosensor networks. – Springer, Berlin.

- DURBIN, J. & WATSON, G.S., 1951: Testing for Serial Correlation in Least Squares Regression. – II. *Biometrika* **38**: 159–177.
- EHRET, U., GÖTZINGER, J., BARDOSSY, A. & PEGRAM, G.G.S., 2008: Radar-based flood forecasting in small catchments, exemplified by the Goldersbach catchment, Germany. – *International Journal of River Basin Management* **6**: 323–329.
- FITZNER, D., SESTER, M., HABERLANDT, U. & RABIEL, E., 2012: Online Calibration of Measurement Devices in Geosensor Networks – An Example Application. – Seventh International Conference on Geographic Information Science, Columbus, Ohio.
- GOOVAERTS, P., 2000: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. – *Journal of Hydrology* **228**: 113–129.
- GOUDENHOOFDT, E. & DELOBBE, L., 2009: Evaluation of radar-gauge merging methods for quantitative precipitation estimates. – EGU General Assembly Conference Abstracts: 5272.
- GRIMES, D.I.F. & DIOP, M., 2003: Satellite-based rainfall estimation for river flow forecasting in Africa. – Rainfall estimates and hydrological forecasts. – *Hydrological sciences journal* **48**: 567–584.
- HABERLANDT, U., 2007: Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. – *Journal of Hydrology* **332**: 144–157.
- HABERLANDT, U. & SESTER, M., 2010: Areal rainfall estimation using moving cars as rain gauges – a modelling study. – *Hydrological and Earth System Science* **14**: 1139–1151.
- HASSE, L., GROSSKLAUS, M., UHLIG, K. & TIMM, P., 1998: A Ship Rain Gauge for Use in High Wind Speeds. – *Journal of Atmospheric and Oceanic Technology* **15**: 380.
- HILL, D.J. & FARZAN, F., 2012: Bayesian fusion for high-resolution rainfall sensing using pervasive sensors. – XIX International Conference on Water Resources, Urbana-Champaign, Illinois, USA.
- KALMAN, R., 1960. A New Approach to Linear Filtering and Prediction Problems. – *Transactions of the ASME – Journal of Basic Engineering* **82** (1): 35–45.
- KRAJEWSKI, W.F. & SMITH, J.A., 2002: Radar hydrology: rainfall estimation. – *Advances in Water Resources* **25**: 1387–1394.
- LEIJNSE, H., UILENHOET, R. & STRICKER, J.N.M., 2007: Rainfall measurement using radio links from cellular communication networks. – *Water Resource Research* **43**: W03201.
- MESSER, H., ZINEVICH, A. & ALPERT, P., 2006: Environmental Monitoring by Wireless Communication Networks. – *Science* **312**: 713–713.
- RAINSENSORS, 2013: <http://www.rainsensors.com/> (11. 2. 2013).
- RAWLINGS, J.O., PANTULA, S.G. & DICKEY, D.A., 1998: *Applied Regression Analysis: A Research Tool*. – Springer.
- SCHULZE, M.-J., BRENNER, C. & SESTER, M., 2010: Cooperative information augmentation in a geosensor network. – International Conference on Theory, Data Handling and Modelling in Geo-Spatial Information Science: 444–449, Hongkong, China.
- SESTER, M., 2009: Cooperative Boundary Detection in a Geosensor Network using a SOM. – International Cartographic Conference, Santiago de Chile.
- SMITH, J.A., BAECK, M.L., MEIERDIERCKS, K.L., MILLER, A.J. & KRAJEWSKI, W.F., 2007: Radar rainfall estimation for flash flood forecasting in small urban watersheds. – *Advances in Water Resources* **30**: 2087–2097.
- THAKURIAH, P. & TILAHUN, N. (Forthcoming): Incorporating Weather Information into Real-Time Speed Estimates: Comparison of Alternative Models. – Forthcoming in *Journal of Transportation Engineering*.
- WARDAH, T., ABU BAKAR, S.H., BARDOSSY, A. & MAZNORIZAN, M., 2008: Use of geostationary meteorological satellite images in convective rain estimation for flash-flood forecasting. – *Journal of Hydrology* **356**: 283–298.
- XANONEX, 2013: <http://www.xanonex.de/> (11. 2. 2013).
- YUTER, S.E. & PARKER, W.S., 2001: Rainfall Measurement on Ship Revisited: The 1997 PACS TEPPS Cruise. – *Journal of Applied Meteorology* **40**: 1003–1018.

Addresses of the Authors:

DANIEL FITZNER & Prof. Dr.-Ing. MONIKA SESTER, Institut für Kartographie und Geoinformatik, Tel.: +49-511-762-3589, Fax: +49-511-762-2780, e-mail: {daniel.fitzner}{monika.sester}@ikg.uni-hannover.de

Prof. Dr.-Ing. UWE HABERLANDT & EHSAN RABIEL, Institut für Wasserwirtschaft, Hydrologie und landwirtschaftlichen Wasserbau, Tel.: +49-511-762-2237, Fax: +49-511-762-3731, e-mail: {haberlandt}{rabiei}@iww.uni-hannover.de

Leibniz Universität Hannover, Appelstraße 9a, 30167 Hannover

Manuskript eingereicht: Dezember 2012
Angenommen: Januar 2013



Texturanalyse mittels diskreter Wavelet Transformation für die objektbasierte Klassifikation von Orthophotos

PHILIPP TOSCANI, MARKUS IMMITZER & CLEMENT ATZBERGER, Wien, Österreich

Keywords: texture, gray level co-occurrence matrices (GLCM), discrete stationary wavelet transformation (DSWT), random forest, land cover classification

Summary: *Wavelet-based texture measures for object-based classification of aerial images.* Object-based approaches permit the use of textural information in addition to the spectral signatures. However, it is not yet clear which textural features and which spatial resolution offer the highest information content. The aim of the research was to evaluate textural features based on discrete stationary wavelet transformation (DSWT) against metrics from gray level co-occurrence matrices (GLCM) within a typical land cover classification setting using 4-band UltraCam data with up to 20 cm ground resolution. The study clearly demonstrated the superiority of wavelet based approaches compared to classical GLCM-features both in terms of classification accuracy and necessary computer resources. The random forest (RF) based classification of three broad land cover classes (forest, open land and pinus mugo plus shadows) of the Bavarian Alps (Germany) using the coiflet-wavelet-family in addition to the spectral data gave for example an overall classification accuracy >90% with all individual class accuracies (user and producer accuracies) being >88%. Using the spectral data alone, the corresponding overall classification accuracy was 87%. The positive effect of the textural measures was particularly noticeable with the highest ground resolution. With decreasing resolution, the positive effect of the tested textures vanished and overall classification accuracies decreased. Both feature selection criteria offered by RF (mean decrease in accuracy and mean decrease in gini) permitted the identification of a small subset of features (~10) yielding similar or higher classification accuracies compared to the use of all textural features. This permits to significantly reduce the processing time for large scale applications.

Zusammenfassung: Bei der Klassifikation der Landbedeckung mittels objektbasierten Ansätzen bietet sich die Möglichkeit, die spektrale Information durch Textur beschreibende Parameter zu ergänzen. Dafür stehen verschiedenste Texturmaße zur Verfügung, welche zusätzlich auf unterschiedlicher räumlicher Auflösung berechnet werden können. In dieser Arbeit wurden auf Basis von diskreter stationärer Wavelet Transformation (DSWT) generierte Maße mit den weitverbreiteten Texturmaßen aus Gray Level Co-occurrence Matrizen (GLCM) verglichen. Für die Klassifikation der drei Landbedeckungsklassen Wald, Latsche und Offenland (+ Schatten) in den Bayerischen Alpen (Deutschland) wurden 4-Kanal-Orthophotos (UltraCam) mit 20 cm Bodenauflösung verwendet. Die DSWT-Maße zeigten deutliche Vorteile in Bezug auf Klassifikationsgenauigkeit und Rechenzeit. Bei der Klassifikation mittels Random Forest (RF) anhand spektraler Information und Kennwerten der Coiflet-Wavelet-Familie wurden Gesamtgenauigkeiten von über 90% erreicht. Dabei lagen alle klassenspezifischen Nutzer- und Produzentengenauigkeiten über 88%. Der positive Effekt der Texturmaße war bei der Klassifikation der Bilder mit der höchsten räumlichen Auflösung am deutlichsten und nahm mit abnehmender Auflösung bedeutend ab. Zusätzlich konnte gezeigt werden, dass sich die bei der Klassifikation mit RF ermittelten Wichtigkeitsmaße (Mean Decrease in Accuracy und Mean Decrease in Gini) für eine Variablenselektion eignen, wobei Mean Decrease in Accuracy Vorteile zeigte. Die Klassifikationsgenauigkeiten bei Verwendung der 10 wichtigsten Texturmaße lagen immer über den mit allen Variablen erzielten Werten. Die Variablenselektion verringert zusätzlich die Rechenzeiten, was vor allem bei großflächigen Anwendungen Vorteile bringt.

1 Einleitung

Objektbasierte Ansätze zeigen insbesondere bei der Klassifikation von räumlich sehr hochaufgelösten Fernerkundungsdaten Vorteile gegenüber pixelbasierten Verfahren (JÄHNE 2005, KAYITAKIRE et al. 2006, CASTILLA & HAY 2008, WASER et al. 2010, IMMITZER et al. 2012a). Das Zusammenfügen von ähnlichen und benachbarten Bildelementen zu Segmenten, welche realweltlichen Objekten entsprechen, ermöglicht neben der spektralen Information die Verwendung von zusätzlicher, das Objekt beschreibender Information für die Klassifikation (BLASCHKE & STROBL 2001, HAY et al. 2005, BLASCHKE 2010). Neben geometrischen Parametern wie Größe und Form kann vor allem die Oberflächenstruktur (Textur) bei der Klassifizierung von segmentierten Bildern einen wesentlichen Beitrag leisten (RENGERS & PRINZ 2009, VIEIRA et al. 2012). Das Ziel ist die Nachahmung bzw. mathematische Beschreibung der menschlichen Wahrnehmung eines Interpreten (ZHU & YANG 1998). Ein zusätzlicher Vorteil von objektbasierter Klassifikation ist der meist geringere Daten- und Rechenaufwand (JÄHNE 2005).

Etablierte Ansätze für die Beschreibung der Objekttextur basieren auf der statistischen Analyse der Grauwerte des Bildes. Ein guter Überblick zu Grundlagen, Modellen und Verwendung in der Landbedeckungsklassifikation findet sich in RENGERS & PRINZ (2009). Der in der Fernerkundung am weitesten verbreitete Ansatz stammt von HARALICK et al. (1973). Sie beschreiben 14 Texturmaße, welche mit Hilfe von Gray Level Co-occurrence Matrizen (GLCM) berechnet werden. Für die Berechnung der GLCM wird die Nachbarschaft eines Pixels in einer Aufnahmezelle betrachtet, gegebenenfalls in unterschiedlichen Distanzen, und die Zusammenhänge in verschiedenen, meist vier, Richtungen berechnet. Von diesen Messungen werden texturale Kennwerte abgeleitet. Der Informationsgehalt dieser Texturmaße ist abhängig vom spektralen Bereich, der Auflösung und der Charakteristik der zu erfassenden Objekte (Dimension, Form und räumliche Verteilung) und dem analysierten Bild (KAYITAKIRE et al. 2006). Verschiedene Studien zeigen, dass einige der 14 Texturparameter redundante Information lie-

fern (HARALICK et al. 1973, BARALDI & PARMIGIANI 1995, SOH & TSATSOUKIS 1999, SHABAN & DIKSHIT 2001, METZLER et al. 2002). Ein Nachteil der Texturmaße nach Haralick besteht darin, dass die Texturanalyse in nur einem Maßstab durchgeführt wird, weshalb der Größe der Auswahlzelle eine hohe Bedeutung zukommt (ZHU & YANG 1998, SHABAN & DIKSHIT 2001).

Methoden basierend auf der Wavelet Transformation arbeiten hingegen auf der Basis von multiresolution-Analysen. Die Bilder werden dabei einer Frequenzanalyse in mehreren Auflösungen unterzogen. Die Wavelet Transformation stammt aus dem Gebiet der Signalanalyse. Dabei wird versucht, durch Transformation eines Signales Information zu gewinnen, welche als Maß für die Textur genutzt werden kann. Bei der Analyse des Signales wird die Fenstergröße variabel gehalten und an die auftretende Frequenz angepasst (LOUIS et al. 1998, ZHU & YANG 1998, POLIKAR 1999). Dadurch können unterschiedliche Oberflächenbeschaffenheiten von feiner bis grober Textur unterschieden werden (GRAPS 1995). Der Vorteil der diskreten stationären Wavelet Transformation (DSWT) besteht darin, dass durch das Fehlen von Up- und Downsampling-Schritten die Abmessungen des Ausgabebildes jenen des Eingangsbildes entsprechen (NASON & SILVERMAN 1995). Jede Auflösungsstufe gibt vier Layer mit Ergebnissen zurück: A-Bild (Approximation), H-Bild (horizontale Details), V-Bild (vertikale Details) und D-Bild (diagonale Details) (LIVENS et al. 1997, BUSCH & BOLES 2002). Die verschiedenen Ergebnisse liegen pixelweise vor und können daher ohne weiteres auf Objektebene aggregiert werden.

Ziel dieser Studie ist die systematische Untersuchung des Potenzials verschiedener Texturmaße für die objektbasierte Klassifikation von 4-kanaligen Orthophotos im Hochgebirge. Dabei soll die etablierte Methode der Texturmaße nach Haralick mit der bisweilen weniger verwendeten DSWT-basierten Texturanalyse im Hinblick auf die Klassifikationsgenauigkeit untersucht werden. Zusätzlich soll der Einfluss der räumlichen Auflösung der Orthophotos auf die Klassifikationsgenauigkeit analysiert werden. Wert wird auf eine unabhängige Validierung der Klassifikationsergebnisse sowie auf die räumliche Übertrag-

barkeit der verwendeten Methoden gelegt. Für Fragen der praktischen Umsetzbarkeit wird auch die Rechenzeit mit betrachtet.

