Photogrammetrie Fernerkundung Geoinformation

Journal for Photogrammetry, Remote Sensing and Geoinformation Science

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

Jahrgang 2012, Heft 4

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

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

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

International Society for Photogrammetry and Remote Sensing

Deutsche Arbeitsgemeinschaft für Mustererkennung e.V.

GeoUnion Alfred-Wegener-Stiftung

Herausgeber:

Präsident: Prof. Dr. Thomas Kolbe, Technische Universität Berlin, Institut für Geodäsie und Geoinformationstechnik, Straße des 17. Juni 135, 10623 Berlin, Germany, Tel.: +49-30-314-23274
Geschäftsstelle: Dr. Klaus-Ulrich Komp, c/o EFTAS Fernerkundung Technologietransfer GmbH, Oststraße 2–18, 48145 Münster, Germany, e-mail: klaus.komp@eftas.com
Published by: E. Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller), Johannesstraße 3A, 70176 Stuttgart, Germany, Tel.: +49-711 351456-0, Fax: +49-711 351456-99, e-mail: mail@schweizerbart.de
Internet: http://www.schweizerbart.de
∞ Gedruckt auf alterungsbeständigem Papier nach ISO 9706-1994

ISSN 1432-8364
Science Citation Index Expanded (also known as SciSearch®) Journal Citation Reports/Science Edition
Hauptschriftleiter: Prof. Dr.-Ing. Wolfgang Kresse, Hochschule Neubrandenburg, Fachbereich Landschaftsarchitektur, Geoinformatik, Geodäsie und Bauingenieurwesen, Brodaer Straße 2, 17033 Neubrandenburg, Germany, e-mail: kresse@hs-nb.de
Schriftleiter: Prof. Dr.-Ing. Stefan Hinz, Karlsruher Institut für Technologie – KIT, Institut für Photogrammetrie und Fernerkundung, Englerstraße 7, 76131 Karlsruhe, Germany, e-mail: stefan.hinz@ipf.uni-karlsruhe.de, Prof. Dr. rer. nat. Carsten Jürgens, Ruhr-Universität Bochum, Geographisches Institut, Gebäude NA7/133, 44780 Bochum, Germany, e-mail: carsten.juergens@rub.de, Prof. Dr. rer.nat. Lars Bernard, Technische Universität Dresden, Fachrichtung Geowissenschaften, Helmholtzstraße 10, 01062 Dresden, Germany, e-mail: lars.bernard@tu-dresden.de, Privatdozent Dr. techn. Franz Rottensteiner, Leibniz Universität Hannover, Institut für Photogrammetrie und GeoInformation, Nienburger Straße 1, 30167 Hannover, Germany, e-mail: rottensteiner@ipi.uni-hannover.de und Dr.-Ing. Eckhardt Seyfert, Landesvermessung und Geobasisinformation Brandenburg, Heinrich-Mann-Allee 103, 14473 Potsdam, Germany, e-mail: eckhardt.seyfert@geobasis-bb.de
Erscheinungsweise: 6 Hefte pro Jahrgang.
Bezugspreis im Abonnement: € 219,– pro Jahrgang. Mitglieder der DGPF erhalten die Zeitschrift kostenlos. Der Online-Zugang ist im regulären Subskriptionspreis enthalten.
Anzeigenverwaltung: E. Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller), Johannesstraße 3A, 70176 Stuttgart, Germany, Tel.: +49-711 351456-0; Fax: +49-711 351456-99.
e-mail: mail@schweizerbart.de, Internet: http://www.schweizerbart.de
Bernhard Harzer Verlag GmbH, Westmarkstraße 59/59a, 76227 Karlsruhe, Germany, Tel.: +49-721 944020, Fax: +49-721 9440230, e-mail: Info@harzer.de, Internet: www.harzer.de
Printed in Germany by Tutte Druckerei GmbH, 94121 Salzweg bei Passau, Germany.

PFG – Jahrgang 2012, Heft 4
Inhaltsverzeichnis

Editorial

Originalbeiträge

Mitteilungen

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

PFG 2012 / 4, 0315–0316
Stuttgart, August 2012

High Resolution Earth Imaging for Geospatial Information

The number and quality of digital images of the Earth taken from air and space has significantly increased over the last years. We can now observe our globe at a daily rate, and in principle the turn-around time of the data allows for real-time monitoring. As a result Earth images play a more and more important role in a growing number of applications. Examples comprise global change, disaster management, agriculture, forestry, insurance, security, traffic monitoring, and – of course – topographic mapping. All of these applications rely on fast and complete data exploitation with high geometric accuracy.

While digital aerial and space sensors, multiple-echo laser scanners and high resolution SAR sensors were introduced a number of years ago and are in operational use today, current trends in Earth imaging comprise an ever increasing ground resolution from space, full waveform laser scanners, unconventional platforms like unmanned aircraft systems (UAS) and geosensor networks. Developments such as those initiated by the Group of Earth Observation (GEO) and GMES (Global Monitoring of Environment and Security) will lead to even many more datasets. Automatic processing thus becomes a sheer necessity because interactive screening of the acquired images, leave alone detailed exploitation, becomes impossible due to the huge amount of bits and bytes.

These developments have formed the background for the latest edition of the ISPRS Hannover Workshop *High Resolution Earth Imaging for Geospatial Information* which was held at the Institute of Photogrammetry and Geo-Information (IPI), Leibniz Universität Hannover, between June 14 and 17, 2011 (see also the workshop report by Markus Gerke in PFG issue 5/2011). This biennial workshop, which was first held nearly two decades ago, has become a constant in the calendar of photogrammetry and remote sensing and has continually attracted around 100 participants from all around the world. It is worth mentioning that after 2005 this is already the second time that PFG dedicates a special issue to the topic of *High Resolution Earth Imaging for Geospatial Information* and to the workshop outcomes.

In parallel to the workshop, a call for papers was distributed to experts around the world. As a result, extended versions of eight workshop papers (out of 13 contributions submitted to the special issue), which have undergone a rigorous peer review process, are published in this special issue of PFG. A further paper submitted to this special issue, entitled *Assessment of reflectance calibration methods for ADS40 imagery* by Lauri Markelin et al., had to be published in the PFG issue 3/2012 due to restrictions of space.

The first paper by Daniela Poli and Pierre Soille presents a method for the refinement of digital surface models (DSM) by integrating a satellite image of higher resolution. Using a new segmentation technique, the heights inside homogeneous segments are adapted so that sharp edges at segment boundaries are obtained, which makes the method particularly well-suited for densely built-up urban areas, potentially improving the prospects of automated building detection. In the second paper, Norbert Haala and Mathias Rothermel apply the semi-global matching technique in a scenario consisting of multiple aerial images and assess the quality of the resulting DSM in terms of the accuracy and density of the underlying 3D point clouds as a function of the image configuration. Their results highlight the potential of dense matching techniques in combination with modern digital aerial cameras for producing high-quality DSMs. The third paper, written by Ali Özgün Ok et al., also deals with an aspect of image matching, but here the authors work on the reconstruction of 3D straight line segments from stereo images. Their focus is on improving the accuracy of 3D lines that are nearly aligned with the stereo base line, which is achieved by introducing fictitious 3D point entities into the estimation process.

DOI: 10.1127/1432-8364/2012/0119 1432-8364/12/0119 $ 0.50

The fourth and the fifth paper of this special issue are dedicated to different aspects of classification. First, Txomin Hermosilla et al. describe a method for periurban change detection. They achieve their goal by comparing classification results obtained from aerial data acquired at two epochs. In classification, they combine features from image and airborne laserscanner data, additionally using local context features. Their experiments highlight the impact of using 3D information on the quality of the overall process. In the fifth paper, Andreas Braun et al. investigate the potential of a state-of-the-art classifier, support vector machines (SVM), for data fusion. They exploit the typical application of SVM for multi-class problems to integrate expert knowledge into the fusion process, using data from a particular sensor only when it can actually contribute to the separation of two classes, thus reducing the dimensionality of the feature space to a minimum. They show two applications of their framework, fusing hyperspectral with airborne laser scanner data.

The last group of papers deals with satellite SAR data. Thierry Toutin et al. investigate the accuracy of of a digital elevation model (DEM) derived from Radarsat-2 high-resolution images by means of SAR stereo. In contrast to interferometry which is restricted to small spatial baselines the stereo technique requires large viewing angle differences. The authors show that for moderate terrain the bias and the standard deviation of the DEM are similar to the pixel spacing. The potential of urban classification from fully polarimetric SAR imagery is addressed by Mariam Soheili Majd et al. They derive 20 polarimetric and textural features from the data and investigate the performance of their approach, which is based on a Fisher distribution model, in comparison to other statistical parametric models as well as to non-parametric SVM. For some important classes their method provides best results. Even though the orientation of satellite SAR images is commonly rather accurate shortcomings in the DEM used during orthoprojection may lead to considerable misalignment. In order to tackle this problem Dimitra Vassilaki propose a novel method for matching and evaluating free-form linear features for georeferencing space-borne SAR imagery.

In summary, both the workshop and the selection of papers of this special issue show a number of trends. (1) Orientation of optical satellite images, an intensively debated topic in previous workshop editions, was hardly mentioned, probably owing to the fact that most questions have been successfully solved. (2) The automatic generation of digital terrain and digital surface models seems to be well established, the latest progress in dense matching allows tackling also urban areas, which was not possible a few years ago. (3) Radar images are on their way to becoming a standard input source in our field, both in isolation and in combination with optical image data. (4) The automatic extraction of GIS objects and the update of GIS databases remain the great challenges, full automation is not in sight, and thus semi-automatic approaches are needed for practical applications.

Finally, we would like to sincerely thank everybody involved in the preparation of this special issue. We are very grateful to all IPI staff for their invaluable help in organizing the workshop, to the workshop participants who have made the meeting a success, and to the authors of this special issue for making available their excellent papers, and for keeping a tough timeline. We thank the old and the new editors-in-chief of PFG, Helmut Mayer and Wolfgang Kresse, and their teams for all the freedom we have had when assembling this special issue and for the help they extended to us in the technical preparation. In particular, we are very grateful for the possibility to have the results of the workshop published in the PFG issue prepared for the XXII ISPRS Congress in Melbourne, Australia, in order to bring them to the attention of the world-wide community of photogrammetry, remote sensing and spatial information sciences.