2 Daten und Methodik

2.1 Projekthintergrund

Bei der großflächigen Erfassung und Bewertung von FFH-Lebensraumtypen in den Natura 2000 Gebieten können fernerkundliche Analysen einen wesentlichen Beitrag leisten. Die Studie ist Teil einer Arbeit, in welcher Fernerkundungsdaten für diesen Zweck im bayerischen Alpenraum eingesetzt werden (BINNER & SEITZ 2009). Das Institut für Vermessung, Fernerkundung und Landinformation der Universität für Bodenkultur (BOKU) Wien wurde von der Bayerischen Landesanstalt für Wald und Forstwirtschaft (LWF) mit der Entwicklung und Umsetzung einer großflächigen Klassifikation der Natura 2000 Gebiete des bayerischen Alpenraumes beauftragt. Ziel war die Erfassung der Klassen Offenland, Wald und Latsche in den FFH- und SPA-Schutzgebieten des bayerischen Hochgebirges mit einer Gesamtfläche von 96.000 ha. Die drei Klassen resultieren aus der Tatsa-

che, dass die LWF nur für die Kartierung der Waldflächen zuständig ist. Daraus ergab sich die Notwendigkeit, die aktuelle Waldfläche von der Offenland- sowie Latschenfläche abzugrenzen. Einzusetzen waren vorrangig die digitalen Orthophotos des Bayerischen Landesamtes für Vermessung und Geoinformation (LVG) (IMMITZER et al. 2012c).

Die flächendeckende Kartierung wurde zwischenzeitlich erfolgreich abgeschlossen. Für kleine Teilbereiche wurden methodische Fragestellungen untersucht, die hier dargestellt werden.

2.2 Untersuchungsgebiet und Datengrundlage

Aus den Natura 2000-Gebieten des bayerischen Hochgebirges wurde für diese Studie das Teilgebiet Karwendel östlich von Garmisch-Partenkirchen ausgewählt (Abb. 1). Für die Klassifikation standen 4 Kanal Orthophotos (Blau, Grün, Rot, Infrarot) mit einer Bodenauflösung von 20 cm zur Verfügung. Die Bilder wurden am 1. September 2009 mit einer UltraCamX aufgenommen. Für die Untersuchung wurden 3 Orthophotokacheln mit jeweils 4 km² Fläche verwendet. Die wesent-

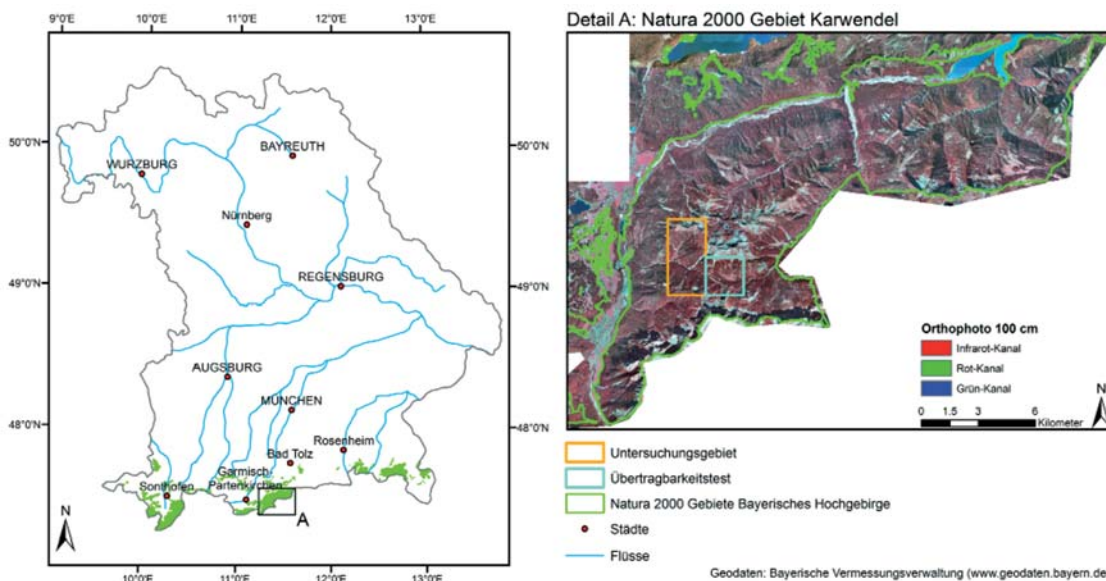


Abb. 1: Lage der Natura 2000 Gebiete des Bayerischen Hochgebirges (Untersuchungsgebiet des Projekts „Wald-Latsche-Offenland Abgrenzung“) und Detailausschnitt des Natura 2000 Gebietes Karwendel mit den verwendeten Orthophotokacheln.

lichen Analysen wurden auf 2 Kacheln durchgeführt, welche in Abb. 1 rechts orange markiert sind. Eine weitere Kachel (hellblau dargestellt) wurde lediglich für die Untersuchung der Übertragbarkeit herangezogen.

2.3 Segmentierung und visuelle Klassifikation

Die Segmentierung des Testgebietes erfolgte mit der Software Definiens unter Verwendung der Kanäle Grün, Rot und Infrarot. Die Ursprungsdaten wurden zu diesem Zweck auf 100 cm Bodenauflösung degradiert. Die für die multiresolution-Segmentierung (BAATZ & SCHÄPE 2000) geeigneten Parameter wurden wie für derartige Fragestellungen üblich (DRÄGUT et al. 2010) mittels Trial-and-Error Herangehensweise ermittelt. Die Qualität der Segmentierung wurde durch visuelle Kontrolle der Segmente bewertet. Dabei wurde vor allem auf die Klassenreinheit der einzelnen Segmente geachtet. Zusätzlich sollte die Übertragbarkeit der Segmentierungseinstellung auf Orthophotokacheln anderer Befliegungstage gewährleistet bzw. mit geringem Adaptionaufwand möglich sein. Mit folgenden Parametern wurde ein den Anforderungen entsprechendes Ergebnis erzielt: Scale =

40, Shape = 0,3 und Compactness = 0,5. Die drei zur Segmentierung verwendeten Kanäle wurden wie folgt gewichtet: Infrarot = 4, Rot = 1 und Grün = 1.

Die Segmentierung ergab 4.681 Polygone mit einer Größe von $3 \text{ m}^2 - 26.849 \text{ m}^2$. Die am Ausgangsbild mit 100 cm Bodenauflösung generierten Polygone wurden auch für die Analysen mit höherer und mit niedrigerer Auflösung verwendet. Um die Übertragbarkeit sicherzustellen, wurden Polygone mit einer Fläche $< 100 \text{ m}^2$ eliminiert und dem benachbarten Objekt mit der längsten gemeinsamen Grenze klassenunabhängig zugeordnet. Durch die Eliminierung ergaben sich 4.549 Polygone mit einer Größe von $100 \text{ m}^2 - 26.849 \text{ m}^2$.

Anschließend wurden alle Polygone visuell einer der sechs Klassen Nadelwald, Laubwald, Latsche, Offenland Gras, Offenland Fels bzw. Schatten zugeordnet. In Abb. 2 sind für die einzelnen Klassen Beispielpolygone dargestellt. Dieser Datensatz stellt die Referenzinformation dar, mit der die Klassifikation durchgeführt und validiert wurde. Für zusammenfassende Analysen wurden die Klassen Nadel- und Laubwald zur Klasse Wald und die Klassen Offenland Gras bzw. Offenland Fels zu Offenland entsprechend den Projektanforderungen (Kapitel 2.1) zusammengefasst.

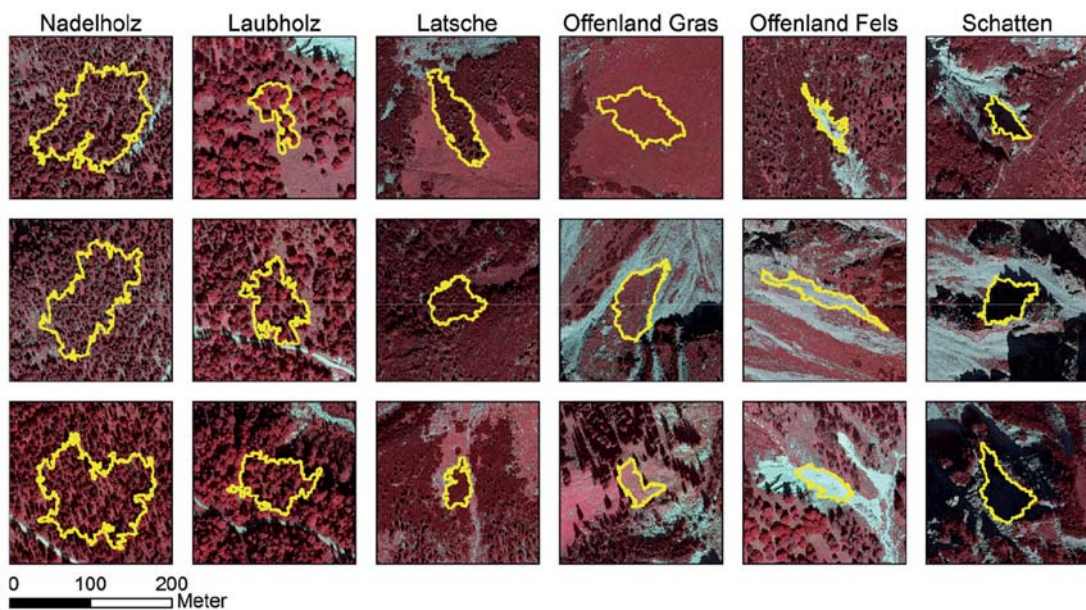


Abb. 2: Beispieloobjekte der 6 untersuchten Klassen (CIR Darstellung aus den Kanälen Infrarot, Rot und Grün).

Tab. 1: Aufteilung der Objekte auf die 6 Klassen und deren objekt- bzw. flächenbezogene Anteile sowie der Mittelwert und Median der Objektgrößen je Klasse des gesamten Testgebietes.

Klasse	Objekte	Anteil Objekte	Fläche (m ²)	Anteil Fläche	Mittelwert der Objektgröße (m ²)	Median der Objektgröße (m ²)
Nadelholz	1.166	0,26	2.752.042	0,34	2.360,2	1.508
Laubholz	490	0,11	954.202	0,12	1.947,4	1.412
Latsche	772	0,17	1.388.250	0,17	1.798,3	1.470
Offenland Gras	1.109	0,24	1.657.608	0,21	1.494,7	1.275
Offenland Fels	544	0,12	701.066	0,09	1.288,7	1.148
Schatten	468	0,10	546.832	0,07	1.168,4	1.125
Summe	4.549		8.000.000		1.758,6	1.322

Die Aufteilung der Objekte auf die einzelnen Klassen zeigt Tab. 1. Die Klassen Nadelholz und Offenland Gras weisen die höchsten Anteile auf, sowohl bezogen auf die Objektanzahl als auch auf die bedeckte Fläche. Die durchschnittliche Objektgröße ist bei den beiden Waldklassen (Nadelholz und Laubholz) am größten. Die Objekte der Klassen Offenland Fels und Schatten sind durchschnittlich am kleinsten.

Die Orthophotokachel für den Test der Übertragbarkeit wurde mit denselben Einstellungen segmentiert und ebenfalls visuell klassifiziert.

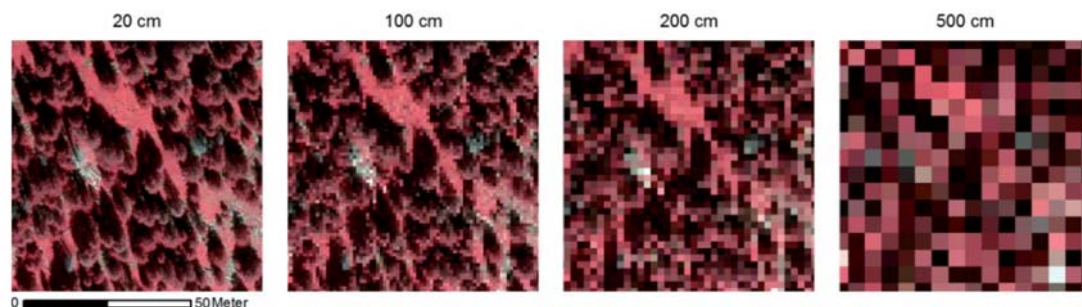
2.4 Erstellung der Eingangsvariablen

Für die Verringerung der Rechenzeiten und zur Erleichterung der Datenmanipulation bietet sich für sehr großräumige Analysen die Verringerung der räumlichen Auflösung der verwendeten Daten an (SCHLERF et al. 2003). Aus diesem Grund wurden aus den Aus-

gangsdaten (20 cm Bodenauflösung), Bilder mit 100 cm, 200 cm und 500 cm berechnet (Abb. 3). Alle Eingangsvariablen wurden für die 4 Auflösungen berechnet und für die Klassifikation parallel untersucht.

Die zur Klassifizierung verwendeten Variablen können vier Gruppen zugeordnet werden (Abb. 4). Beim ersten Verfahren „*Spektrale Information*“ wurden die Grauwerte der einzelnen Pixel innerhalb eines Segmentes zur Merkmalsgewinnung herangezogen, z. B. mittlerer Grauwert des IR-Kanals.

Beim zweiten Verfahren „*Texturmaße, Bezugsfläche Segment*“ wurde aus den einzelnen Pixeln der Segmente eine Grauwertmatrix (GLCM) erstellt, aus welcher dann die objektbasierten Texturparameter, z. B. *Energy*, ermittelt wurden. Beim dritten Verfahren „*Texturmaße, Moving Window*“ wurde in einer Zwischenberechnung zunächst die Information über die Nachbarschaft eines Pixels innerhalb einer quadratischen Bezugsfläche gesammelt, daraus eine GLCM erstellt und dann die Texturparameter, z. B. *Energy*, dem zen-

**Abb. 3:** Detailausschnitt in den verwendeten Auflösungen.

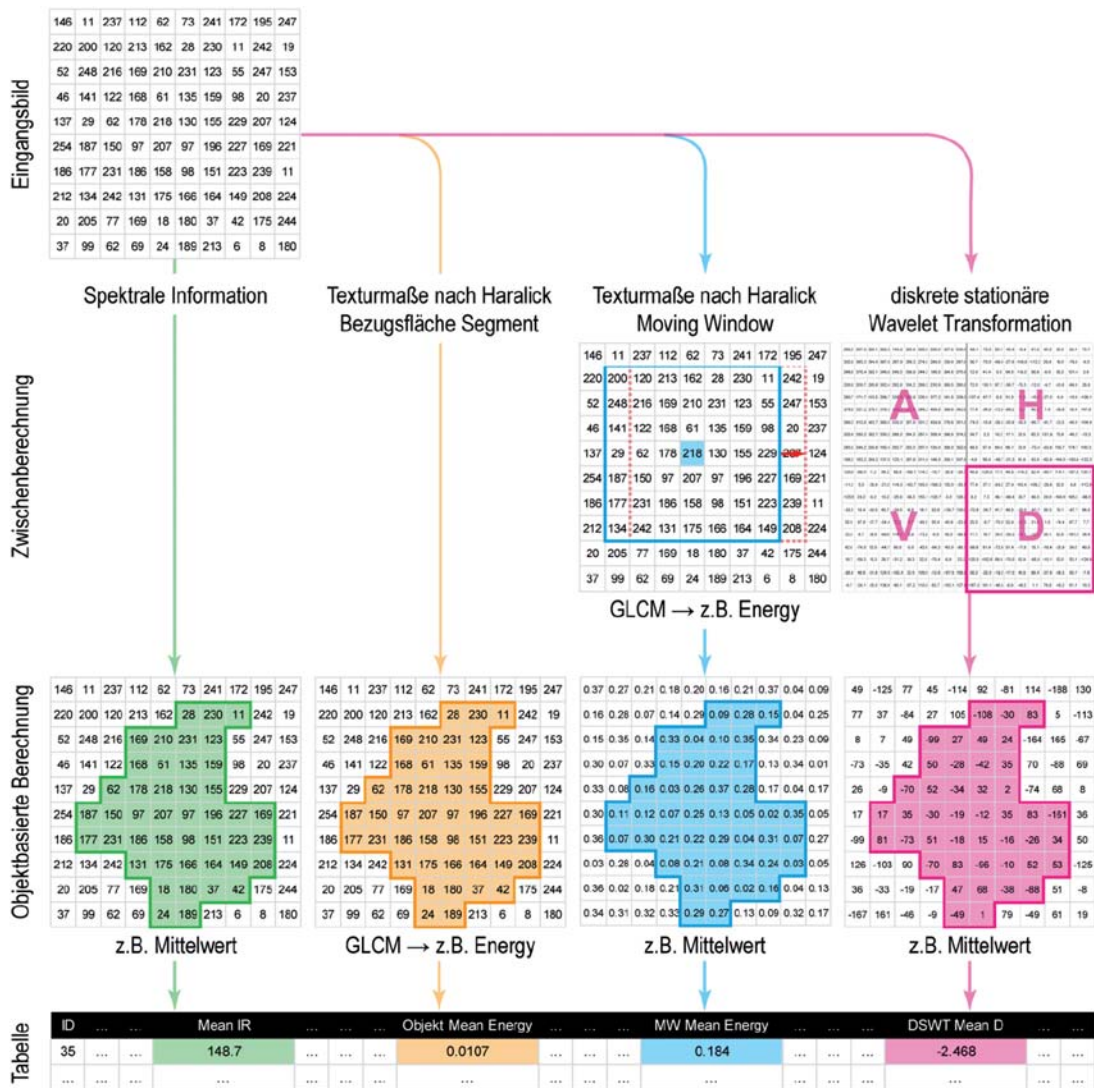


Abb. 4: Ablaufschema für die Berechnung der objektbezogenen Variablen (GLCM = Gray-level co-occurrence matrix).

tralen Pixel zugewiesen. Nach Verschiebung der Bezugsfläche um ein Pixel (Moving Window) wurde die Berechnung für das nächste zentrale Pixel vorgenommen. Dadurch wurde ein neuer Layer mit den Abmessungen des Eingangsbildes generiert. Dessen Merkmale wurden im nächsten Schritt wieder objektweise aggregiert, z. B. Mittelwert. Beim vierten Verfahren wurde das Eingangsbild in einer Zwischenberechnung einer „diskreten stationären Wavelet Transformation“ unterzogen. Diese liefert pro Transformationsdurchgang vier Layer mit den Abmessungen des Ein-

gangsbildes, welche im nächsten Schritt wieder objektbasiert ausgewertet wurden, z. B. Mittelwert.

2.4.1 Haralick, Bezugsfläche Segment (Haralick Segment)

Die Berechnung der GLCM für die einzelnen Objekte erfolgte in MATLAB (2011) getrennt für die vier Spektralkanäle mit folgenden Einstellungen: Lag-Distanz 1, Richtungen 0°, 45°, 90° und 135°, NumLevels (NL) 16, GrayLimits 0 – 255, Symmetric true. Aus den

vier GLCM je Spektralkanal wurden Mittelwert, Maximalwert und Standardabweichung der Texturparameter *Contrast*, *Correlation*, *Energy* und *Homogeneity* für jedes Objekt berechnet ($4 \text{ Kanäle} \times 4 \text{ Texturmaße} \times 3 \text{ Statistische Kennwerte im Objekt} = 48 \text{ Variablen pro räumlicher Auflösung}$).

2.4.2 Haralick, Moving Window (Haralick MW)

Für die Berechnung der GLCM wurden die gleichen Einstellungen wie bei der Bezugsfläche Segment verwendet. Die GLCM wurde jedoch nicht für die einzelnen Segmente, sondern in quadratischen Moving Windows (MW) erstellt. In Anlehnung an die Arbeiten von SHABAN & DIKSHIT (2001) und RODRIGUEZ-GALIANO et al. (2012a) wurden MW-Größen von 35×35 und 51×51 Pixel verwendet. Aus den vier GLCM je Spektralkanal wurden der Mittel- und Maximalwert der Texturparameter *Contrast*, *Correlation*, *Energy* und *Homogeneity* berechnet und dem zentralen Pixel des MW zugewiesen. Im nächsten Rechenschritt wurden die Mittelwerte, Maximalwerte und die Standardabweichung der Texturparameter direkt im Objekt berechnet ($4 \text{ Kanäle} \times [4 \text{ Texturmaße} \times 2 \text{ Statistische Kennwerte}] \times 3 \text{ Statistische Kennwerte im Objekt} = 96 \text{ Variablen pro räumlicher Auflösung}$).

2.4.3 Diskrete Stationäre Wavelet Transformation (DSWT)

Die Eingangsbilder wurden getrennt nach den einzelnen Spektralkanälen in MATLAB (2011) einer diskreten stationären Wavelet Transformation (DSWT) unterzogen. Dabei wurde aus folgenden fünf Wavelet-Familien je ein Wavelet (Bezeichnungen in Klammer) getestet: *Biorthogonale Wavelets (bior1.1)*, *Coiflets (coif1)*, *Daubechies Wavelets (db1)*, *Discrete Meyer (dmey)* und *Symlets (sym2)*. Wavelet-Familien unterscheiden sich hinsichtlich ihrer Form, dem Frequenzbereich und ihrer Maßstäbe (GRAPS 1995). In Abhängigkeit der Bilddimensionen war eine unterschiedlich hohe Anzahl von Transformationsdurchgängen möglich. Für die Bodenauflösungen 500 cm und 200 cm wurden zwei Durchgänge, für 100 cm und 20 cm vier Durchgänge

gewählt. Für jeden Spektralkanal und jeden Transformationsdurchgang entstehen vier Ergebnisbilder: Approximation (A), horizontale Details (H), vertikale Details (V), diagonale Details (D). Das A-Bild des Durchganges i dient dabei als Eingangsbild für den Durchgang $i+1$. Von jedem Ergebnisbild wurde für jedes Objekt Mittelwert, Maximalwert und Standardabweichung berechnet ($4 \text{ Kanäle} \times 4 \text{ Durchgänge} \times 4 \text{ Ergebnisbilder} \times 3 \text{ Statistische Kennwerte im Objekt} = 192 \text{ Variablen pro räumlicher Auflösung bei } 20 \text{ cm und } 100 \text{ cm}$; $4 \text{ Kanäle} \times 2 \text{ Durchgänge} \times 4 \text{ Ergebnisbilder} \times 3 \text{ Statistische Kennwerte im Objekt} = 96 \text{ Variablen bei } 200 \text{ cm und } 500 \text{ cm}$).

2.5 Klassifikation und Variablenselektion mittels Random Forest

Für die Klassifikation wurde der von BREIMAN (2001) entwickelte Random Forest Klassifikator (RF) verwendet. RF ist ein nicht parametrisches Verfahren, welches zuverlässige Klassifikationsergebnisse bei geringer Feinabstimmung liefert (LIAW & WIENER 2002). Das Verfahren wird zur Gruppe der Ensemble learning Klassifikatoren gezählt. Dabei wird durch die Kombination mehrerer Klassifikatoren die Erkennungsleistung gesteigert (WASKE & BRAUN 2009, KROSCHER et al. 2011). RF stellt somit eine wesentliche Weiterentwicklung von traditionellen Entscheidungsbäumen dar (RODRIGUEZ-GALIANO et al. 2012b).

Die Vorteile dieses Verfahrens sind (BREIMAN 2001, 2002, WASKE & BRAUN 2009, HASTIE et al. 2009, RODRIGUEZ-GALIANO et al. 2012a):

- die Klassen müssen nicht normalverteilt und können auch multimodal sein,
- hochdimensionale und redundante Datensätze können effizient verarbeitet werden,
- robustes Verhalten bei der Reduktion der Trainingsdaten,
- neigt nicht zum Overfitting,
- auch bei stark heterogenen Klassen sind zuverlässige Ergebnisse zu erwarten.

Für die Durchführung der RF Klassifikation wird aus dem Referenzdatensatz durch zufälliges Ziehen mit Zurücklegen (Bootstrapping) ein Trainingsdatensatz mit der Größe des Re-

ferenzdatensatzes generiert. In diesem Datensatz sind ca. 2/3 der Referenzdaten enthalten, mit welchen ein Entscheidungsbaum erstellt wird. Die nicht gezogenen Daten (out-of-bag, OOB) dienen als unabhängiger Validierungsdatsatz und werden mit dem erstellten Entscheidungsbaum klassifiziert. Damit jedes einzelne Objekt mehrmals im OOB-Datensatz enthalten ist und somit klassifiziert wird, wird dieser Vorgang oftmals wiederholt (Anzahl an Entscheidungsbäumen). Neben der Anzahl der Entscheidungsbäume ist für die Erstellung einer RF Klassifikation auch noch die Anzahl an Variablen pro Entscheidungsknoten festzulegen. Im Unterschied zu herkömmlichen Entscheidungsbäumen werden bei RF nicht alle Variablen verwendet, sondern nur eine zufällige Auswahl (BREIMAN 2001). Dadurch kommt es zu einer Reduktion der Korrelation zwischen den Entscheidungsbäumen (HASTIE et al. 2009). Der Standardwert für die Anzahl an Variablen wird aus der Wurzel aller verfügbaren Variablen berechnet (LIAW & WIENER 2002). Nachdem alle Bäume erstellt wurden, weisen alle Objekte mehrere Klassifikationsergebnisse auf. Aus diesen wird letztendlich der Modus (majority vote) dem jeweiligen Objekt zugeordnet. Durch den Vergleich mit der tatsächlichen Klasse erhält man einen unverfälschten Schätzwert der Klassifizierungsgenauigkeit. Für die Annäherung dieses Wertes an die tatsächliche Klassifizierungsgenauigkeit ist eine relativ hohe Anzahl an Entscheidungsbäumen (> 500) sinnvoll (BREIMAN 2001, HASTIE et al. 2009).

Da Variablen ohne Aussagekraft zu einer Verschlechterung des Ergebnisses führen können (HASTIE et al. 2009), ist in bestimmten Fällen eine Variablenselektion sinnvoll. Dazu bieten sich die von RF gelieferten Aussagen über die Wichtigkeit der einzelnen Variablen für die Klassifikation an. Die zwei gängigsten Maße sind „Mean decrease in Gini“ (MDG) und „Mean decrease in accuracy“ (MDA). Für die Berechnung der MDA-Wichtigkeit einer Variablen werden deren Werte in den OOB-Daten zufällig verändert. Alle anderen Variablen bleiben dabei unverändert. Anschließend erfolgt die erneute Klassifikation der OOB-Daten. Der Vergleich der dabei erzielten Klassifizierungsgenauigkeit mit der ursprünglichen Genauigkeit ergibt die Wichtigkeit der jewei-

ligen Variablen. Dieser Vorgang wird für jede Variable wiederholt. MDG errechnet sich aus dem Anstieg des Gini-Trennwertes (Trennmaß in den jeweiligen Entscheidungsknoten) hervorgerufen durch die jeweilige Variable. Die Werte der einzelnen Entscheidungsknoten werden dabei innerhalb des Baumes summiert und über alle Bäume normalisiert (BREIMAN 2002).