Christian Heipke, Karsten Jacobsen, Franz Rottensteiner, Sönke Müller, Uwe Sörgel, Hannover

This PFG-issue also includes four articles from universities in Munich, Halle, Dresden, and Potsdam providing an insight into their current fields of research in Ph & RS as well as geographic information pursued in Germany.

Wolfgang Kresse

PFG 2012 / 4, 0317–0330
Stuttgart, August 2012 **Article**

Digital Surface Model Extraction and Refinement through Image Segmentation – Application to the ISPRS Benchmark Stereo Dataset

Daniela Poli & Pierre Soille, Ispra (VA), Italy

Keywords: high resolution, data fusion, digital surface models, segmentation

Summary: This paper proposes a methodology for the geometric refinement of a digital surface model (DSM) using high or very high resolution satellite scenes through an advanced hierarchical image segmentation and shows the results obtained on a dataset of Catalonia, Spain, provided by ISPRS WG I/4 through the project "Benchmarking and quality analysis of DEM generated from high and very high resolution optical stereo satellite data". During this experiment a WorldView-1 quasi-nadir scene was used for the enhancement of the DSM generated by a Cartosat-1 stereopair. The original and final DSMs are compared to the lidar DSM on the same area for quality analysis.

Zusammenfassung: Erstellung von Oberflächenmodellen und deren Verbesserung durch hierarchische Bildsegmentierung am Beispiel der ISPRS Benchmark-Stereodaten. Dieser Artikel stellt eine Methode zur geometrischen Verbesserung eines digitalen Oberflächenmodells (DSM) durch hierarchische Bildsegmentierung von hoch bzw. sehr hoch auflösenden Satellitenbildern vor. Die Methode wurde an einem Datensatz aus Katalonien in Spanien getestet, der im Rahmen des ISPRS WG I/4 Projektes „Benchmarking and quality analysis of DEM generated from high and very high resolution optical stereo satellite data" zur Verfügung gestellt wurde. In diesem Experiment wurde ein aus einem Cartosat-1-Stereopaar erstelltes DSM durch eine WorldView-1 Quasi-Nadir-Szene verbessert. Zur Bewertung der Methode wurde ein Lidar DSM als Referenz verwendet. Der Artikel beschreibt die Methode sowie die Ergebnisse des Tests und der Qualitätsanalyse.

1 Introduction

In the last years the number of Earth-observation platforms equipped with sensors delivering very high resolution (VHR) images with ground sample distances (GSD) smaller than 1 m has increased. Those sensors are characterized by different geometric, radiometric, spectral and operational specifications, and in most cases allow for the acquisition of multi-angle imagery for automatic digital surface models (DSM) generation and 3D feature extraction. One of the tasks of the Joint Research Centre of the European Commission in Ispra (Italy) is to assess changes in land cover after events like natural disasters or conflicts. VHR satellite images are generally used for this scope, as they almost cover any area of the planet, in particular remotely located areas. To describe the pre-event and post-event scenarios the available data are analysed and eventually fused in order to achieve the most accurate and reliable information. With respect to automatic 3D information extraction, the availability of accurate and detailed DSMs is a crucial issue for automatic building detection and subsequent damage estimation. While the acquisition of VHR stereo scenes can be planned after the event and used to model the post-event 3D scenario, VHR stereo scenes are generally not available for data acquisition before the event; on the other hand, single

DOI: 10.1127/1432-8364/2012/0120 1432-8364/12/0120 $ 3.25

VHR scenes are likely to be available in the archives, as well as DSMs at medium or low resolution from multi-line optical sensors with simultaneous along-track stereo acquisition, like SPOT-5/HRS, Cartosat-1, ALOS-PRISM, or other sources. In this context we developed a strategy for data fusion, in order to enhance the surface models using information from single scenes. It is enhanced in the sense that the final DSM is sharper and therefore likely to be more suitable for subsequent 3D shape detection. We tested the method on the dataset provided by ISPRS Working Group I/4 within the project "Benchmarking and quality analysis of DEM generated from high and very high resolution optical stereo satellite data" (Reinartz et al. 2010). The approach and first results were presented at the Workshop "High-Resolution Earth Imaging for Geospatial Information" held in Hannover (Germany) in June 2011 (Poli & Soille 2011); in this work the latest developments in the methodology and the achieved results are reported and discussed.

2 Methodology

2.1 *Related Works*

The combined use of digital imagery and surface models mainly finds applications in automatic object extraction and 3D reconstruction from aerial data or very high resolution satellite sensors (Baltsavias et al. 1995, Jaynes et al. 1996, Paparoditis et al. 1998, Lu et al. 2002, Tao & Yasuoka 2002, Lu et al. 2006, Li et al. 2008, Arefi 2009). Few works aim at using satellite imagery for DSM enhancement. In Maire (2010), user-defined semantic contents (sea, lakes, buildings, roads, etc.) are extracted using a supervised classification in high resolution satellite imagery; then in correspondence of each class the surface is modelled by plane surfaces with geometric constraints given by the topological properties of each class and neighbour regions. This approach was applied, for example, for the enhancement of SRTM-X using a 2.5 m ground resolution SPOT-5 scene. In the computer vision community, several matching approaches employ image segmentation during the generation of the DSM (Klaus et al. 2006, Krauss & Reinartz 2010). For example, in the approach proposed by Krauss & Reinartz (2010), one stereo image is segmented and transferred to the disparity map, then for each segment the original disparity map is filled with suitable interpolation of the disparities. In the proposed method, segmentation of a VHR scene is used to refine a given DSM at a coarser resolution.

2.2 *Proposed Method*

The principal steps of the proposed approach are summarized in Fig. 1. The initial DSM can be any surface model generated by matching techniques from stereo images or other sources. Unless already orthorectified, the VHR scene is oriented using ground control information and orthorectified on the initial DSM. The generated orthophoto is segmented with alpha-omega connectivity (Soille 2008) and the segments are overlaid to the existing DSM. After calculating the statistics (mean, median, sigma, minimum and maximum values) of the heights of the points falling into each segment, the new surface model is calculated from the existing one by imposing that the height values that fall into the same segment follow a certain mathematic function, i.e. constant value or planar surface. In the initial version of the algorithm, the new height values for each segment were calculated as the minimum, mean or maximum values of the original DSM heights, or as the result of planar fitting.

A more advanced approach for the calculation of the enhanced height values was then developed. It is based on exploiting not only the radiometric information of the VHR scene, but also the information contained in the DSM shape. The aim is to enhance the surface model using the maximum value in segments deemed to correspond to a building, and using the minimum value for segments supposed to match ground areas between buildings. Buildings are generally represented in the DSM as convex surfaces while ground areas between buildings are likely to match concave surfaces (the term convex and concave are used in the sense "convex from above" and "concave from above", respectively). Therefore, the original DSM is segmented to extract concave, convex

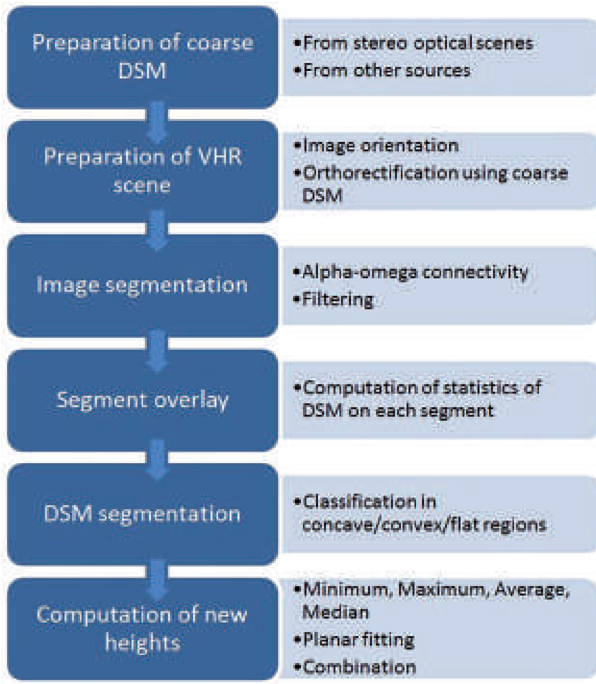


Fig. 1: Proposed workflow for DSM enhancement.

and flat regions, and then the new height values are computed differently, depending on whether the segment falls in a concave, convex or flat region. The decision rule is:

- Use minimum value if the segment falls in a concave region;
- Use maximum value if the segment falls in a convex region;
- Use median value or planar fitting if the segment falls in a flat region.

In practice, to exclude possible outliers, the 10 % (resp. 90 %) percentile is used instead of the minimum (resp. maximum) value. In the next paragraphs the algorithms used for image and DSM segmentation are described.

2.3 *Image Segmentation by Constrained Connectivity*

Given a logical predicate, the segmentation of an image is defined as the partition of the image definition domain into non-overlapping regions called segments, such that the logical predicate returns true on every segment but false on the union of any pair of adjacent segments (Horowitz & Pavlidis 1976). Recently, a segmentation technique producing a hierarchical partitioning of the image definition domain under connectivity constraints corresponding to logical predicates was proposed (Soille 2008). Given the input constraints, the resulting partition is uniquely defined, contrary to most segmentation techniques. By increasing the threshold levels associated to each connectivity constraint, one obtains a series of fine-to-coarse partitions with the pixels at the lowest level and the whole image definition domain at the highest level of the hierarchy.

Constrained connectivity relies on a dissimilarity measure computed between each pair of adjacent pixels. The absolute difference between the intensity values was used for all experiments hereafter (see Soille 2011 for other dissimilarity measures). Two pixels are said to be alpha-connected if there exists a path linking these two pixels in such a way that the dissimilarity between all pairs of adjacent pixels of the path does not exceed the value of alpha. The relation 'to be alpha-connected' is an equivalence relation (reflexive, symmetric, and transitive relation) and therefore it leads to a unique partition of the image definition domain into regions of maximal extent called alpha-connected components.