2.6 Testablauf

Für die Klassifikation wurde der Datensatz entsprechend der Objektanzahl der einzelnen Klassen in zwei gleichgroße Teildatensätze geteilt. Anschließend wurden mit einem Teil die RF-Modelle erstellt, inklusive der Ermittlung der OOB Fehler und MDA/MDG Wichtigkeiten. Dabei wurden die spektrale Information alleine (12 Variablen), sowie kombiniert mit den 3 Gruppen der Texturmaße verwendet. Anschließend wurde mit den erstellten Modellen der zweite Teil des Datensatzes als unabhängiger Validierungsdatsatz klassifiziert. Damit stehen neben dem OOB-Fehler des Trainingsdatensatzes auch Fehlermaße aus dem unabhängigen Validierungsdatsatz zur Verfügung.

Zusätzlich zur Klassifikation mit allen Variablen erfolgte auch eine Klassifikation mit einem durch Variablenselektion reduzierten Datensatz, d. h. mit den 10 wichtigsten Texturvariablen nach MDA bzw. MDG. Die Texturvariablen wurden immer zusätzlich zu den 12 Variablen aus den Spektralkanälen verwendet. Dieser Vorgang, von der Aufteilung des Datensatzes bis zur Klassifikation, wurde 100-mal wiederholt. Aus den Ergebnissen der einzelnen Wiederholungen wurden Mittelwert und Standardabweichung der Gesamtgenauigkeiten berechnet. Die Unterschiede der einzelnen Verfahren wurden mittels ANOVA analysiert.

Für die Erstellung der flächendeckenden Kartendarstellung wurde das Majority Vote der 100 Klassifizierungswiederholungen verwendet. Diese Werte wurden auch für die Erstellung der Konfusionsmatrizen herangezogen.

Die Berechnungen der Texturmaße erfolgte mit der Image Processing bzw. der Wave-

let Toolbox in MATLAB 7.13.0 (MATLAB 2011). Für die Klassifikation und die Analysen der Ergebnisse wurde das Statistikprogramm R 2.15.0 (R DEVELOPMENT CORE TEAM 2012) mit dem Zusatzpaket randomForest in der Version 4.6 (LIAW & WIENER 2002) verwendet.

3 Ergebnisse

3.1 Bedeutung der spektralen Information und der Texturvariablen

Die Ergebnisse zeigen eine deutliche Verbesserung des Klassifikationsergebnisses, wenn zusätzlich zur spektralen Information Texturvariablen verwendet werden (Abb. 5). Der orange Balken stellt dabei die Gesamtgenauigkeit unter alleiniger Verwendung der spektralen Information dar (Orthophoto mit 100 cm Bodenauflösung). Alle getesteten Texturmaßgruppen konnten eine höchst signifikante Verbesserung der Klassifikationsgenauigkeit bewirken (ANOVA post-hoc Tukey's HSD-text $p < 0.01$).

Die Steigerung der Genauigkeit ist bei Verwendung der im Objekt generierten Texturmaße nach Haralick am geringsten und bei den mittels DSWT erstellten Maßen am höchsten. Zwischen den getesteten Wavelets sind nur geringe Unterschiede feststellbar, wobei die Ergebnisse, die mit DSWT *coif1*, *dmey* und *sym2* erzielt wurden, die höchsten Genauigkeiten zeigten.

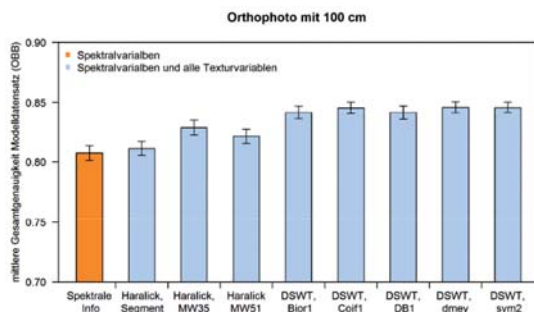


Abb. 5: Klassifikationsgenauigkeit (OOB) für 6 Klassen unter alleiniger Verwendung der spektralen Information (orange) und bei zusätzlicher Verwendung von Texturmaßen (Orthophoto mit 100 cm Auflösung).

3.2 Vergleich unterschiedlicher Auflösungen

Mit abnehmender Bodenauflösung der Eingangsdaten ist generell eine Verschlechterung der Klassifikationsgenauigkeit erkennbar. Der Einfluss hängt jedoch stark von den verwendeten Klassifikationsmerkmalen ab. Verwendet man beispielsweise für die Unterscheidung der 6 Klassen nur die spektrale Information, sind zwischen 100 cm und 200 cm keine Genauigkeitsunterschiede erkennbar. Das Ergebnis mit 20 cm liegt über, jenes mit 500 cm deutlich darunter. Wesentlich deutlicher sind die Unterschiede, wenn für die Klassifikation zusätzliche Texturmaße verwendet werden. Während am Orthophoto mit 20 cm Auflösung Ergebnisse um 85% erzielbar sind, liegen die Gesamtgenauigkeiten beim Orthophoto mit 500 cm Auflösung um 10 Prozentpunkte darunter. Bei der Bewertung macht es keinen Unterschied, ob die OOB Fehler aus dem Trainingsdatensatz betrachtet werden oder die Ergebnisse aus dem unabhängigen Validierungsdatensatz (Abb. 6).

Zwischen den drei getesteten Gruppen von Texturmaßen schneidet bis 200 cm die DSWT (*coif1*) immer am besten ab. Die schlechtesten Ergebnisse liefern die im Objekt berechneten Haralick Texturmaße. Die mittels Moving Window berechneten Haralick Texturmaße nehmen eine Mittelstellung ein. Ab 200 cm Bodenauflösung können die direkt im Objekt berechneten Haralick Texturmaße nicht mehr zu einer Steigerung der Klassifikationsgenauigkeit beitragen. Am Orthophoto mit 500 cm Auflösung bringen die mittels Moving Window berechneten Texturmaße nach Haralick die besten Ergebnisse. Sie sind die einzigen, die bei dieser Auflösung noch einen nennenswerten Beitrag zur Verbesserung des Klassifikationsergebnisses liefern können. Allerdings sind die bei dieser groben Auflösung erzielten Genauigkeiten wesentlich schlechter als beispielsweise jene mit den 20 cm Daten.

3.3 Variablenselektion

In Abb. 7 sind die mittleren Gesamtgenauigkeiten aus 100 Berechnungen bei Verwendung unterschiedlich vieler „wichtigsten Va-

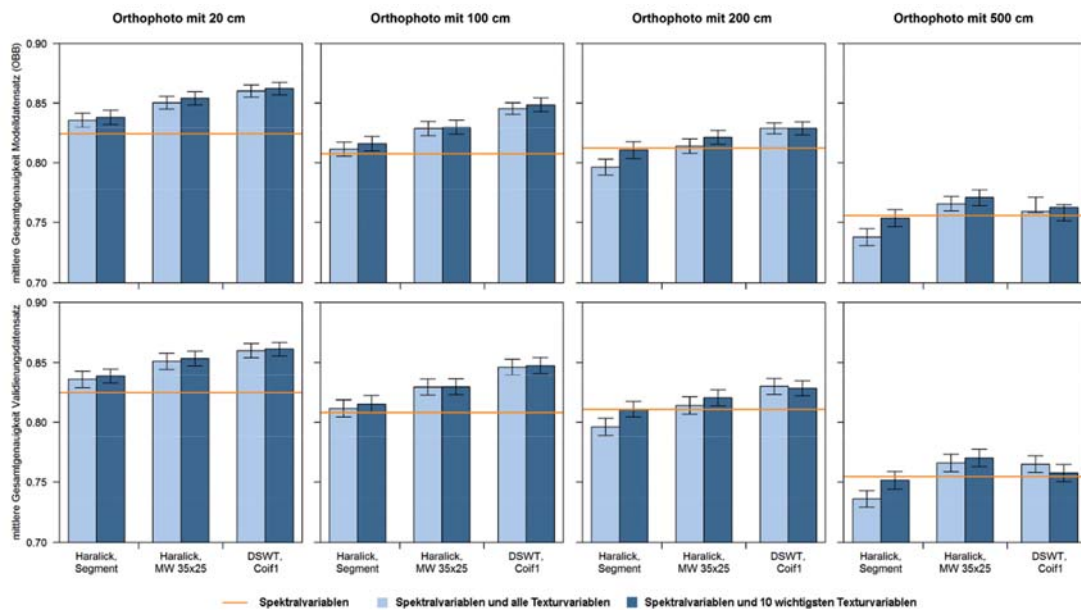


Abb. 6: Mittlere Gesamtgenauigkeit und Standardabweichung aus 100 Wiederholungen bei Verwendung aller bzw. der nach MDA 10 wichtigsten Texturvariablen für die unterschiedlichen Orthophotoauflösungen, obere Reihe: OOB-Genauigkeiten des Modellansatzes, untere Reihe: Gesamtgenauigkeit des unabhängigen Validierungsdatensatzes.

riablen“ nach MDA bzw. MDG des DSWT coif1 dargestellt. Vergleicht man diese Ergebnisse (Punktreihen) mit dem Ergebnis bei Verwendung aller coif1 Texturvariablen zeigt sich, dass durch gezielte Auswahl an Variablen höhere Klassifikationsgenauigkeiten erreicht werden können. Die Variablenselektion anhand der MDA-Wichtigkeit liefert bereits mit den 6 wichtigsten Variablen ein besseres Ergebnis. Die Ergebnisse auf Basis von

MDG liegen allgemein unter jenen auf Basis von MDA. Daher sind die 11 wichtigsten Variablen nach MDG notwendig, um die Klassifikationsgenauigkeit unter Verwendung aller Variablen zu erreichen. Aus diesem Grund sind in Abb. 6 neben den mit allen Variablen erzielten Ergebnissen (in hellblau) auch jene bei Verwendung der 10 wichtigsten Variablen nach MDA dargestellt (in dunkelblau).

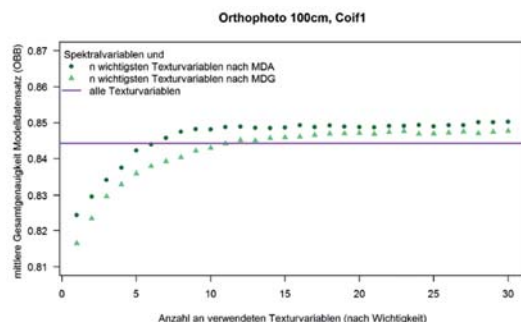


Abb. 7: Klassifikationsgenauigkeit (Mittelwerte aus 100 Wiederholungen) mit unterschiedlicher Anzahl (1-30) an nach MDA bzw. MDG wichtigsten Variablen im Vergleich zum Ergebnis mit allen Variablen (Orthophoto mit 100 cm Auflösung, Variablen: DSWT coif1).

3.4 Detaildarstellung des besten Klassifikationsergebnisses

Die Klassifikationsmatrizen basierend auf den Majority Votes aus den 100 Wiederholungen der Klassifikation mit und ohne den 10 wichtigsten Texturmaßen der DSWT coif1 zeigen, dass alle Klassen mit hohen Genauigkeiten klassifiziert wurden. Die geringsten Nutzer- und Produzentengenauigkeiten finden sich bei der Klasse Laubholz (LH), wobei die meisten Verwechslungen mit der zweiten Waldklasse Nadelholz (NH) zu beobachten sind. Die Latschenflächen wurden am zahlreichsten mit Nadelholz bzw. mit Offenland Gras verwechselt. Die Schattenobjekte wurden am häufigs-

Tab. 2: Klassifikationsmatrix aus dem Majority Vote der 100 Wiederholungen für 6 Klassen bei Verwendung der spektralen Information und den nach MDA 10 wichtigsten DSWT coif1 Variablen aus dem 100 cm Orthophoto (in Klammern die Werte bei Klassifikation nur mit spektraler Information). NH = Nadelholz, LH = Laubholz, OL = Offenland, NG = Nutzergenauigkeit, PG = Produktionsgenauigkeit.

		Referenz						Σ	NG
		NH	LH	Latsche	OL Gras	OL Fels	Schatten		
Klassifiziert als	NH	1001 (953)	134 (126)	36 (130)	40 (42)	5 (1)	67 (67)	1283 (1319)	0,780 (0,723)
	LH	54 (61)	333 (309)	2 (6)	18 (18)	0 (0)	11 (12)	418 (406)	0,797 (0,761)
	Latsche	38 (61)	1 (15)	682 (590)	44 (45)	1 (2)	7 (4)	773 (717)	0,882 (0,823)
	OL Gras	42 (59)	17 (24)	41 (33)	985 (985)	22 (20)	0 (0)	1107 (1121)	0,890 (0,879)
	OL Fels	0 (0)	0 (0)	4 (5)	22 (19)	510 (515)	3 (5)	539 (544)	0,946 (0,947)
	Schatten	31 (32)	5 (16)	7 (8)	0 (0)	6 (6)	380 (380)	429 (442)	0,886 (0,860)
	Σ	1166 (1166)	490 (490)	772 (772)	1109 (1109)	544 (544)	468 (468)	4549 (4549)	
PG	0,858 (0,817)	0,680 (0,631)	0,883 (0,764)	0,888 (0,888)	0,938 (0,947)	0,812 (0,812)		0,855 (0,820)	

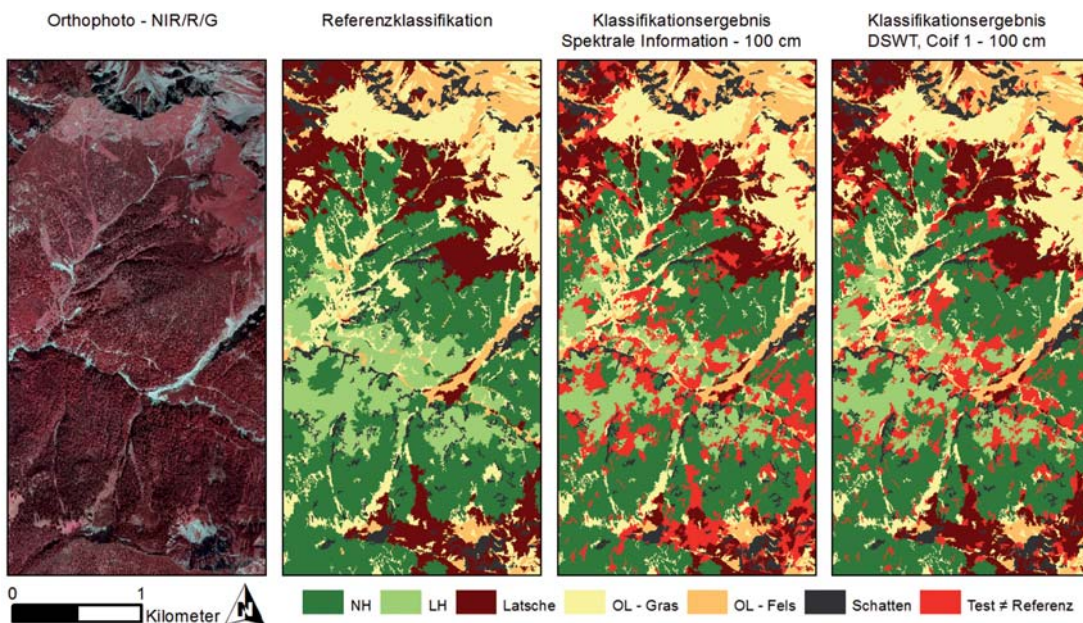


Abb. 8: Orthophoto, Referenzklassifikation und Klassifikationsergebnisse Spektraler Information bzw. Spektraler Information + DSWT coif1 (Mittelwerte aus 100 Wiederholungen) des Untersuchungsgebietes für 6 Klassen bei Verwendung des 100 cm Orthophotos.

ten fälschlich als Nadelholz klassifiziert. Im Vergleich zum Klassifikationsergebnis ohne Texturinformation ist eine Verbesserung der Klassifikationsgenauigkeit vor allem bei der Klasse Latsche und im geringeren Ausmaß bei den beiden Waldklassen (NH und LH) erkennbar.

In der Kartendarstellung in Abb. 8 sind diese beiden Klassifikationsergebnisse (mit/ohne coif1) der Referenzklassifikation gegenübergestellt. Abweichungen zwischen den Klassifikationsergebnissen und der Referenzklassifizierung sind in Rot dargestellt. Insbesondere im Norden und im südlichen Bereich des Bil-

Tab. 3: Klassifikationsmatrix aus dem Majority vote der 100 Wiederholungen für 3 Klassen (+ Schatten) bei Verwendung der spektralen Information und den nach MDA 10 wichtigsten DSWT coif1 Variablen aus dem 100 cm Orthophoto (in Klammern die Werte bei Klassifikation nur mit spektraler Information).

		Referenz					Σ	NG
		Wald	Latsche	Offenland	Schatten			
Klassifiziert als	Wald	1522 (1449)	38 (136)	63 (61)	78 (79)	1701 (1725)	0,895 (0,840)	
	Latsche	39 (76)	682 (590)	45 (47)	7 (4)	773 (717)	0,882 (0,823)	
	Offenland	59 (83)	45 (38)	1539 (1539)	3 (5)	1646 (1665)	0,935 (0,924)	
	Schatten	36 (48)	7 (8)	6 (6)	380 (380)	429 (442)	0,886 (0,860)	
	Σ	1656 (1656)	772 (772)	1653 (1653)	468 (468)	4549 (4549)		
PG		0,919 (0,875)	0,883 (0,764)	0,931 (0,931)	0,812 (0,812)		0,906 (0,870)	

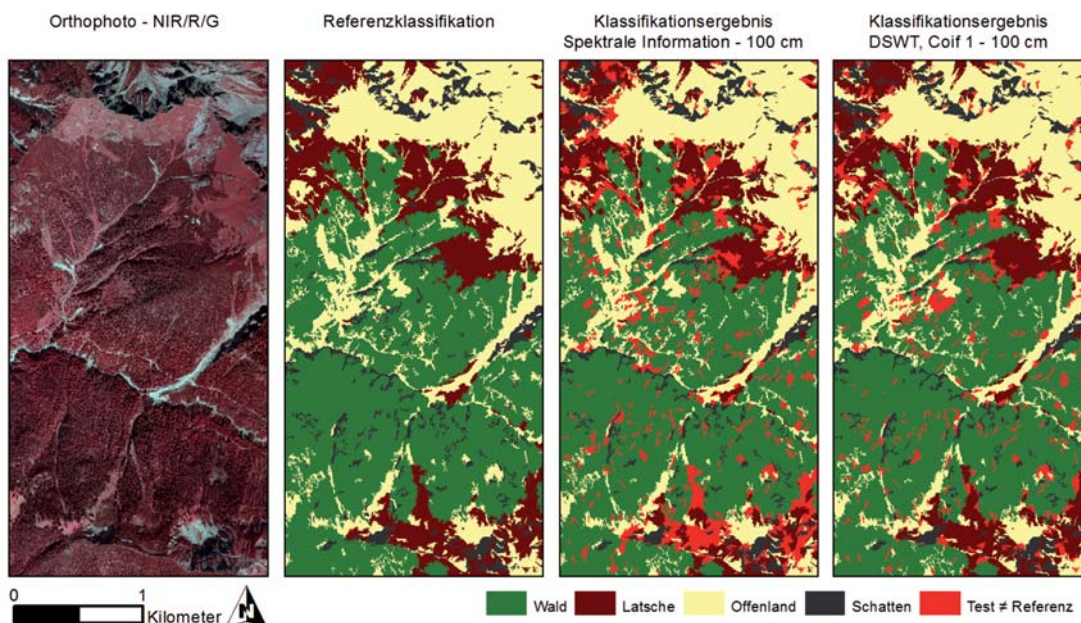


Abb. 9: Orthophoto, Referenzklassifikation und Klassifikationsergebnisse Spektraler Information bzw. Spektraler Information + DSWT coif1 (Mittelwerte aus 100 Wiederholungen) des Untersuchungsgebietes für 3 Klassen (plus Schatten) bei Verwendung des 100 cm Orthophotos.

des ist die deutliche Verbesserung der Klassifikation der Latschenflächen erkennbar, wenn die *coif1* Texturinformation zur Klassifikation mitverwendet wurde.

Die Untersuchung des Einflusses der Polygongröße auf das Klassifikationsergebnis zeigt, dass sich die Größenverteilung der falschklassifizierten Polygone nur geringfügig von den restlichen Polygonen unterscheidet. Lediglich bei den Klassen Nadelholz und Schatten konnte ein signifikanter Einfluss der Objektgröße festgestellt werden. Bei diesen beiden Klassen ist der Median der falsch klassifizierten Polygone deutlich geringer als im gesamten Datensatz. Das bedeutet, dass für diese beiden Klassen die größeren Polygone tendenziell besser klassifiziert wurden.

Die Aggregation der Ergebnisse auf die im ursprünglichen Projekt geforderten 3 Klassen (+ Schatten) verbessert die Gesamtklassifikationsgenauigkeit ohne Texturvariablen von 82,0% auf 87,0% und mit *coif1* Texturvariablen von 85,5% auf 90,6%. Mit Texturvariablen liegen alle Nutzer- und Produzentengenauigkeiten der drei Klassen Wald, Latsche und Offenland über 88% (Tab. 3).

Die Kartendarstellung in Abb. 9 zeigt den Vergleich der Klassifikationsergebnisse mit und ohne Texturvariablen (*coif1*) für die 3 ursprünglich geforderten Klassen und der zusätzlichen Schattenklasse.

Um die Übertragbarkeit der RF Modelle auf andere Gebiete zu untersuchen, wurde eine benachbarte Orthophotokachel (Abb. 1) ohne weitere Anpassungen klassifiziert. Es wurde eine mittlere Gesamtgenauigkeit von 83,2% bei 100 cm und 85,3% bei 20 cm Auflösung erzielt.

4 Diskussion

Durch die Einbeziehung von texturbeschreibenden Variablen bei der Klassifikation konnte in dieser Arbeit die Klassifikationsgenauigkeit signifikant gesteigert werden. Diese Erkenntnis deckt sich mit zahlreichen Studien zur Klassifikation der Landbedeckung (RYHERD & WOODCOCK 1996, WULDER et al. 1998, FRANKLIN et al. 2000, LALIBERTE & RANGO 2009, RODRIGUEZ-GALIANO et al. 2012a). Da-

bei ergaben alle drei getesteten Texturansätze zum Teil deutlich verbesserte Klassifikationsergebnisse. Die höchste Steigerung der Klassifikationsgenauigkeit erzielte in unserer Studie der Ansatz „*diskrete stationäre Wavelet Transformation*“ (im Mittel eine um 3,3 Prozentpunkte verbesserte Klassifikationsgenauigkeit am 100 cm Orthophoto), gefolgt vom Ansatz „*Texturmaße nach Haralick, Moving Window*“ (mittlere Steigerung: 2,2 Prozentpunkte). Die geringste Genauigkeitssteigerung brachte der Ansatz „*Texturparameter nach Haralick, Bezugsfläche Segment*“ (1,0 Prozentpunkte am 100 cm Orthophoto). Vor allem für die Unterscheidung von spektral sehr ähnlichen Klassen erwiesen sich die getesteten Texturmaße als vorteilhaft. Im konkreten Fall sind dies vor allem die Klassen Nadelholz und Latsche. Objekte dieser beiden Klassen würde auch ein Bildinterpret nur anhand der für Latschen eindeutigen Textur (Abb. 2) unterscheiden. Damit wird auch deutlich, wie wichtig es ist, bei der Arbeit mit hochauflösenden Sensoren auf objektbasierte Ansätze zurückzugreifen (JÄHNE 2005, KAYITAKIRE et al. 2006, CASTILLA & HAY 2008). Im Vergleich zu pixelbasierten Verfahren ermöglichen objektbasierte Ansätze auf eine einfache Art und Weise die Berechnung zusätzlicher, das Objekt beschreibender Information wie z. B. Textur.

Die positiven Erfahrungen in dieser Arbeit mit wavelet-basierten Texturmaßen bezüglich Klassifikationsgenauigkeit und Rechenzeitaufwand sind besonders herauszuheben. Sie bestätigen frühere Arbeiten von ZHU & YANG (1998), VAN COILLIE et al. (2011), SINGH et al. (2012). Insbesondere stellt die diskrete stationäre Wavelet Transformation unseres Erachtens eine interessante Alternative zu den gängigen Texturparametern nach Haralick dar. Das Potential der Frequenzanalyse durch Wavelet Transformation wird schon seit längerem in vielen Bereichen wie beispielsweise der Computer Vision für bildgebende Verfahren in der Medizin und Datenkompression in der Bildverarbeitung genutzt (GRAPS 1995, LOUIS et al. 1998). Eine verbreitete Anwendung in der Fernerkundung blieb aber bisweilen noch aus, wahrscheinlich auch, weil solche Verfahren noch nicht in gängigen Softwareprodukten implementiert sind.

LIVENS et al. (1997) zeigten, dass die einzelnen Wavelets nur geringe Unterschiede in den Klassifikationsergebnissen bewirken. Dieses Erkenntnis deckt sich mit den Ergebnissen dieser Arbeit. Die fünf aus unterschiedlichen Wavelet-Familien stammenden Wavelets zeigten ebenfalls nur geringe Unterschiede in der Klassifikationsgenauigkeit. Wesentlich wichtiger als die verwendete Wavelet-Familie, war die räumliche Auflösung der Daten. Generell konnten die besten Ergebnisse bei der höchsten Auflösung erzielt werden. Dies wurde auch bereits durch andere Arbeiten bestätigt (SCHLERF et al. 2003).

Neben den Klassifikationsergebnissen stellt der Aufwand der Datengewinnung einen wesentlichen Aspekt für die Bewertung der Anwendbarkeit dar. Bezüglich der erforderlichen Rechenzeit unterscheiden sich die drei getesteten Ansätze deutlich voneinander. Die Berechnung der Parameter „Haralick, Bezugsfläche Segment“ erfolgte am schnellsten. Aber auch die Berechnungen der DSWT-Parameter war deutlich weniger zeitintensiv als die Haralick Parameter mittels Moving Window. Die Rechenzeiten der drei getesteten Ansätze verhielten sich dabei ca. im Verhältnis 1:2:80 (Haralick Segment, DSWT, Haralick MW). Noch rechenzeitkritischer ist die Auflösung des analysierten Bildes. Hier steigt die Rechenzeit mit dem Quadrat des Auflösungs-faktors (RANDEN & HUSØY 1999, LILLESAND et al. 2004). Die Berechnung der Texturparameter auf dem 20 cm Bild benötigt daher rund 25-mal so lange wie jene auf dem 100 cm Bild.

Für großflächige Anwendungen ist daher eine Anpassung der Auflösung zu erwägen. Die dabei zu erwartende Verschlechterung der Klassifikationsgenauigkeit wurde auch bei dieser Untersuchung bestätigt. Die Verringerung der Auflösung der analysierten Orthophotos von 20 cm auf 100 cm und weiter auf 200 cm bewirkte eine Verschlechterung der mittleren Klassifikationsgenauigkeit von 82,3% auf 80,7% bei alleiniger Verwendung der spektralen Information. Zwischen 100 cm (80,7%) und 200 cm (81,2%) ist hingegen keine Verschlechterung feststellbar. Werden aber die Texturvariablen in der Klassifikation mitverwendet, sind die Vorteile der besseren Auflösung deutlicher. Die durch die Hinzunahme der DSWT coif1 Variablen gesteigerten

Klassifikationsgenauigkeiten lagen dabei bei 86,2% am 20 cm, bei 84,9% am 100 cm und bei 82,9% am 200 cm Bild.