The constrained connected component of a given pixel is then defined as the largest alpha-connected component of this pixel that satisfies a series of constraints. Any number of constraints may be considered. Given a local dissimilarity defined as the absolute difference, the most natural constraint is defined in terms of a threshold level on the range of the grey level values of the connected components. This constraint is called the omega-constraint and its associated threshold value is the omega threshold. Denoting by α and ω the values of the alpha and omega thresholds, respectively, the (α,ω)-connected component of a pixel is defined as the largest α'-connected component such that α' is less than or equal to α and the range of the α'-connected component is less than or equal to ω (Soille 2008). The omega constraint prevents linking through transitions, a well-known problem of single-linkage clustering.

A typical choice is to consider the same values for the alpha and omega thresholds. By increasing this value, a fine-to-coarse hierarchy of connected components is obtained. In practice, it is desirable to obtain segmentation as coarse as possible but retaining all structures of interest in the image given the considered application. In this study, since we are interested in buildings and other man-made structures, the parameters need to be selected in such a way that these objects are simplified as much as possible, i.e. ideally are matched by one connected component, but not merged with other objects.

To further favour linking within homogeneous regions while preventing it through transitions, an image pre-processed by the edge sharpening technique described in (Soille 2011) is applied before segmentation. The pre-processing consists in considering as seeds all local maxima of the input image and propagating their values so as to cover the whole image definition domain. The propagation is controlled by the original image values (the smaller the difference between a seed and a non-seeded adjacent point, the greater the speed). Alternatively, rather than pre-processing the image, the dissimilarity measure proposed in (Soille 2011) could be used to prevent linking through transitions while favouring it within homogeneous regions.

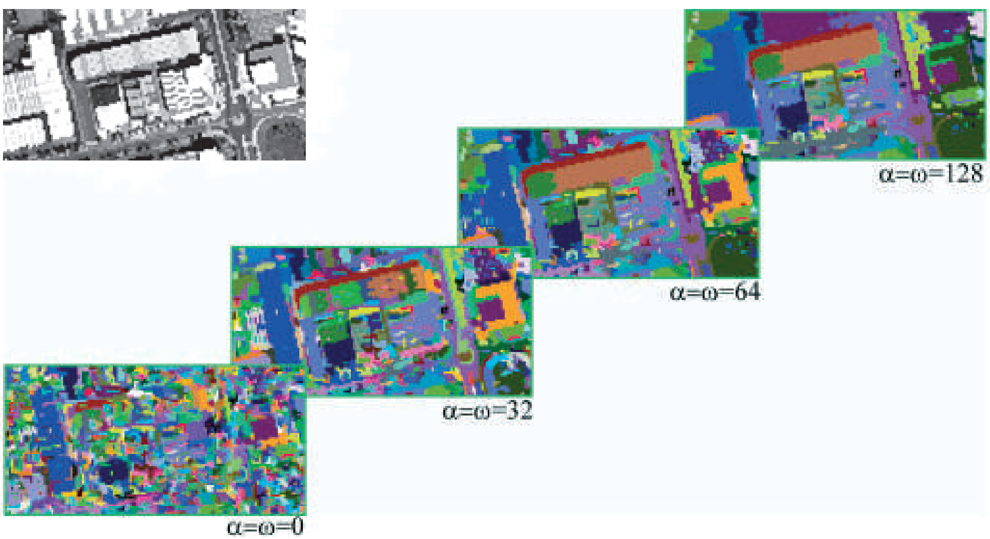


Fig. 2: Example of constrained connectivity partitioning of a 512 × 256 sample of a WorldView-1 image resampled at 2.5 m, rescaled to byte data type, and contrast enhanced. The partitions are displayed for the alpha and omega thresholds values equal to 0, 32, 64, and 128 respectively (from fine to coarse partitions).

An example of constrained connectivity partitioning is shown in Fig. 2. In this figure, the original image is partitioned with increasing values of alpha and omega parameters. To highlight the hierarchical property of the resulting partitions, random colours were given to the connected components of the finest partition while the colour of a given connected component of a subsequent level is inherited from the largest connected component of the previous level, contained by this connected component.

2.4 *DSM Segmentation*

The approach used for DSM segmentation and identification of concave, convex and flat regions was proposed by Pesaresi & Benediktsson (2001) and applied to satellite images. The method is a morphological segmentation by the derivative of the morphological profile, based on the use of residuals from opening and closing (levelling) by reconstruction.

A structure or an "object" in an image is defined as a connected component of pixels sharing the same morphological characteristics, i.e. "flat", "concave", "convex" in case of the local curvature of the grey level function surface. The residuals between any opening by reconstruction (or closing by reconstruction) and their original function are interpreted as a measure of the relative brightness of the structure (or relative darkness). Then, one membership function can be defined corresponding to the class "convex" and another membership function corresponding to the class "concave". The levelling algorithm is rewritten as a decision rule based on the greatest value of the membership function. The segmented image of the characteristic is a tessellation of three different labels like "convex", "concave" and "flat". For this segmented image, pixels where the lower levelling is strictly lower than the original image are labelled "convex", and pixels where the upper levelling is strictly greater than the grey function are labelled as "concave". Pixels labelled "flat" have maintained the same value of the grey function in both the upper and lower levelling. Therefore, those pixels have been indifferent to the erosion/dilation-reconstruction process with a given structuring element. In experiments with satellite images, the proposed method demonstrated good performance in terms of shape description and retained small but significant regions in images.

3 Experimental Results

In our experiment the algorithms described in section 2 were used for the enhancement of a DSM generated from a stereopair acquired by Cartosat-1 (C1) sensor using a WorldView-1 (WV1) scene. In the next paragraphs the data and detailed processing will be presented and discussed.

3.1 *Data Description*

The Working Group 4 of Commission I on "Geometric and Radiometric Modelling of Optical Spaceborne Sensors" is providing on its website several stereo datasets from high and very high resolution spaceborne stereo sensors on three areas in Catalonia, Spain, covering urban, rural, and forest areas in flat and medium undulated terrain as well as steep mountainous terrain. In addition to these datasets, a lidar DSM generated by the Institut Cartogràfic de Catalunya (ICC) is provided as a reference for comparison (Reinartz et al. 2010). Among the three available regions (Terrassa, La Mora, Vacarisses), the dataset located over Terrassa was chosen for our experiments.

The dataset consists of:

- a subset of a stereopair acquired by WV1 on August 28, 2008, composed by a nadir (-1.3º) and off-nadir (33.9º) scene, with ground sample distances of 0.50 m and 0.76 m, respectively; both scenes are panchromatic and have size 10,000 x 10,000 pixels; rational polynomial coefficients (RPC) are available; for our experiment only the nadir scene is used;
- a stereopair acquired by C1 in February 2008, composed of a backward (-5º) and forward (33.9º) looking scene, with ground sample distance of 2.5 m; both scenes are panchromatic; RPCs are available;
- ground coordinates of points;

- a dense 3D point cloud acquired on November 27, 2007 with airborne lidar.

The area is characterized by urban, rural and forest cover on flat and hilly terrain (Reinartz et al. 2010). The terrain elevation ranges from 200 m to 430 m.

As explained in the introduction, in this paper we propose an approach to refine a coarse resolution DSM using a single VHR scene, as generally VHR stereopairs are not available to describe pre-event scenarios. Therefore, even if a VHR stereopair is available, from the Terrassa dataset we used only the C1 stereopair to generate a DSM and the nadir WV1 scene to refine it. The processing workflow includes the following steps:

- orientation of C1 stereo scenes,
- DSM generation at 2.5 m grid spacing (DSM_C1),
- orientation of WV1 nadir scene,
- orthorectification of WV1 scene using DSM_C1 at ground sample distance of 2.5 m,
- segmentation of the WV1 orthoimage,
- enhancement of DSM_C1 with generation of a new DSM at 2.5 m grid spacing.

3.2 *Orthorectification and DSM Generation*

For the photogrammetric processing of C1 and WV1 scenes the commercial software SAT-PP (SATellite image Precision Processing) by 4DiXplorer AG was used (Zhang 2005).

The C1 stereopair was oriented by estimating the correction of the available RPC through an affine transformation using nine ground control points (GCPs). Then a DSM was generated using the advanced approach in SAT-PP based on a coarse-to-fine hierarchical solution with an effective combination of several image matching methods and automatic quality indication (Zhang & Grün 2004). To improve the conditions for feature extraction and matching, the images were preprocessed using Wallis filter. Few seed points were measured in the epipolar images in stereo mode to provide the initial approximation of the surface. The DSM was generated with a regular grid space of 2.5 m (DSM_C1, Fig. 3). No further editing was applied on the surface model. In the DSM it is possible to distinguish urban areas, but buildings are not well defined from adjacent buildings or roads. In open areas, the surface model follows the terrain shape without outliers.

The nadir WV1 scene was georeferenced, with the estimation of affine correction parameters of the given RPCs by means of GCPs. This was followed by the orthorectification, using the DSM_C1 as elevation model. As a result an orthoimage at 2.5 m GSD was produced.