Die Vergleiche der einzelnen Klassifikationen erfolgte in dieser Arbeit in erster Linie auf der im Random Forest (RF) Modell berechneten OOB Genauigkeit. Zusätzlich wurden die einzelnen Modelle aber auch auf einen unabhängigen Datensatz angewendet und dadurch validiert. Dabei konnte bestätigt werden, dass Random Forest durch das interne Bootstrapping bereits eine ausreichende unabhängige Testgröße liefert. Die Modellgenauigkeit, also die OOB-Klassifikationsgenauigkeit, lieferte dabei idente Genauigkeiten im Vergleich zum Validierungsdatensatz.

Die beiden gängigen Wichtigkeitsmaße Mean Decrease in Accuracy (MDA) und Mean Decrease in Gini (MDG) aus der RF Klassifikation eignen sich in der vorliegenden Arbeit gut für die Variablenselektion. So konnte mit beiden Maßen die Variablenzahl verringert und gleichzeitig die Klassifikationsgenauigkeit gesteigert werden. Auch wenn Random Forest prinzipiell mit einer großen Anzahl an Variablen umgehen kann, können zu viele Variable mit geringer Trennkraft das Klassifikationsergebnis verschlechtern (HASTIE et al. 2009). Eine Variablenselektion vermindert darüber hinaus den Zeitaufwand für die Berechnung der Klassifikationsmerkmale. Die Auswahl anhand MDA lieferte in unserer Studie das bessere Ergebnis und ermöglichte bereits mit den 10 wichtigsten Variablen bessere Ergebnisse als mit allen Texturvariablen. Somit sind für die Variablenselektion sowohl die MDA- als auch die MDG-Wichtigkeit geeignet, wobei in der Regel die MDA-Wichtigkeit zu bevorzugen ist, was auch die Ergebnisse anderer Studien bestätigt (CHAN & PAELINCKX 2008, NICODEMUS 2011, STUMPF & KERLE 2011, IMMITZER et al. 2012b).

5 Schlussfolgerungen und Ausblick

Durch das Hinzufügen von texturbeschreibenden Parametern in die objektbasierte Klassifikation von hochauflösenden Orthophotos ist eine deutliche Verbesserung der Klassifikationsgenauigkeit möglich. Mit den aus der

diskreten stationären Wavelet Transformation (DSWT) abgeleiteten Texturmaßen stehen interessante Alternativen zu den etablierten GLCM-Texturmaßen nach Haralick zur Verfügung. Die getesteten DSWT Texturmaße konnten dabei bezüglich Klassifikationsgenauigkeit aber auch in punkto geringerer Rechenzeit gegenüber den Haralick-Texturmaßen überzeugen. Diese Vorteile sind vor allem bei sehr hoher räumlicher Auflösung wesentlich. In dieser Studie konnten die besten Ergebnisse mit der höchsten Auflösung des Bildmaterials erzielt werden. Die Texturmaße zeigten bei den Orthophotos mit 20 cm bzw. 100 cm ihr Verbesserungspotenzial am deutlichsten.

Rechenzeitbetrachtungen sind insbesondere für großflächige Anwendungen von Bedeutung. So steigt der Rechenaufwand für die Berechnung der Texturvariablen im Quadrat zur Auflösung. Für großflächige Anwendungen sind daher Orthophotos mit 100 cm Auflösung trotz der etwas geringeren Klassifikationsgenauigkeiten gegenüber den höher auflösenden Daten oft zu bevorzugen.

Alle fünf getesteten Wavelets lieferten ähnlich gute Ergebnisse, wobei keine der fünf getesteten Wavelet-Familien besonders positiv hervorstach. Aus diesem Grund sind noch weitere Tests geplant, wobei auch noch nicht verwendete Wavelet-Familien einbezogen werden. Neben den Variablen und dem Klassifikationsverfahren wird vor allem den Objekten selbst ein wesentlicher Einfluss auf die Klassifikation zugeschrieben. Je homogener die Objekte bezüglich Größe und Form sind, desto aussagekräftiger werden auch die Texturmaße sein. Daher werden weitere Aktivitäten für die Optimierung der Segmentierung unternommen.

6 Danksagung

Die Arbeit wurde im Rahmen eines Projektes zwischen der Bayerischen Landesanstalt für Wald und Forstwirtschaft (LWF) und dem Institut für Vermessung, Fernerkundung und Landinformation (IVFL) der Universität für Bodenkultur Wien (BOKU) erstellt. Wir möchten uns bei Rudolf Seitz und Anna Kanold (LWF) für die bereitgestellten Daten

bedanken. Unser Dank geht auch an Tatjana Koukal, Reinfried Mansberger und Matteo Mattiuzzi für die gewährte technische und methodische Unterstützung.

Referenzen

- BAATZ, M. & SCHÄPE, A., 2000: Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation. – STROBL, J., BLASCHKE, T. & GRIESEBNER, G. (Hrsg.): *Angewandte Geographische Informationsverarbeitung XII*: 12–23, Wichmann-Verlag, Heidelberg.
- BARALDI, A. & PARMIGGIANI, F., 1995: An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. – *IEEE Transactions on Geoscience and Remote Sensing* **33** (2): 293–304.
- BINNER, S. & SEITZ, R., 2009: Mit Fernerkundung und GIS zuverlässig und kostensparend FFH-Gebiete erfassen und bewerten. – *LWF aktuell* **69**: 17–19.
- BLASCHKE, T., 2010: Object based image analysis for remote sensing. – *ISPRS Journal of Photogrammetry and Remote Sensing* **65**: 2–16.
- BLASCHKE, T. & STROBL, J., 2001: What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. – *GIS – Zeitschrift für Geoinformationssysteme* **6** (14): 12–17.
- BREIMAN, L., 2001: Random forests. – *Machine learning* **45** (1): 5–32.
- BREIMAN, L., 2002: *Manual on Setting Up, Using, and Understanding Random Forests V3.1.* – http://oz.berkeley.edu/users/breiman/Using_random_forests_V3.1.pdf (3.5.2012).
- BUSCH, A. & BOLES, W., 2002: *Texture Classification Using Multiple Wavelet Analysis.* – *DICTA2002: Digital Image Computing Techniques and Applications*, Melbourne, Australia.
- CASTILLA, G. & HAY, G., 2008: Image objects and geographic objects. – BLASCHKE, T., LANG, S. & HAY, G. (Hrsg.): *Object Based Image Analysis*: 92–110.
- CHAN, J.C.-W. & PAELINCKX, D., 2008: Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. – *Remote Sensing of Environment* **112** (6): 2999–3011.
- VAN COILLIE, F.M.B., VERBEKE, L.P.C. & DE WULF, R.R., 2011: Wavelet-based texture measures for semicontinuous stand density estimation from very high resolution optical imagery. – *Journal of Applied Remote Sensing* **5** (1).

- DRĂGUT, L., TIEDE, D. & LEVICK, S.R., 2010: ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. – *International Journal of Geographical Information Science* **24** (6): 859–871.
- FRANKLIN, S.E., HALL, R.J., MOSKAL, L.M., MAUDIE, A.J. & LAVIGNE, M.B., 2000: Incorporating texture into classification of forest species composition from airborne multispectral images. – *International Journal of Remote Sensing* **21** (1): 61–79.
- GRAPS, A., 1995: An introduction to wavelets. – *IEEE Computational Science Engineering* **2** (2): 50–61.
- HARALICK, R.M., SHANMUGAM, K. & DINSTEN, I., 1973: Textural Features for Image Classification. – *IEEE Transactions on Systems, Man and Cybernetics* **3** (6): 610–621.
- HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J., 2009: *The elements of statistical learning: Data mining, inference, and prediction*. – 2. Auflage, 768 S., Springer, New York.
- HAY, G.J., CASTILLA, G., WULDER, M.A. & RUIZ, J.R., 2005: An automated object-based approach for the multiscale image segmentation of forest scenes. – *International Journal of Applied Earth Observation and Geoinformation* **7** (4): 339–359.
- IMMITZER, M., ATZBERGER, C. & KOUKAL, T., 2012a: Eignung von WorldView-2 Satellitenbildern für die Baumartenklassifizierung unter besonderer Berücksichtigung der vier neuen Spektralkanäle. – *PGF* **2012** (5): 573–588.
- IMMITZER, M., ATZBERGER, C. & KOUKAL, T., 2012b: Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. – *Remote Sensing* **4** (9): 2661–2693.
- IMMITZER, M., KOUKAL, T., KANOLD, A., SEITZ, R., MANSBERGER, R. & ATZBERGER, C., 2012c: Abgrenzung der Natura 2000 Waldflächen – Klassifikation von Wald, Offenland und Latschenfeldern im bayerischen Hochgebirge unter Verwendung digitaler Luftbild- und Laserscannerdaten. – *LWF aktuell* **88**: 49–51.
- JÄHNE, B., 2005: *Digitale Bildverarbeitung*. – Springer-Verlag Berlin, Heidelberg, New York.
- KAYITAKIRE, F., HAMEL, C. & DEFOURNY, P., 2006: Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. – *Remote Sensing of Environment* **102** (3-4): 390–401.
- KROSCHEL, K., RIGOLL, G. & SCHULLER, B., 2011: *Statistische Informationstechnik: Signal - und Mustererkennung, Parameter- und Signalschätzung*. – Springer-Verlag Berlin, Heidelberg.
- LABERTE, A.S. & RANGO, A., 2009: Texture and Scale in Object-Based Analysis of Subdecimeter Resolution Unmanned Aerial Vehicle (UAV) Imagery. – *IEEE Transactions on Geoscience and Remote Sensing* **47** (3): 761–770.
- LIAW, A. & WIENER, M., 2002: Classification and Regression by randomForest. – *R News* **2** (3): 18–22.
- LILLESAND, T.M., KIEFER, R.W. & CHIPMAN, J.W., 2004: *Remote sensing and image interpretation*. – Wiley, New York.
- LIVENS, S., SCHEUNDERS, P., VAN DE WOUWER, G. & VAN DYCK, D., 1997: Wavelets for texture analysis, an overview. – *Sixth International Conference on Image Processing and Its Applications 1997*: 581–585.
- LOUIS, A.K., MAASS, P. & RIEDER, A., 1998: *Wavelets: Theorie und Anwendungen*. – Teubner, Stuttgart.
- MATLAB, 2011: Version 7.13.0 (R2011b). – The MathWorks Inc., Natick, Ma, USA.
- METZLER, V., PALM, C., LEHMANN, T. & AACH, T., 2002: Texture classification of gray-level images by multiscale cross co-occurrence matrices. – *IEEE Computer Society*: 549–552.
- NASON, G.P. & SILVERMAN, B.W., 1995: The stationary wavelet transform and some statistical applications. – *Lecture Notes in Statistics*: 281–281.
- NICODEMUS, K.K., 2011: Letter to the editor: on the stability and ranking of predictors from random forest variable importance measures. – *Briefings in bioinformatics* **12** (4): 369–373.
- POLIKAR, R., 1999: The Wavelet Tutorial. – <http://users.rowan.edu/~polikar/WAVELETS/WTtutorial.html> (27.9.2012).
- R DEVELOPMENT CORE TEAM, 2012: *R: A language and environment for statistical computing*. – R Foundation for Statistical Computing, Vienna, Austria.
- RANDEN, T. & HUSØY, J., 1999: Filtering for Texture Classification: A Comparative Study. – *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21** (4): 291–310.
- RENGERS, N. & PRINZ, T., 2009: JAVA-based Texture Analysis Employing Neighborhood Gray-Tone Difference Matrix (NGTDM) for Optimization of Land Use Classifications in High Resolution Remote Sensing Data. – *PGF* **2009** (5): 455–467.
- RODRIGUEZ-GALIANO, V.F., CHICA-OLMO, M., ABARCA-HERNANDEZ, F., ATKINSON, P.M. & JEGANATHAN, C., 2012a: Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. – *Remote Sensing of Environment* **121**: 93–107.
- RODRIGUEZ-GALIANO, V.F., GHIMIRE, B., ROGAN, J., CHICA-OLMO, M. & RIGOL-SANCHEZ, J.P., 2012b: An assessment of the effectiveness of a random forest classifier for land-cover classification. –

- ISPRS Journal of Photogrammetry and Remote Sensing **67**: 93–104.
- RYHERD, S. & WOODCOCK, C., 1996: Combining spectral and texture data in the segmentation of remotely sensed images. – *Photogrammetric Engineering and Remote Sensing* **62** (2): 181–194.
- 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.
- SHABAN, M.A. & DIKSHIT, O., 2001: Improvement of classification in urban areas by the use of textural features: The case study of Lucknow city, Uttar Pradesh. – *International Journal of Remote Sensing* **22** (4): 565–593.
- SINGH, A., DUTTA, R., STEIN, A. & BHAGAT, R.M., 2012: A wavelet-based approach for monitoring plantation crops (tea: *Camellia sinensis*) in North East India. – *International Journal of Remote Sensing* **33** (16): 4982–5008.
- SOH, L.-K. & TSATSIOULIS, C., 1999: Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. – *IEEE Transactions on Geoscience and Remote Sensing* **37** (2): 780–795.
- STUMPF, A. & KERLE, N., 2011: Object-oriented mapping of landslides using Random Forests. – *Remote Sensing of Environment* **115**: 2564–2577.
- VIEIRA, M.A., FORMAGGIO, A.R., RENNÓ, C.D., ATZBERGER, C., AGUIAR, D.A. & MELLO, M.P., 2012: Object Based Image Analysis and Data Mining applied to a remotely sensed Landsat time-series to map sugarcane over large areas. – *Remote Sensing of Environment* **123**: 553–562.
- 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.
- WASKE, B. & BRAUN, M., 2009: Classifier ensembles for land cover mapping using multitemporal SAR imagery. – *ISPRS Journal of Photogrammetry and Remote Sensing* **64** (5): 450–457.
- WULDER, M.A., LEDREW, E.F., FRANKLIN, S.E. & LAVIGNE, M.B., 1998: Aerial Image Texture Information in the Estimation of Northern Deciduous and Mixed Wood Forest Leaf Area Index (LAI). – *Remote Sensing of Environment* **64** (1): 64–76.
- ZHU, C. & YANG, X., 1998: Study of remote sensing image texture analysis and classification using wavelet. – *International Journal of Remote Sensing* **19** (16): 3197–3203.

Adresse der Autoren:

PHILIPP TOSCANI, MARKUS IMMITZER & Dr. CLEMENT ATZBERGER, Universität für Bodenkultur, Institut für Vermessung, Fernerkundung und Landinformation, Peter-Jordan-Straße 82, A-1190 Wien, Tel: +43-1-47654-5100, Fax: +43-1-47654-5142, e-mail: {philipp.toscani}{markus.immitzer}{clement.atzberger}@boku.ac.at

Manuskript eingereicht: November 2012

Angenommen: Januar 2013

Berichte von Veranstaltungen

ISPRS-Kommission VIII: Bericht zum ISPRS-Kongress 25. 8. – 1. 9. 2012 in Melbourne, Australien

Die Kommission VIII zu den Anwendungen und Richtlinien der Fernerkundung hatte sich in Peking 2008 unter der Leitung von HARUSHI SHIMODA (Japan) neu strukturiert und präsentierte in Melbourne die Ergebnisse der zehn Arbeitsgruppen.

Die Kommission VIII veranstaltete insgesamt 26 Vortragsitzungen mit jeweils vier bis fünf Beiträgen sowie vier Postersitzungen und veröffentlichte 107 Beiträge im Tagungsband (ISPRS Archives Vol. XXXIX-B8). Alle zehn Arbeitsgruppen waren mit mindestens einer Vortragsitzung vertreten – allerdings trugen die einzelnen Arbeitsgruppen in sehr unterschiedlichem Ausmaß zur Konferenz bei, wie ein Überblick über die einzelnen Arbeitsgruppen deutlich macht, gegliedert nach abnehmender Präsenz:

- WG VIII/6 *Agriculture, Ecosystems and Bio-Diversity*: 7 Vortragsitzungen (*Crop Assessment; Vegetation and Ecosystems (2); Vegetation Dynamics; Biophysical Parameter Retrieval; Agricultural Management; Wetland Estimation*), 1 Postersitzung, 28 Beiträge.
- WG VIII/7 *Forestry*: 4 Vortragsitzungen (*Biomass Estimation with Active Sensors; Forest Change Detection; Classification and Lidar; Biomass*), 1 Postersitzung zusammen mit WG VIII/10, 15 Beiträge.
- WG VIII/8 *Land*: 3 Vortragsitzungen (*Landscape Biophysical Characterization; Drought, Moisture and Stress Detection; Land Cover Dynamics*), keine Postersitzung, 15 Beiträge.
- WG VIII/1 *Disaster Management*: 3 Vortragsitzungen (*Earthquakes and Damage Assessment; Mass Movements; Projects and Systems*), 1 Postersitzung mit WG VIII/3, 13 Beiträge.
- WG VIII/4 *Water*: 2 Vortragsitzungen (*Coastal Regions and Wetlands; Hydrologi-*

cal Estimations, Sediment and Geomorphology), 1 Postersitzung, 12 Beiträge.

- WG VIII/3 *Atmosphere, Climate and Weather*: 2 Vortragsitzungen (*Coastal Regions and Wetlands, Hydrological Estimations, Sediment and Geomorphology*), 1 Postersitzung zusammen mit WG VIII/1, 7 Beiträge.
- WG VIII/2 *Health*: 2 Vortragsitzungen (*Environment, Health*), keine Postersitzung, 6 Beiträge.
- WG VIII/10 *Cryosphere*: 2 Vortragsitzungen (*Sea Ice; Glaciers and Ice Sheets*), 1 Postersitzung zusammen mit WG VIII/7, 5 Beiträge.
- WG VIII/9 *Oceans*: 2 Vortragsitzungen (*Ocean Temperature, Salinity and Carbon; Ecology*), keine Postersitzung, 4 Beiträge.
- WG VIII/5 *Energy and Solid Earth*: 1 Vortragsitzung, keine Postersitzung, 2 Beiträge.

Im Vergleich zum Kongress in Peking 2008 fiel die zahlenmäßige Bilanz für Kommission VIII 2012 also erheblich schlechter aus. Zwar fanden in Melbourne drei Vortragsitzungen mehr, aber acht Postersitzungen weniger statt, so dass sich die Zahl der Beiträge im Tagungsband im Vergleich zu Peking (251) mehr als halbierte. Offensichtlich gab es für Melbourne eine deutlich geringere Anzahl von Beitragsanmeldungen, mit denen zunächst die Vortragsitzungen gefüllt worden waren.

Über die Gründe dieses Rückganges kann nur spekuliert werden. Abgesehen von den hohen Reise- und Teilnahmekosten, die allgemein viele potentielle Teilnehmer/innen abschreckten, ist es den neuen Arbeitsgruppen offenbar nicht in ausreichender Weise gelungen, die früheren Mitglieder der Kommission VIII für die seit 2008 geänderten thematischen Schwerpunkte zu begeistern bzw. neue Mitglieder hinzu zu gewinnen. Möglicherweise wurde auch die Abgrenzung zur Kommission VII (*Thematic Processing, Modelling and Analysis of Remotely Sensed Data*), in der viele Anwendungen von Kommission VIII mit stärkerem methodischem Fokus bearbeitet werden, nicht deutlich genug herausgestellt.

HARUSHI SHIMODA zog dennoch ein positives Resümee seiner Amtszeit. Er betonte die immer größer werdende Rolle von Satellitendaten beim Katastrophenschutz und Krisenmanagement, für die Wettervorhersage sowie die Atmosphären- und Klimaforschung. Dies sei auch auf die Arbeit und den Einfluss von Kommission VIII zurückzuführen. Immer häufiger würden heutzutage zum Beispiel räumlich sehr hochauflösende Satellitendaten zur Überwachung von Flüchtlingslagern oder in Feuerwarnsystemen genutzt. Viele Wetterdienste verwendeten operationell Daten meteorologischer Satelliten für die numerische Wettervorhersage. Satelliten wie GOSAT ermöglichten die Bestimmung der Bruttoprimärproduktion sowie des Kohlendioxid-Gehaltes in der Atmosphäre. Mit Hilfe von Satellitendaten sei außerdem die Kartierung der Gletscher weltweit gelungen und aus dem Vergleich von historischen und aktuellen Aufnahmen ein signifikanter Massenverlust der Eisschilde Grönlands und der Antarktis sowie ein dramatisches Schmelzen des arktischen Meereises festgestellt worden. Satellitendaten ermöglichten heute detailliertere Aussagen über den Einfluss des Klimawandels auf die verschiedenen Ökosysteme der Erde. Der Einsatz von Lidar habe zudem die Oberflächenmodelle von Waldgebieten sowie die Bestimmung der Biomasse von Wäldern deutlich verbessert.

SHIMODA hob besonders die gute Kooperation mit der Gesundheitsforschung hervor; die Leiterin der WG VIII/2 *Health*, Amelie Budge, erhielt wie in Peking die Ehrenurkunde des Präsidenten.

Zwei Bücher entstanden im Rahmen der ISPRS Book Series, eines mit dem Titel *Recent Advances in Remote Sensing and Geoinformation Processing for Land Degradation Assessment* (Volume 8), ein zweites zu *Environmental Tracking for Public Health Surveillance* (Volume 11), außerdem steuerte die Kommission VIII ein Kapitel über *Remote Sensing Applications* dem *Handbook of Satellite Applications* (Springer) bei. Kommission VIII führte zwei internationale Workshops durch und veröffentlichte die Beiträge in den Archiven der ISPRS (Vol. XXXVIII-8/W3 und Vol. XXXVIII-8/W20), ferner wurden weitere Workshops und Tagungen mit anderen Kommissionen veranstaltet.

Der Kongress verabschiedete als Leitlinien für Kommission VIII im Zeitraum 2012 – 2016 Resolutionen zu den folgenden drei Themen:

- 1) Global change studies;
- 2) Applications for the benefit of society;
- 3) Collaboration with global and applied sciences programmes and user communities.

Die Leitung der Kommission VIII im Zeitraum 2012 – 2016 übernahm VINAY K. DADHWAL vom National Remote Sensing Centre (NRSC), Indien.

IRMGARD NIEMEYER, Jülich

Racurs 12th International Scientific and Technical Conference, 24.–27. September 2012, Portimão, Portugal

Since the year 2001 Racurs organises international photogrammetric conferences. From September 24 to 26, 2012, the conference subtitled “From imagery to map: digital photogrammetric technologies” was held in the Algarve in Portugal at the Postano Hotel in Portimão. RACURS is a Russian Photogrammetric software company founded in 1993. It produced its first Photomod System in 1994. In 2005 the fifth version Photomod 5 placed it in one of the leading photogrammetric processing systems. It can process analog and digital aerial imagery, satellite imagery and images from UAVs. It is available in Russian, English and Spanish.

During this two-day conference numerous papers dealing with technical and economic issues of data acquisition and mapping for geospatial applications have been presented. The Conference 2012 demonstrated again, that photogrammetry and remote sensing has many important applications especially in the Russian Federation, where the use of satellite images is not only possible, but necessary. The Racurs conferences have generally become an important interface for Western and Central Europeans for communication with colleagues in the Russian Federation. Where else is it possible to listen to a conference in Russian with simultaneous translation into English?

GOTTFRIED KONECNY, Hannover

Hochschulnachrichten

Technische Universität Wien, Dissertation MARCELA DOUBKOVÁ

M.A. MARCELA DOUBKOVÁ promovierte am 20.11.2012 am Department für Geodäsie und Geoinformation der Technischen Universität Wien mit der Arbeit „*Error characterization methods for surface soil moisture products from remote sensing*“ zur Dr. techn.

1. Gutachter: Prof. Dr. WOLFGANG WAGNER, TU Wien, Department für Geodäsie und Geoinformation;
2. Gutachter: Prof. Dr. GÜNTER BLÖSCHL, TU Wien, Institut für Wasserbau und Ingenieurhydrologie;
3. Gutachter: Prof. Dr. KATARZYNA DĄBROWSKA-ZIELIŃSKA, Universität Warschau, Institute of Geodesy and Cartography.

Kurzfassung:

Zur Unterstützung des operationellen Einsatzes von Erdbeobachtungssystemen wie Synthetic Aperture Radar (SAR) entwickelt die Europäische Weltraumbehörde ESA den Radar-Satelliten Sentinel-1, der im C-Band arbeitet. Ähnlich seinen SAR-Vorgängern auf den Plattformen ERS, ENVISAT, oder RADARSAT wird der Sensor Sentinel-1 bei einer mittleren räumlichen Auflösung im Bereich von 5 bis 40 m arbeiten, allerdings mit einer vielfach erhöhten Wiederholrate, die über Europa etwa im Bereich von 2 Tagen liegen wird. Aufgrund dieser hohen zeitlichen Auflösung sowie dem operationellen Design wird der Satellit einen großen Beitrag zur Überwachung von dynamischen Prozessen in Hydrologie und Phänologie leisten. Der Nutzen von C-Band SAR-Überwachungssystemen in der Hydrologie wurde in der Vergangenheit bereits im Rahmen des Projects SHARE (Soil Moisture for Hydrometeorologic Applications) gezeigt, bei dem Daten des Instruments ASAR (Advanced Synthetic Aperture Radar) im Global Mode (GM) verwendet wurden. Um das volle Potential von SAR-Produkten im Bereich Bodenfeuchtigkeit auszuschöpfen, ist ein mitgeliefertes Fehler-

maß unerlässlich. Das Verständnis dieses Fehlermaßes ist unentbehrlich für die Anwendung von Bodenfeuchteprodukten in Modellen, die die Extraktion oder Erstellung von Produkten, sowie den Vergleich mit anderen Bodenfeuchteprodukten unterstützen.

Die vorliegende Arbeit umfasst mehrere Ziele. Erstens werden Grundlagen sowie ein aktueller Stand der Technik im Bereich der Fehlermaße von Bodenfeuchteprodukten dargestellt, wie etwa dem quadratischen mittleren Fehler, Korrelationskoeffizienten oder erweiterten Methoden wie Fehlerfortpflanzung und triple collocation. Des Weiteren wird eine Übersicht der Anwendungsbereiche von Bodenfeuchteprodukten präsentiert und Evaluierungsmethoden je nach Bereich und Qualitätsanforderung vorgeschlagen. Die Evaluierung von ASAR GM Bodenfeuchteprodukten mit eben diesen Methoden stellt ein zweites Ziel dieser Arbeit dar. Um dies zu erreichen, wurden Daten des australischen hydrologischen Modells Water Assessment System (AWRA-L), Feldmessungen des australischen Netzwerkes OzNET, sowie weitere geringer aufgelöste Fernerkundungsdaten verwendet, um die Qualität der Bodenfeuchteprodukte umfassend zu beschreiben. Die dritte Zielsetzung dieser Arbeit ist das Bereitstellen von Richtlinien für eine Evaluierungsmethode, die auf beliebige Bodenfeuchteprodukte angewendet werden kann. Zu diesem Zweck wurden die ASAR GM Ergebnisse vor einem breiteren Hintergrund analysiert.