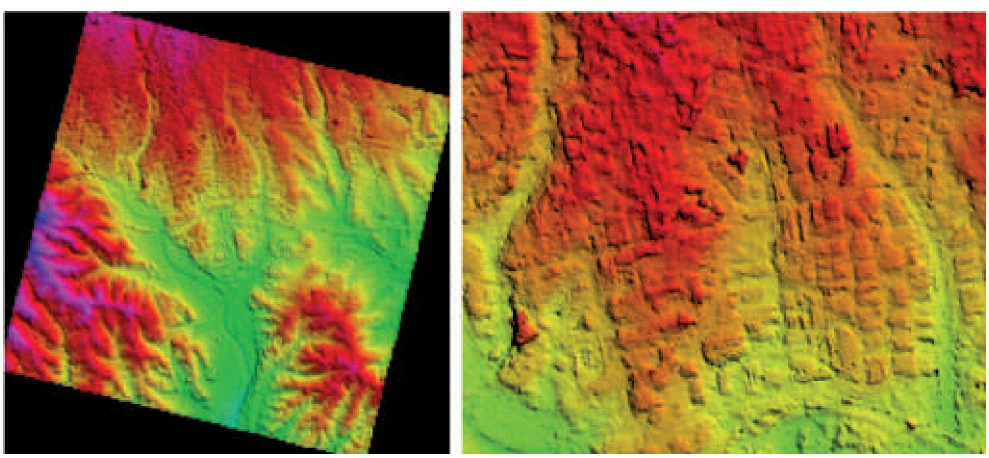
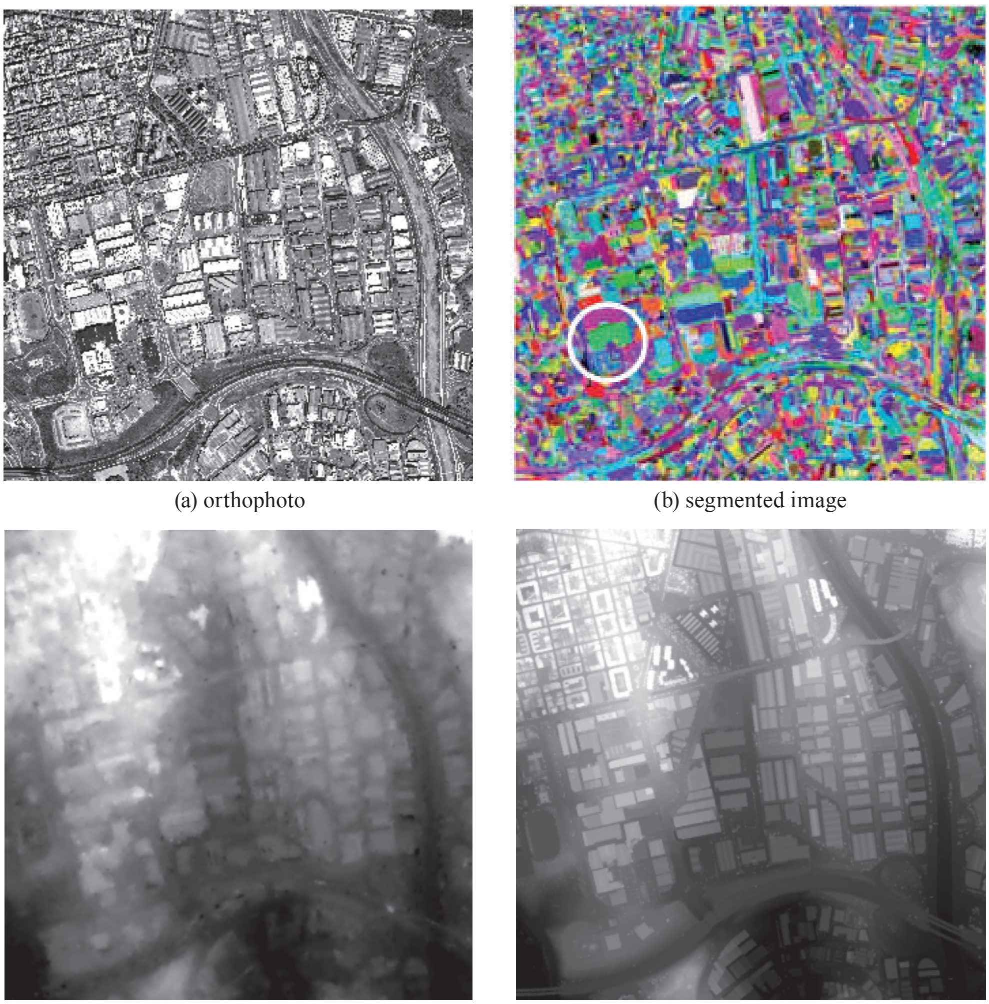


Fig. 3: Colour-shaded visualization of DSM generated from Cartosat-1 stereo scenes over Terrassa area. Overview of the full model (left) and zoom on urban and open areas (right).

3.3 *Segmentation and DSM Enhancement*

The WV1 orthoimage was pre-processed with edge-sharpening and segmented. For the segmentation, a value of 50 intensity levels for both the alpha and omega thresholds was found to be optimal, in the sense that it enabled the formation of connected components matching the subparts of the building roofs while avoiding their merging with adjacent objects. An example is given in Fig. 4 (a) and (b) on a subset with buildings, main and secondary roads on different levels, a river with bridges, open areas with vegetation. Following the methodology previously described, four DSMs were calculated from the original DSM, using the minimum value (ENH_MIN), the maximum value (ENH_MAX), the mean value (ENH_AVG) of the heights in each segment, and planar fitting (ENH_PLA). From a visual comparison of the original DSM (Fig. 4 (c)) and the enhanced DSMs with average and planar fitting (Fig. 5 (a) and (b) respectively) on the subset mentioned above, it is evident that in the enhanced DSMs the building edg-



(a) orthophoto (b) segmented image

(c) DSM_C1 (d) Lidar DSM

Fig. 4: Results of DSM enhancement over a subset. (a) WV1 orthoscene at 2.5 m grid spacing, (b) segmentation, using alpha = omega = 50 over a subset, (c) original Cartosat-1 DSM, (d) reference lidar DSM used for quality analysis.

es are more delineated and objects at different heights, like rivers and bridges, are better defined than in DSM_C1. In homogeneous areas (roof faces, roads, parking areas, fields, and so on) small details due to noise are removed. For the purposes of our research, we focused the analysis on buildings. With respect to the function used to enhance the DSM, there is not a significant difference between average and planar fitting at this resolution, as in DSM_C1 itself it is not possible to recognize shape details, like roof faces, with sufficient accuracy.

The enhanced DSMs obtained using the minimum and maximum values (ENH_MIN and ENH_MAX) were compared to the reference one and to DSM_C1 to evaluate if the original errors of the DSM_C1 were increasing or decreasing with the respect to the lidar after the refinement. As expected, we notice in Fig. 6 that by applying the minimum values, the errors in correspondence of ground level decreases, while using the maximum value the enhanced DSM improves in correspondence of buildings. Therefore the criteria should be combined. The original DSM was segmented in concave, convex and flat regions and the new DSM was calculated with minimum and maximum height values in segments falling in concave and convex regions respectively, and medium value elsewhere (ENH_HYB).

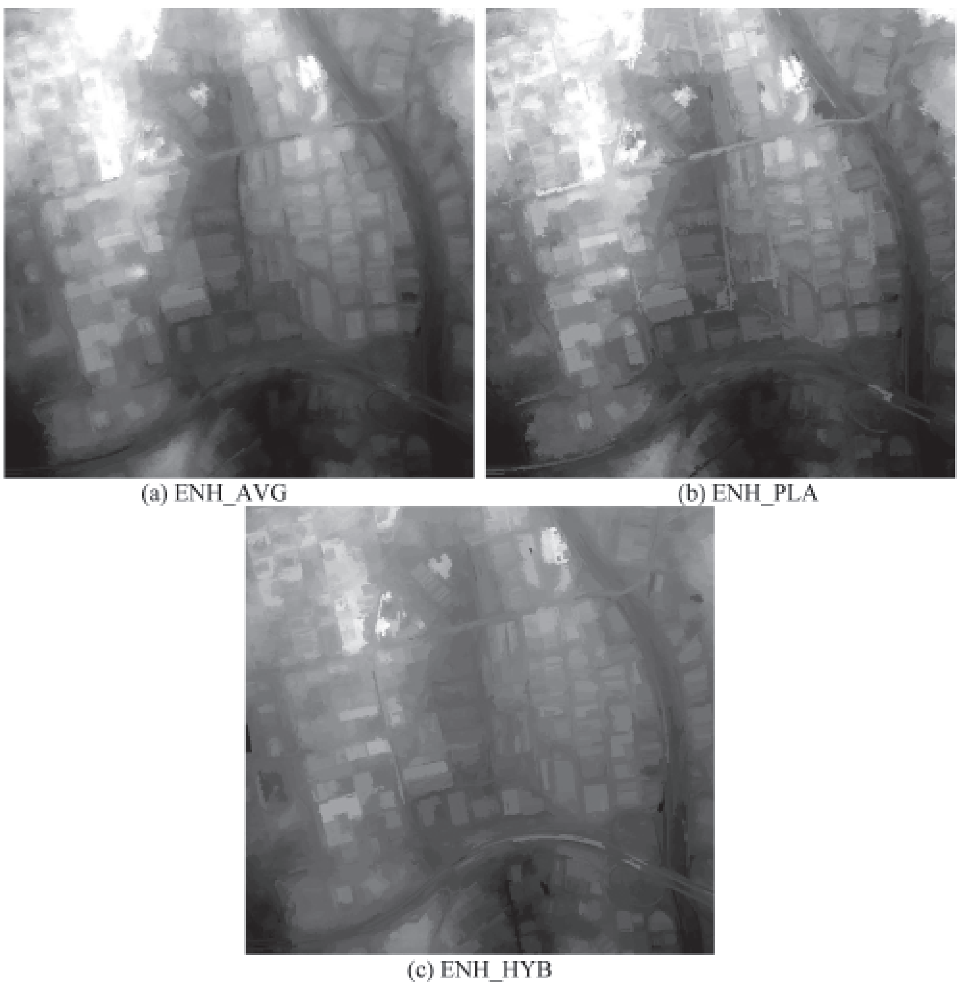


Fig. 5: Results of DSM enhancement over a subset using (a) average values, (b) planar fitting, (c) a combination of minimum, maximum and median values.

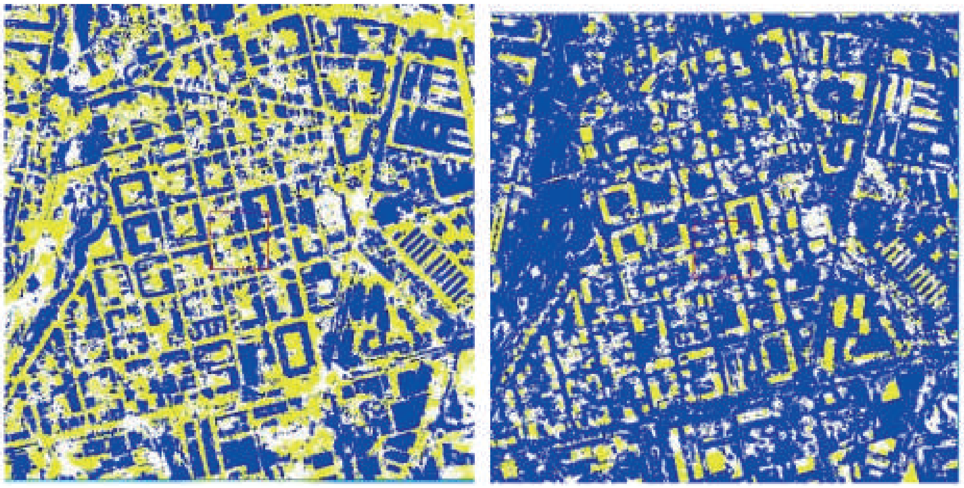


Fig. 6: Zoom on height gradients of ENH_MIN (left) and ENH_MAX (right): yellow = error decreases, blue = error increases, white = error does not change.

3.4 *Quality Analysis*

For the quality analysis of the enhanced DSM, the lidar DSM was used as reference (Fig. 4 (d)). We evaluated the height differences through the surface profiles in dense urban and industrial areas. Some examples are reported in Fig. 7.

With respect to the original DSM, the enhanced one produces step profiles in correspondence of buildings external edges, as in profile 2, and in some cases the algorithm detected buildings that were missing in the original DSM (profiles 1 and 5). The enhancement is not sensitive to the terrain slope (profile 4). Obviously the accuracy of the enhanced DSM depends on the accuracy of the initial DSM. Indeed small corridors between buildings are well defined in the lidar DSMs, but not in DSM_C1; as a consequence the ground height in the enhanced DSM is similar to the one in the original DSM, but the profiles of the buildings are improved (profiles 1, 2 and 5).

The presented method shows a limitation in case of complex buildings with homogeneous texture. With reference to profile 2, the building portions at different heights cannot be distinguished in the panchromatic channel (white circle in Fig. 4 (a)), resulting in a single segment after partitioning (white circle in Fig. 4 (b)), and a single height value for that segment in the enhanced DSM. A refining of the segmentation with additional information, like multispectral channels, could improve the identification of sub-structures of the roofs at different heights.

The statistics on the height differences between the original and enhanced DSMs with respect to the lidar one are summarized in Tab. 1. The mean error, the mean absolute

Tab. 1: Statistics of height differences between the original and refined DSM with respect to lidar along five transects (μ = mean value, MAD= mean absolute deviation, RMSE = root-mean-square error) (m).

	Profile 1			Profile 2			Profile 3			Profile 4			Profile 5		
	μ	MAD	RMSE	μ	MAD	RMSE	μ	MAD	RMSE	μ	MAD	RMSE	μ	MAD	RMSE
Original	-1.64	2.91	3.78	-2.00	3.37	4.15	0.70	2.41	3.25	-1.22	2.59	3.21	-4.21	4.86	5.88
Refined	-1.03	2.90	3.89	-1.18	2.08	2.77	0.40	1.90	1.87	-0.51	2.50	3.35	-2.87	4.06	4.92
# points		37				70			140			209			74

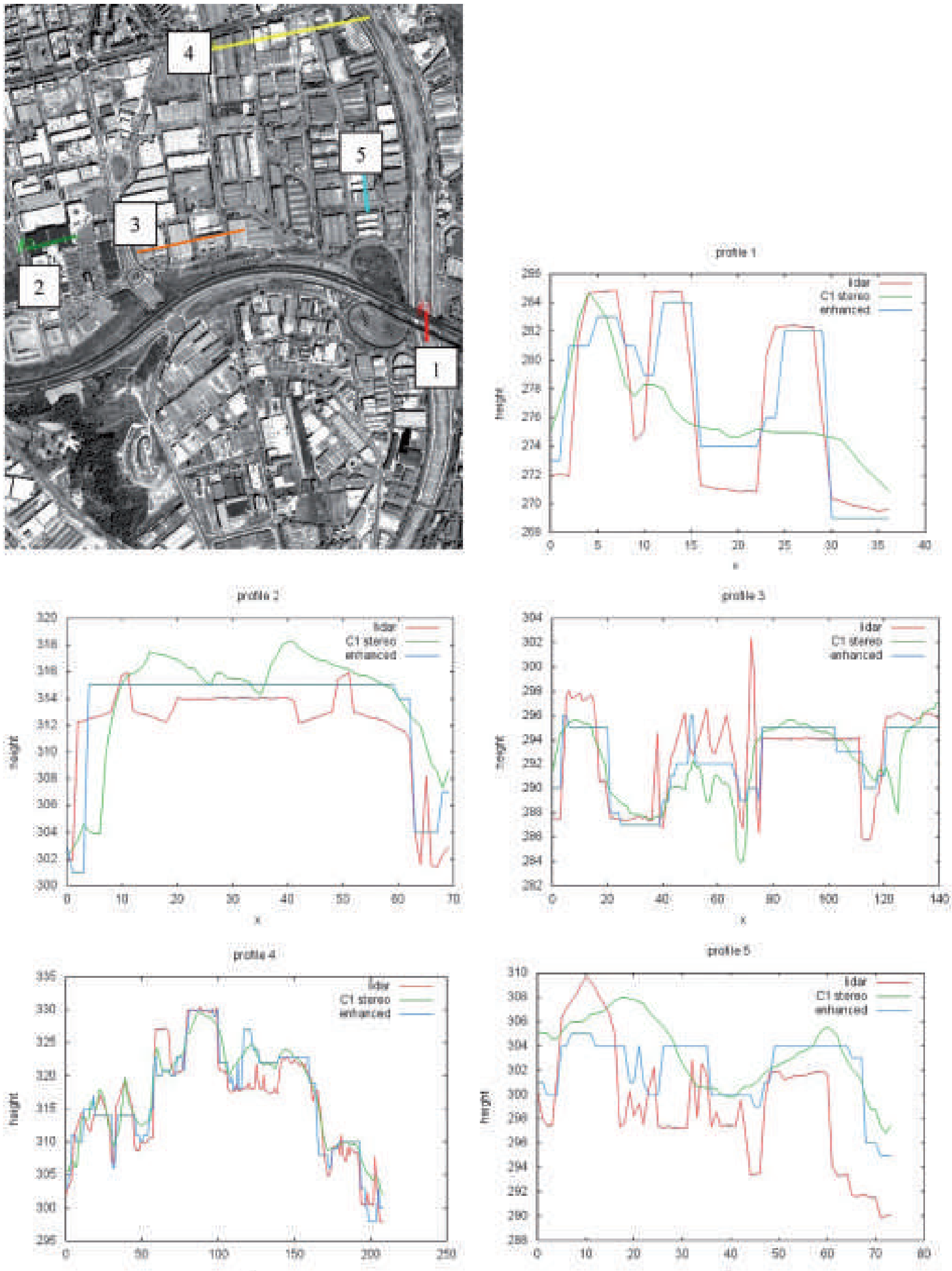


Fig. 7: Profiles of original C1 stereo DSM (green), lidar DSM (red) and ENH (blue) along the five transects shown in the upper left image.

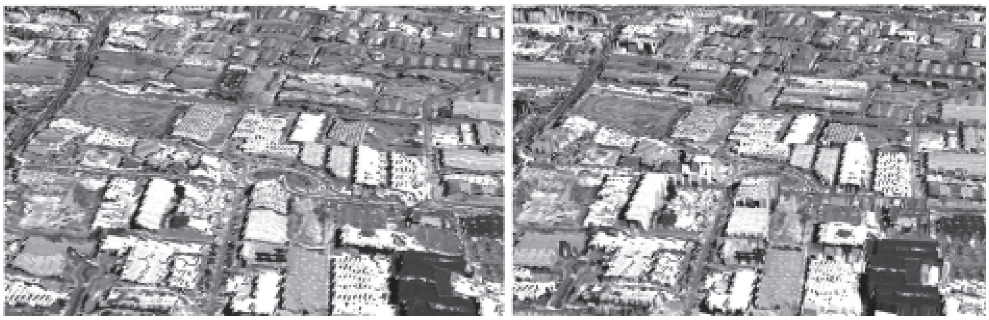


Fig. 8: Perspective 3D visualization of original DSM (DSM_C1, left) and enhanced DSM (ENH_HYB, right), both with vertical exaggeration 2.0, texture with C1 orthoimage (2.5 m).

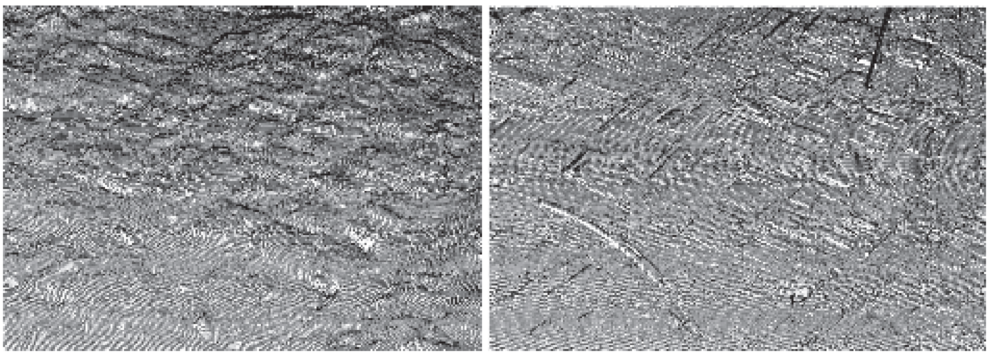


Fig. 9: Perspective 3D visualization in wireframe mode of original DSM (DSM_C1, left) and enhanced DSM (ENH_HYB, right), both with vertical exaggeration 2.0.

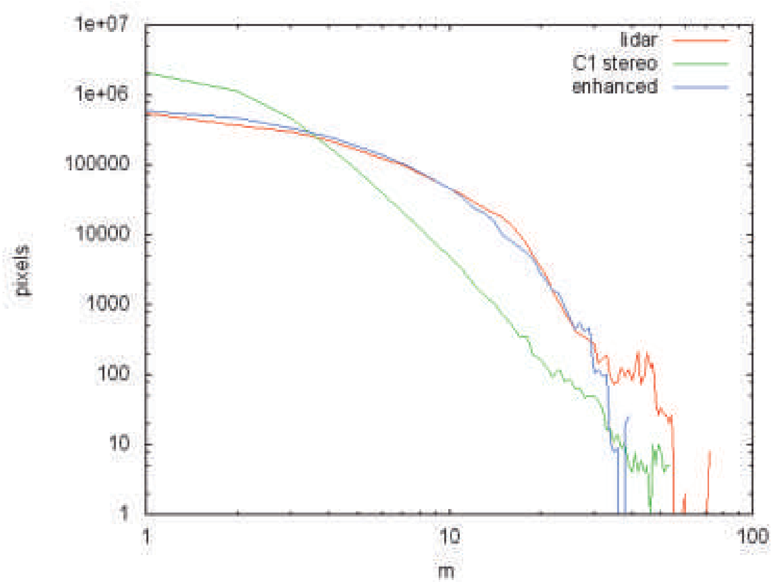


Fig. 10: Histogram of the local slope of lidar, original C1 stereo and enhanced DSM (logarithmic scale).

deviations and RMSE are decreasing for the refined DSMs, except for non-significant increase of the RMSE of profile 1.