Die Arbeit liefert Antworten und Ergebnisse, die auch auf weitere satellitenbasierte Bodenfeuchteprodukte angewendet werden können. Besonders die Übertragung auf den geplanten Sentinel-1 Sensor ist von besonderem Interesse, da dieser Sensor zwar ähnliche technische Eigenschaften, aber ein verbessertes Fehlermaß im Vergleich zu ASAR GM besitzt. Die operationell verfügbaren Bodenfeuchteprodukte von Sentinel-1 werden wesentlich zur Modellierung und Beobachtung von Land-Atmosphäre-Interaktionen, Erntertrag sowie Anwendungen im Bereich der Wasserbilanz beitragen.

Technische Universität Dresden, Habilitation Dr. TOBIAS BOLCH

Herr Dr. TOBIAS BOLCH habilitierte am 7.5.2012 am Institut für Kartographie, Fakultät Forst-, Geo- und Hydrowissenschaften der Technischen Universität Dresden mit der Arbeit „*Cryosphere Variations in Different Climatic Regimes and its Impacts Using Geomatics*“ zum Dr. rer. nat. habil.

1. Gutachter: Prof. Dr. MANFRED BUCHROITHNER, TU Dresden, Institut für Kartographie;
2. Gutachter: Prof. Dr. WILFRIED HAEBERLI, Universität Zürich, Geographisches Institut;
3. Gutachter: Prof. Dr. BRIAN MENOUNOS, University of Northern British Columbia, Canada

Kurzfassung

Glaciers are typical elements of the mountain cryosphere and are key indicators for assessing climate change in remote mountain areas where climate stations are rare or non-existent. Glacier run-off can be an essential source of freshwater especially during summer and early autumn when the water demand is usually highest. Water release from permafrost bodies can also play an important role especially in arid areas. Since the end of the Little Ice Age, a nearly global recession of the glaciers with an accelerating trend in recent decades is noticeable. Glacial lakes can develop and expand concomitant with glacier recession. Outbursts of these lakes can represent a serious threat to the infrastructure and society downvalley. Hence, assessing the cryospheric variations is of uttermost importance.

However, glaciers are usually situated in remote and rugged mountain regions which are difficult to access. Remote sensing, GIS analysis and modelling are therefore ideal means to study glacier fluctuations and their impacts. A prerequisite for these investigations is a precise digital terrain model (DTM) and suitable remote sensing imagery. Clean ice is relatively easy to map using multi-spectral imagery. Debris-cover, however, which is common for many glaciers of the world, hampers the automated mapping. Own analysis revealed that a combination of morphometric and thermal information leads to promising results for larger debris-covered glaciers. Slope was identified

as the key parameter for this task. The uncertainties of the mapping approaches are, however, too large for change assessments and manual adjustment is required. A glacier inventory can be generated within less than one year for large areas such as western Canada with more than 15,000 glaciers based on remote sensing if suitable imagery and previous glacier outlines are available and the area of debris-covered glaciers is limited.

Early declassified imagery from the 1960s and 1970s such as Corona and Hexagon are valuable sources for extending the remote-sensing based glacier analysis back in time and to evaluate previous inventories based on topographic maps from similar years. Glacier area and length changes represent indirect signals of climate variability only while glacier mass balance shows the most direct signal. Multi-temporal DTM comparisons allow the estimation of the glacier mass balance. Own investigations showed that stereo Corona is also a valuable source for the generation of precise DTMs for this task despite of image distortion and unstable camera parameters. Multi-temporal DTMs from different sources such as ASTER, aerial images or Cartosat-1 data allow the generation of a mass balance time series. However, tilts and inaccuracies in absolute elevation of the generated DTMs or the use of different software packages require careful relative adjustments in order to minimize the uncertainties.

Almost all glaciers showed a significant and continuous recession in all investigated regions in western Canada and high Asia in the last decades despite of different climatic controls. The identified recessions were, however, less than previously published based on topographic maps. Also, glaciers at Mt. Everest clearly lost mass despite thick debris cover. However, the mass loss is within the range of the global mean. A case study about rockglaciers in the Tien Shan based on remote sensing and terrain analysis provided some insights into their occurrence and characteristics. The spatial distribution of rockglaciers is to a certain degree related to the distribution of glaciers and the availability of debris. The permafrost bodies contain also a significant amount of ice which can be of importance for the run-off. Geomatics was also proven to be

an effective tool to assess the developments and to provide a first order classification of the potential danger of the glacial lakes. The current and future availability of high resolution imagery and DTMs will allow similar investi-

gations with higher accuracy which will lead to an improved understanding of the reaction of the cryosphere to the changing climate and its impacts.

Deutsche Geodätische Kommission (DGK)

60 Jahre DGK

Die DGK bei der Bayerischen Akademie der Wissenschaften beging im Rahmen der Jahrestagung 2012 die 60. Wiederkehr ihrer Gründung im Jahre 1952. Die Tagung fand gemeinsam mit den Schwesterkommissionen aus Österreich und der Schweiz in Diessenhofen bei Schaffhausen am Rhein statt. Anwesend waren auch Vertreter der Niederländischen Geodätischen Kommission und Kollegen aus weiteren Staaten.

Im Rahmen einer festlichen Abendveranstaltung skizzierte der Ständige Sekretär der DGK, THOMAS WUNDERLICH, TU München, lebhaft die Entstehungsgeschichte vor allem im Kontext der internationalen Geodäsie. Die Jahrestagung selber war von Diskussionen zu aktuellen Themen in Forschung und insbesondere in der universitären Lehre geprägt. Dabei stellten die Referenten speziell in der Lehre den europäischen Aspekt heraus, der für die Zukunft der Geodäsie immer wichtiger wird.

Die DGK vertritt die geodätische Forschung und Lehre in Deutschland und hat 45 Ordentliche Mitglieder, die in der Regel Universitätsprofessoren für alle Teilgebiete der Geodäsie sind. Ausländische Wissenschaftler wirken an den Arbeiten als Korrespondierende Mitglieder mit, außeruniversitäre geodätische Organisationen als Ständige Gäste. Zu den Aufgaben der DGK gehört vorrangig die wissenschaftliche Forschung auf allen Gebieten der Geodäsie. Die DGK gliedert sich in die vier forschungsorientierten Sektionen *Erdmessung*, *Ingenieurgeodäsie*, *Geoinformatik und Land- und Immobilienmanagement*, sowie in eine Sektion für *Lehre*. Seit 2010 wird die DGK von CHRISTIAN HEIPKE, Leibniz Universität Hannover, als Vorsitzendem und THOMAS WUNDERLICH, Technische Universität München, als Ständigem Sekretär geleitet.

Ausführliche Informationen über die DGK finden sich unter <http://dgk.badw.de/>.

CHRISTIAN HEIPKE, Hannover



Die Führungsriege der DGK: THOMAS WUNDERLICH (Ständiger Sekretär), HELMUT HORNIK (Geschäftsführer), CHRISTIAN HEIPKE (Vorsitzender) während der 60. Jahrestagung der DGK.

Persönliches

Nachruf auf Enrico Clerici (1939–2012)



Am 26. Oktober 2012 verstarb Dr.-Ing. ENRICO CLERICI nach langer schwerer Krankheit im 74. Lebensjahr an seinem neuen Wohnort Rivergaro. Er wurde in engstem Familienkreis im natürlich gelegenen Grab der Familie in den Hügeln der nördlichen Apenninen bestattet und war damit nach vielen Lebensstationen zu seinen Wurzeln zurückgekehrt.

CLERICI kam auf einem für deutsche Verhältnisse sehr ungewöhnlichen Weg zur Photogrammetrie. Geboren am 20. Januar 1939 in Mailand, ging er gleich nach der Schule im Alter von 17 Jahren nach Afrika und half bei Vermessungsarbeiten einer Baufirma in Uganda. Von 1959 bis 1964 fuhr er in der Handelsmarine zur See und betrieb gleichzeitig ein Fernstudium in Land Surveying, welches er mit einem Diplom der ICS Scranton, PA/USA abschloss. Mit dieser Ausbildung arbeitete er für eine niederländische Ingenieur-Firma drei Jahre lang in Nigeria. 1967 begann er seine photogrammetrische Ausbildung am ITC, die er 1972 in Enschede mit dem M.Sc. abschloss. In Holland hatte er auch seine Frau kennen gelernt und die niederländische Staatsbürgerschaft angenommen. Parallel zum Masterstudium arbeitete er zunächst an Forschungs- und Entwicklungsarbeiten zu Fernerkundungssensoren (Radar, IR Line Scanner

usw.). Vom ITC ging CLERICI zum Rijkswaterstaat. In dieser Zeit vervollständigte er auch seine Erfahrungen mit Sonar-Verfahren, so dass er darüber 1976 an der TU Hannover unter dem Titel „Über die Anwendbarkeit von Sidescan Sonar zur Erstellung von topographischen Karten des Meeresbodens“ promovieren konnte.

1975 wurde CLERICI als Senior Lecturer an das Department of Surveying der University of Queensland nach Brisbane berufen, wo die Familie – inzwischen mit Tochter und Sohn – für die kommenden 11 Jahre ihren Lebensmittelpunkt hatte, allerdings mit einer Unterbrechung von 1981 bis 1983, als er bei CARL ZEISS in Oberkochen als Projektleiter in der Software-Entwicklung für analytische Plotter arbeitete.

Diesem passionierten Weltenbummler war es aber nicht zu verdenken, dass er bald von der rauen Schwäbischen Alb zurück in die Subtropen Asiens und Australiens strebte. Er blieb aber ZEISS als Consultant für Australien und Südostasien (SEA) verbunden, ab Ende der 1980er Jahre mit Sitz in Singapur. Selbst während der ihn ab 2010 zunehmend behindernden Krankheit blieb er bis zuletzt in das laufende Geschäft eingebunden. CLERICIS Stärke als Selfmademan war die Verbindung seines anhaltenden theoretischen Interesses an Mathematik und Physik mit den praktischen Erfordernissen und Anwendungen, was sich in 15 Originalveröffentlichungen niederschlug.

Die Verbindung von theoretischer Neugier und Freude am praktischen Detail prägte auch seine Hobbys, wie das Foto zeigt. Mit seiner *Vagabond 47* segelte er fast jedes Jahr durch die Straße von Malakka bis nach Phuket und einmal sogar bis zu den Andamanen. Und auch dabei konnte er es nicht lassen, die GPS-Ortung mit Hilfe von Astro-Navigation zu überprüfen.

Wir trauern mit seiner Familie über diesen zu frühen Verlust und werden uns immer angenehm an ihn erinnern.

DIERK HOBBIE, Königsbronn

Veranstaltungskalender

2013

21.–23. April: **Joint Urban Remote Sensing Event (JURSE 2013)** in São Paulo, Brasilien. inpe.br/jurse2013/

30. April – 2. Mai: **8th International Symposium on Mobile Mapping Technology 2013** in Tainan, Taiwan. conf.ncku.edu.tw/mmt2013/

6.–7. Mai: **19. Workshop Computer-Bildanalyse in der Landwirtschaft / 2. Workshop Unbemannte autonom fliegende Systeme (UAS) in der Landwirtschaft**, Beuth Hochschule für Technik in Berlin. www2.atb-potsdam.de/workshop-cba-uas/

6.–8. Mai: **Symposium Königslutter 2013** der DGfK in Königslutter am Elm. angewandte-kartographie.de/

13.–16. Mai: **SPIE Optical Metrology in München – Videometrics, Range Imaging and Applications XII**. www.spie.org/x6506.xml

21.–23. Mai: **IAPR Conference on Machine Vision Applications (MVA 2013)** in Kyoto, Japan. mva-org.jp/mva2013/

21.–24. Mai: **ISPRS Hannover-Workshop „High-Resolution Earth Imaging for Geospatial Information“**. www.ipi.uni-hannover.de/ipi-workshop.html

3.–6. Juni: **33rd EARSeL Symposium in Matera**, Italien. earsel.org/symposia/2013-symposium-Matera/

5.–7. Juni: **International Conference on Earth Observation for Global Changes (EOGC'2013)** in Toronto, Canada. eogc2013.blog.ryerson.ca/

23.–28. Juni: **Computer Vision and Pattern Recognition (CVPR 2013)** in Portland, Oregon, USA. pamitc.org/cvpr13/

2.–5. Juli: **GI_Forum 2013** in Salzburg, Österreich. gi-forum.org

14.–20. Juli: **International Computer Vision Summer School in Punta Sampieri**, Italien. svg.dmi.unict.it/icvss2013/

21.–26. Juli: **IGARSS 2013** in Melbourne, Australien. igarss2013.org/

25.–30. August: **26th International Cartographic Conference (ICC)** in Dresden. icc2013.org/

2.–6. September: **XXIVth CIPA Heritage Documentation Symposium** in Strasbourg, Frankreich. cipa.icomos.org

4.–6. September: **UAV-g 2013** in Rostock. uav-g.org

9.–13. September: **54. Photogrammetrische Woche** in Stuttgart. ifp.uni-stuttgart.de/phowo

9.–13. September: **British Machine Vision Conference (BMVC 2013)** in Bristol, England. bmvc2013.bristol.ac.uk/

15.–18. September: **International Conference on Image Processing (ICIP 2013)** in Melbourne, Australien. www.ieeeicip.org/

26.–29. November: **3D GeoInfo 2013** in Istanbul, Türkei. 3dgeoinfo.com

8.–15. Dezember: **ICCV 2013**, International Conference on Computer Vision, Sydney, Australien. iccv2013.org

Weitere Konferenzen und Workshops finden sich beispielsweise unter:
isprs.org/calendar/
 und
iris.usc.edu/Information/Iris-Conferences.html

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 Luftbilddatenbewertung 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
 Bundesministerium 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 Meereschutz, 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 Photogrammetrie
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