The 3D visualizations of the original and enhanced DSMs in Figs. 8 and 9 show that after the refinement, small blobs at the ground level are removed and the building edges are sharper and more realistic than in the original DSM. This improvement is confirmed by the analysis of the distribution of the local slope, calculated as the difference between the maximum and minimum height within a 3 by 3 neighbourhood. This analysis is carried out over the lidar, C1 stereo, and enhanced DSMs, see Fig. 10. Indeed, the distribution of the slope of the original DSM gets substantially closer to that of the lidar after enhancement, apart from a few outliers at the tail of the distribution.

4 Conclusions

In this paper we presented an approach for the enhancement of digital surface models using very high resolution satellite images. After orthorectification, the images are segmented through an advanced hierarchical image partitioning, then the segments are projected on the existing DSM. The new height values are calculated for each segment with planar fitting or constant (average) value of the original DSM heights contained in that segment. The method was applied on a dataset over Catalonia, Spain, provided by ISPRS WG I/4 through the project "Benchmarking and quality analysis of DEM generated from high and very high resolution optical stereo satellite data". During this experiment a DSM was generated by a Cartosat-1 stereopair and a WorldView-1 (WV1) quasi-nadir scene was used for the DSM enhancement. The original and final DSMs were compared to the lidar DSM on the same area for quality analysis. The analysis of the profiles (with statistics), and local slopes and the 3D visualization shows that in the enhanced DSM buildings contours are sharper and more realistic than in the original one, which is favourable for automatic building detection. Further investigations will also include the use of multispectral information to improve the image segmentation and practical experiments of damage assessment using automatic building detection in the original and enhanced DSMs.

Acknowledgements

Special thanks are given to the data providers for the provision of the stereo datasets, namely: Euromap for the Cartosat-1 data, Digital Globe for the WorldView-1 data and ICC Catalunya for the reference data.

References

Arefi, H., 2009: From LIDAR Point Clouds to 3D Building Models. – PhD thesis, Bundeswehr University Munich.

Baltsavias, E.P., Mason, S. & Stallmann, D., 1995: Use of DTM/DSMs and Orthoimages to Support Building Extraction. – Gruen, A., Kuebler, O. & Agouris, P. (eds.): Automatic Extraction of Man-Made Objects from Aerial and Space Images. – Birkhäuser Verlag, Basel, 199–210.

Jaynes, C.O., Collins, R.T., Stolle, F.R., Hanson, A.R., Riseman, E.M. & Schultz, H.J., 1996: Three-Dimensional Grouping for Site Modeling from Aerial Images. – Arpa Image Understanding Workshop: 223–236.

Horowitz, S. & Pavlidis, T., 1976: Picture segmentation by a tree traversal algorithm. – Journal of the ACM **23** (2): 368–388.

Klaus, A., Sormann, M. & Karner, K., 2006: Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. – **18th** International Conference on Pattern Recognition 2006: 15–18.

Krauss, T. & Reinartz, P., 2010: Enhancement of dense urban digital surface models from VHR optical satellite stereo data by pre-segmentation and object detection. – The International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences **XXXVIII** (1).

Li, Y., Zhu, L. & Shimamura, H., 2008: Integrated method of building extraction from digital surface model and imagery. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XXXVII** (B3b).

Lu, Y., Trinder, J. & Kubik, K., 2002: Automatic building extraction for 3D terrain reconstruction using interpretation techniques. – Techniques, School of Surveying and Spatial Information

Systems, University of New South Wales, NSW 2052 Australia.

Lu, Y., Trinder, J. & Kubik, K., 2006: Automatic Building Detection Using the Dempster-Shafer Algorithm. – Photogrammetric Engineering & Remote Sensing **72** (4): 395–403.

Maire, C., 2010: Image Information Extraction and Modeling for the Enhancement of Digital Elevation Models. – PhD Thesis, Karlsruhe Institute of Technology (KIT).

Paparoditis, N., Cord, M., Jordan, M. & Cocquerez, J.P., 1998: Building Detection and Reconstruction from Mid- and High-Resolution Aerial Imagery. – Computer Vision and Image Understanding **72** (2): 122–142.

Pesaresi, M. & Benediktsson, J.A., 2001: A new approach for the morphological segmentation of high-resolution satellite imagery. – IEEE Transactions on Geoscience and Remote Sensing **39** (2): 309–320.

Poli, D. & Soille, P., 2011: Refinement of digital surface models through constrained connectivity partitioning of optical imagery. – The International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences **XXXVIII** (4/W19).

Reinartz, P., d'Angelo, P., Krauss, T., Poli, D., Jacobsen, K. & Buyuksalih, G., 2010: Benchmarking and quality analysis of DEM generated from high and very high resolution optical stereo satellite data. – The International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences **XXXVIII** (1).

Soille, P., 2008: Constrained connectivity for hierarchical image partitioning and simplification. – IEEE Transactions on Pattern Analysis and Machine Intelligence **30** (7): 1132–1145.

Soille, P., 2011: Preventing chaining through transitions while favouring it within homogeneous regions. – ISMM 2011, Lecture Notes in Computer Science **6671:** 96–107.

Tao, G. & Yasuoka, Y., 2002: Combining high resolution satellite imagery and airborne laser scanning data for generating bare land DEM in urban areas. – The International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences **XXXIV** (5/W3), Kunming, China.

Zhang, L., 2005: Automatic Digital Surface Model (DSM) Generation from Linear Array Images. – PhD Dissertation, No. **88**, Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland.

Zhang, L. & Grün, A., 2004: Automatic DSM generation from linear array imagery data. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XXXV** (3): 128–133.

Address of the Authors:

Dr.-Ing. Daniela Poli, Dr.-Ing. habil. Pierre Soille, European Commission – Joint Research Centre, Institute for the Protection and Security of the Citizen, Global Security and Crisis Management Unit, Geo-Spatial Information Analysis for Global Security and Stability (ISFEREA) Action, Via E. Fermi 2749, I-21027 Ispra (VA) Italy, TP 267, Tel. +39-0332-789241, -785068, Fax +39-0332-785154, e-mail: first name.last name@jrc.ec.europa.eu

Manuskript eingereicht: September 2011
Angenommen: November 2011

PFG 2012 / 4, 0331–0344
Stuttgart, August 2012

Article

Dense Multi-Stereo Matching for High Quality Digital Elevation Models

Norbert Haala & Mathias Rothermel, Stuttgart

Keywords: DEM, surface, point cloud, image matching, geometry

Summary: Until recently, the acquisition of high quality digital elevation models was dominated by the use of airborne lidar. However, meanwhile the increasing quality of digital airborne cameras in combination with recent improvements in matching algorithms allow for an automatic image-based acquisition as a suitable alternative. Within the paper, this progress will be demonstrated on the example of photogrammetric DEM generation using the semi-global matching (SGM) stereo method. Since this approach aims at pixel-wise matching, dense 3D point clouds can be generated. The tests described in the paper are based on data collected from different digital airborne cameras at various flight scenarios during a recent test on photogrammetric 3D data capture. Thus, the impact of different stereo configurations on the quality of the final outcome can be evaluated and compared to test results already available. Special interest is also paid to the combination of multiple stereo pairs with different base-to-height ratios, which can be used efficiently to increase the precision and reliability of the matching results.

Zusammenfassung: *Hochwertige Digitale Höhenmodelle aus dichter Mehrbildstereozuordnung.* Die Erzeugung digitaler Höhenmodelle wurde bis vor kurzem durch den Einsatz von Lidar dominiert. Durch die verbesserte Qualität von Kamerasystemen und die voranschreitende Entwicklung bildbasierter Zuordnungsverfahren stellt die Datenerfassung mittels Luftbildkameras mittlerweile eine ernstzunehmende Alternative dar. In diesem Artikel wird dieser Fortschritt anhand des Beispiels der photogrammetrischen DEM Generierung durch Semi-Global Matching (SGM) aufgezeigt. Mittels dieser Methode werden pixelweise Zuordnungen zwischen zwei Bildpaaren erhalten, so dass extrem dichte Oberflächenmodelle generiert werden können. Für die in diesem Artikel durchgeführten Tests wurde Bildmaterial unterschiedlicher Kamerasysteme und Flugkonfigurationen untersucht. Dabei wurde unter anderem auf Daten im Rahmen kürzlich durchgeführter Tests zur Untersuchung photogrammetrischer Datenerfassung zurückgegriffen. So kann der Einfluss von flugspezifischen Parametern auf die erhaltenen Resultate evaluiert, und selbige mit bestehenden Ergebnissen verglichen werden. Spezielle Beachtung findet außerdem die Fusion mehrerer Stereomodelle resultierend aus Bildpaaren unterschiedlicher Basis-Höhenverhältnisse. Dies ermöglicht eine Steigerung hinsichtlich der Genauigkeit und der geometrischen Qualität.

1 Introduction

For a considerable period, the acquisition of high quality digital elevation models (DEM) was dominated by the use of airborne lidar. Meanwhile, automatic image-based data collection has been revived as a suitable alternative. This development was triggered by the increasing quality of digital airborne cameras as well as recent innovations in matching algorithms. Airborne imagery of good dynamic range and signal-to-noise ratio as it is generated by digital aerial cameras is highly beneficial for automatic image matching. This is especially true for surfaces with relatively little surface texture. Consequently, the quality and precision of image-based point transfer as the basis for 3D surface reconstruction has been improved considerably. Recent tests have already demonstrated the feasibility of image

DOI: 10.1127/1432-8364/2012/0121 1432-8364/12/0121 $ 3.25

matching as a valid alternative to airborne lidar (Haala et al. 2010, Leberl et al. 2010).

Commercial image matching software tools used for DEM generation are still based on algorithms established a relatively long time ago. Usually, they apply standard feature and intensity-based matching. However, standard feature-based approaches can suffer from regions of a rather limited number of matched points in areas of low image texture. On the other hand, with intensity-based approaches, the minimum size of correlation windows frequently results in smoothing effects. This is especially disadvantageous at object borders and height discontinuities. These limitations of traditional approaches can be overcome by current algorithms aiming at a high resolution DEM computation. One example is the semi-global matching (SGM) stereo method as proposed by Hirschmüller (2008), which provides a match for each image pixel. Thus, a computation of dense 3D point clouds and DEM at resolutions similar to the ground sampling distance of the available imagery is feasible. The potential of the SGM algorithm was already demonstrated for different applications and datasets, including aerial images, satellite data or video sequences. This was our motivation to implement and use SGM for our evaluations on dense image matching.

In addition to suitable stereo matching algorithms, the quality of image-based surface reconstruction mainly depends on the proper combination of multiple image information. Multi-view stereo reconstruction has a considerable tradition in photogrammetry (Grün & Baltsavias 1988) as well as for close range applications in the computer vision community. As also documented by the performance evaluation of multi-view stereo reconstruction algorithms (Seitz et al. 2006), a number of high-quality algorithms are available and the state-of-the-art still improves rapidly. Similarly, airborne scenarios also benefit from multiple overlapping stereo pairs. This enables a reliable elimination of outliers and reduction of high frequency noise in DEM generation. In our investigations, dense image matching was evaluated for multiple stereo configurations as available from highly overlapping image flights.

The multi-stereo matching algorithm is briefly introduced in the section 2, whereas the quality of our dense DEM generation is presented in section 3. For this purpose, two different test areas are used. First, matching results for a planar sports field are investigated. For this area, datasets from a recent project on digital photogrammetric camera evaluation, initiated by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) are available (Cramer 2010). The data contain imagery from several large format aerial cameras, thus the potential of our dense matching for different camera systems and illumination conditions can be demonstrated. Furthermore, existing results from the DGPF project produced with commercial software tools and generated from the lidar data can be used as benchmark for our approach. The second test demonstrates the benefit of highly overlapping imagery for consistency checks and quality evaluation for built-up test regions of higher geometric complexity. For this purpose, a test flight with nominal forward and side laps of 80 % and 70 %, respectively, was evaluated.

2 High-resolution DEM from Multi-Stereo Matching

The determination of 3D point clouds at densities corresponding to the resolution of the available stereo images requires matching results for each image pixel. Since such a per-pixel measurement is highly ambiguous, additional constraints such as the assumption of a smooth surface are usually introduced. Algorithms that globally minimize matching costs between corresponding pixels and the respective smoothness constraints are called global image matching techniques. These algorithms provide good results in terms of quality and resolution, but usually suffer from a high complexity and large computational effort. However, this computational complexity can be reduced significantly by the semi-global matching (SGM) stereo method. This approach, used for our investigations, approximates a global approach by minimizing matching costs by aggregating them along a certain number of 1D paths in several directions through the image. In this way, the pixel-wise SGM approach provides a dense point distribution, while the approximation of global cost optimization by cost aggregation on 1D paths allows a reasonable runtime on large images.

2.1 *Semi-Global Matching*

Our implementation of the SGM algorithm is similar to Hirschmüller (2008) to a large extent. In all tests a combination of the census and the mutual information matching cost was used. Both costs have been proven to perform well in Hirschmüller & Scharstein (2009). As the census cost can be computed without any initial disparity estimation, it is utilized in our implementation for matching the low-resolution images during the implemented coarse-to-fine matching. The final matching step at the highest resolution level is realized by exploiting the mutual information matching cost.

Fig. 1 depicts an exemplary result of the matching process. Base and search image are normalized into epipolar geometry. The parallax image produced by SGM is shown on the right. It provides the disparity for each pixel in the base image (left) with respect to the search image. Within this disparity image, black areas are visible. They correspond to wrong parallaxes of the raw SGM results that were eliminated by a filter algorithm. This filter was realized by a simple consistency check with changed roles of base and match images during matching. Only disparity estimations consistent to this forward-backward matching are then considered as valid. Additional erroneous disparity estimations are removed by filtering speckles in the disparity images and a subsequent occlusion check (Hirschmüller 2008).

2.2 *Object Point Triangulation*

The matching process provides parallaxes, which link corresponding pixels between two images. After parallax estimation, object point coordinates are computed by spatial intersection using the corresponding pixel coordinates in the stereo pair as defined by the disparity images. In our implementation, the spatial intersection method described by Hartley & Zisserman (2004) is used, which allows for a computationally efficient solution.

There the transformations $\mathbf{x} = (x,y,1)^T = \mathbf{PX}$ and $\mathbf{x}' = (x',y',1)^T = \mathbf{P}'\mathbf{X}$ of object points $\mathbf{X}$ into the two image planes are given by the projection matrices $\mathbf{P}$ and $\mathbf{P}'$. These matrices represent the respective orientation parameters from bundle block adjustment. The projections are reformulated to $\mathbf{x} \times (\mathbf{PX}) = \mathbf{0}$ and $\mathbf{x}' \times (\mathbf{P}'\mathbf{X}) = \mathbf{0}$. In (1) and (2) the equations are formulated explicitly.

$$\begin{aligned} x\left(\mathbf{p}^{3T}\mathbf{X}\right) - \left(\mathbf{p}^{1T}\mathbf{X}\right) &= 0 \\ y\left(\mathbf{p}^{3T}\mathbf{X}\right) - \left(\mathbf{p}^{2T}\mathbf{X}\right) &= 0 \\ x\left(\mathbf{p}^{2T}\mathbf{X}\right) - y\left(\mathbf{p}^{1T}\mathbf{X}\right) &= 0 \end{aligned} \qquad (1)$$

Fig. 1: Base image of stereo pair with corresponding parallax image.

$$x'\left(\mathbf{p}'^{3T}\mathbf{X}\right)-\left(\mathbf{p}'^{1T}\mathbf{X}\right)=0$$
$$y'\left(\mathbf{p}'^{3T}\mathbf{X}\right)-\left(\mathbf{p}'^{2T}\mathbf{X}\right)=0 \quad (2)$$
$$x'\left(\mathbf{p}'^{2T}\mathbf{X}\right)-y'\left(\mathbf{p}'^{1T}\mathbf{X}\right)=0$$

$$\mathbf{A}=\begin{pmatrix} x\mathbf{p}^{3T}-\mathbf{p}^{1T} \\ y\mathbf{p}^{3T}-\mathbf{p}^{2T} \\ x'\mathbf{p}'^{3T}-\mathbf{p}'^{1T} \\ y'\mathbf{p}'^{3T}-\mathbf{p}'^{2T} \end{pmatrix} \quad (3)$$

In (1) and (2) the term $\mathbf{p}^{iT}$ denotes the i'th row in the 3×4 camera matrix. Using these equations, a linear system of equations of the form **AX = 0** can be generated with **A** given in (3). This system is solved using a single value decomposition (SVD) to compute the coordinates of object point **X**.

2.3 *Multi-Stereo Matching*

Traditional airborne photogrammetry used to limit the cost for film material and the effort for manual stereo measurement by minimizing the number of images to be collected and evaluated. This was realized by the typical 60 % forward and 20 % sideward overlap of images. In this way, the availability of one stereo image pair for each object point and a block geometry of suitable stability could be guaranteed. Meanwhile, extra costs for additional image acquisition can almost be neglected by using digital photogrammetric cameras. Thus, 80 % forward overlap can be easily generated. In contrast, the increase of available side lap requires additional flight lines, which will increase the costs since more flying time is required. However, flights with 60 % side lap are already realized for applications such as true orthophoto generation (Braun 2003). In this context, multiple overlapping images are required especially in dense urban areas to fill areas that are occluded in some of the images.

Overlaps of 80 % along track and 60 % across track will result in at least 10 images per object point. This is a considerable increase of redundancy compared to two image rays as available for image-based surface reconstruction from traditional photogrammetric blocks. These multiple image rays can be used to eliminate mismatches by suitable filter processes and increase the accuracy of image-based 3D point cloud generation. In our implementation, the link between pixels in multiple images representing the same object point is established by defining a central image as the base image, which is then matched with the surrounding images. Depending on the respective flight and camera parameters, different configurations are feasible. Fig. 2 gives some exemplary configurations, which presume imagery captured with 80 % forward and 70 % side lap. The blue rectangle with the bold outline in Fig. 2 represents the base image in the centre. This base image is then matched with the respective search images, which are also represented by blue rectangles. Fig. 2 (left) thus represents a configuration, where the base image is matched with its two nearest neighbours in flying direction. These two match images thus define two stereo image pairs with the base image. Fig. 2 (middle) shows a configuration where the base image is matched with all possible images of the same

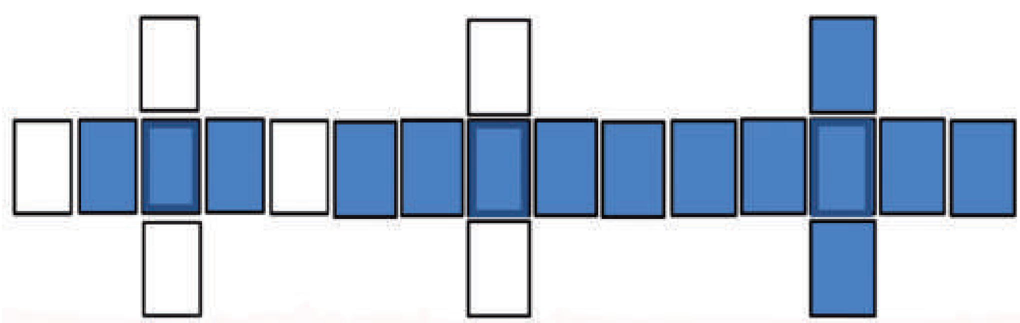


Fig. 2: Three potential multiple stereo configurations in a 80 % / 70 % overlapping image block.

strip. For 80 % forward lap, this results in four image pairs. Finally, Fig. 2 (right) includes two additional images from the upper and lower adjacent strips. Sufficient overlap with respect to the base image is available due to the 70 % side lap.

Matching n images to a common base image results in 2(n+1) pixel coordinates to determine the object point $\mathbf{X}^i$. The corresponding measurements consist of pixel coordinates $\mathbf{x}_b$ in the base image and the respective pixel coordinates $\mathbf{x}_{mi}$ in the multiple match images. Consequently, matrix **A** for calculating the object point $\mathbf{X}^i$ in (3) changes to

$$\mathbf{A} = \begin{pmatrix} x_b \mathbf{p}_b^{3T} - \mathbf{p}_b^{1T} \\ y_b \mathbf{p}_b^{3T} - \mathbf{p}_b^{2T} \\ x_{m_1} \mathbf{p}_{m_1}^{3T} - \mathbf{p}_{m_1}^{1T} \\ y_{m_1} \mathbf{p}_{m_1}^{3T} - \mathbf{p}_{m_1}^{2T} \\ x_{m_2} \mathbf{p}_{m_2}^{3T} - \mathbf{p}_{m_2}^{1T} \\ \dots \end{pmatrix} \tag{4}$$

Now, **A** contains two rows for the base image pixel $\mathbf{x}_b$ and two rows for each of the match images $\mathbf{x}_{mi}$. Each additional search image matched with the base image will add one image ray per point and thus improves the geometric quality of the object points. In this configuration n stereo pairs showing the same object point will provide 2(n+1) observations.

Object point triangulation is again realized by solving the system $\mathbf{AX} = \mathbf{0}$. In principle, this provides an intersection of image rays defined by the corresponding pixel coordinates and projection matrices. However, errors both from aerial triangulation and disparity measurement by SGM will result in small deviations of these image rays from the estimated object point. Thus, the least-squares solution of the linear system $\mathbf{AX} = \mathbf{0}$ can be used to measure the uncertainties of the point $\mathbf{X}^i$ in object space, with its accuracy derived from the diagonal elements σ_X^2, σ_Y^2 and σ_Z^2 of the covariance matrix $\sigma^2(\mathbf{A}^T\mathbf{A})^{-1}$. Despite of the fact that this method is comparably fast, one has to keep in mind that only an algebraic error is minimized. Therefore, the influence of ray geometries resulting from the different camera configurations might not be handled optimally.

Matches from stereo pairs with larger baselines provide a better intersection geometry and thus are advantageous for the determination of the point elevation. To take this into account more properly, one topic of our future work will be the integration of strategies minimizing the geometric error by using the Levenberg-Marquardt algorithm or an alternative linearization of the spatial intersection problem.

The disparity images still contain outliers. Based on the redundant determination of the object points the outliers are removed. The system $\mathbf{AX} = \mathbf{0}$ using (4) is solved for all available measurements $\mathbf{x}_m^i$. The resultant object point is re-projected into each match image according to $\hat{\mathbf{x}}_m^i = \mathbf{P}_m^i \mathbf{X}$ and afterwards, the residuals $r_m^i = \left\| \mathbf{x}_m^i - \hat{\mathbf{x}}_m^i \right\|_2$ are calculated. The measurement corresponding to the largest r_m^i is removed if this largest value r is greater than a stereo model-dependent threshold t^i. The object point coordinates are then determined from the remaining measurements. This procedure is repeated until all remaining residuals fullfil the threshold condition. If less than two detections $\mathbf{x}_m^i$ remain, the triangulation is aborted. The thresholds t^i are computed for each matched image individually by propagating the average expected object point accuracy to the respective image, considering the parameters of the stereo model that image forms with the base image. In a first step all object points $\mathbf{X}^i$ are computed from all corresponding matches without removing any outliers. Then the standard deviations σ_{Zi}^2 of all $\mathbf{X}^i$ are derived by the covariance matrices and averaged. Object points with $\sigma_{Zi}^2 > 3 \cdot \sigma_{Zmean}$ are discarded, and $\sigma_{Z,mean}$ of the remaining object points is recalculated and propagated into each match image. With focal length f, stereo base B and height over ground H the standard deviation σ_I in image space is

$$\sigma_I = \frac{f \cdot B}{H^2} \sigma_{Z,mean}. \tag{5}$$

The final thresholds are then derived by $t = 3\sigma_I$.

3 Performance of Semi-Global Matching

In order to evaluate 3D point cloud generation by dense matching, test data available from the DGPF project on digital photogrammetric camera evaluation (Cramer 2010) are used. These data have already been used to evaluate commercial matching software (Haala et al. 2010). Thus, a comparison with our implementation is feasible. The DGPF dataset includes lidar point clouds as a reference as well as imagery from several different airborne camera systems with nominal ground sample distances (GSD) of 20 cm and 8 cm. For reasons of simplicity our tests presented in section 3.1 were limited to 8 cm GSD imagery. The test area is defined by a planar soccer field, which eases the respective accuracy evaluations.

In section 3.2, multiple image matching formerly proposed in Haala (2011) is investigated for an area of higher geometric complexity. The benefits of multiray photogrammetry are demonstrated by a test flight with 80 % forward and 70 % side laps, respectively. For this purpose, aerial imagery captured by Vexcel's UltraCamXp at a GSD of approximately 10 cm is used.

3.1 *Investigations at a Planar Test Area*

The first test area, 'sports ground' depicted in Fig. 3, is a planar soccer field, which was also used for the investigation of commercial matching tools (Haala et al. 2010). This allows for the comparison of our results to those achieved by the commercial software tool MATCH-T DSM. Since the test area features rather low surface texture, it is also well-suited for evaluating the performance of matching approaches in potentially challenging areas. Thus, in a first test, the precision, accuracy and robustness of SGM regarding different sensors and varying flight-specific conditions, e.g. with respect to illumination, is evaluated.

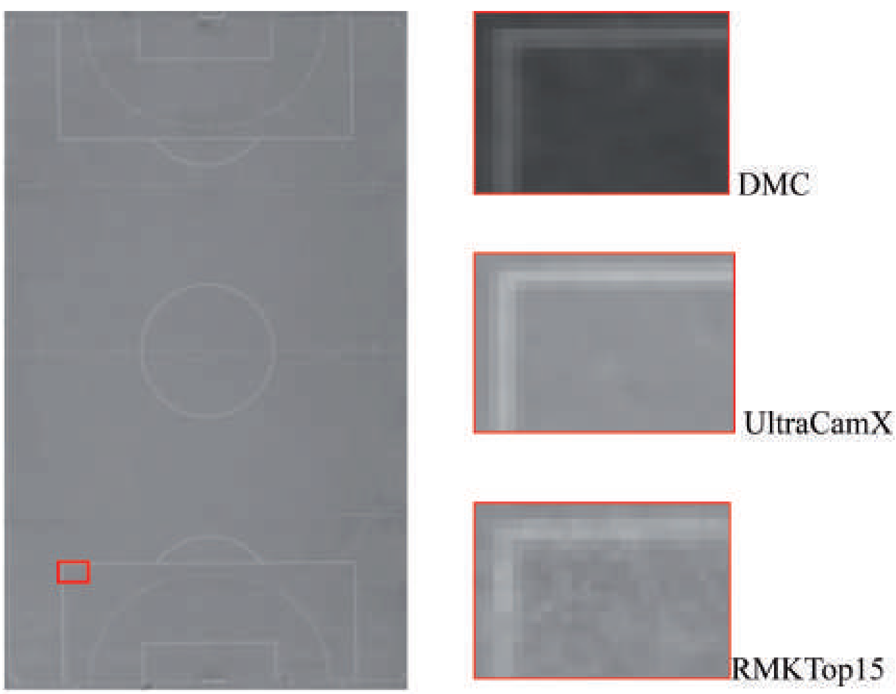


Fig. 3: Test region 'sports ground' with horizontal surface and homogenous texture.

3.1.1 Robustness and Noise Regarding Varying Sensors and Illumination Conditions

In order to examine the performance of SGM for different sensors and varying illumination conditions, the matching success rate and subpixel accuracies are analysed. Stereo models of the planar sports ground were generated using imagery of the three camera systems UltraCamX, DMC and RMKTop15. Obviously, the quality of 3D point clouds at planar areas like the soccer field can be influenced positively by specifying large smoothness constraints for SGM. This is avoided in our investigations by using the same standard set of penalty settings for all camera platforms.

Tab. 1 provides the results of SGM matching for DMC and UltraCamX imagery. In order to obtain comparable results, stereo image pairs of similar base-to-height ratios of 0.29 and 0.26, respectively, were used. In order to determine the accuracy of the generated 3D point clouds at the planar sports field, an approximating plane was estimated and the respective point-to-plane residuals were computed to provide the RMS errors σ_r. Mismatches were eliminated by removing points with residuals $>3\sigma_r$. After this filtering the standard deviation of object-point-to-plane residuals were recalculated to provide σ_r after filtering. These errors in object space were then propagated to image space using (5) to compute the accuracies σ_I in image space after filtering. For comparison, Tab. 1 additionally includes point cloud quality parameters of lidar measurements obtained from an ALS50 scanner. As can be seen in Tab. 1, stereo matching provides rather complete surface reconstructions at matching success rates larger than 99 % and point densities larger than 140 pts/m^2. The number of points is significantly larger than the one provided by the ALS50. As expected, the lidar data is more accurate than SGM results, which nevertheless provide accuracies better than 5 cm. As mentioned before, image-based precision measures were derived by error propagation; the image-based measures obtained for SGM are smaller than 0.14 pixels.

In the DGPF project, RMK imagery is only available at base-to-height ratios of 0.56. Therefore also image pairs of DMC and UltraCamX where matched providing base-to-height ratios of 0.57 and 0.51 respectively. As shown in Tab. 2 the matching success rates of RMK imagery amounts to 99.9 % which is

Tab. 1: SGM for single stereo pairs at test area 'sports ground' for 8 cm GSD imagery.

Sensor	DMC	UltraCamX	ALS
Base-to-height ratio	0.29	0.26	–
Matching success rate (%)	99.9	99.9	–
Point density (pts/m^2)	143.3	141.5	11.2
σ_r before filtering (cm)	4.20	3.57	1.52
σ_r after filtering (cm)	4.12	3.19	1.50
σ_I (pix)	0.14	0.10	–

Tab. 2: Results for single stereo pairs at test area 'sports ground' based on SGM and 8 cm GSD imagery, base-to-height ratios 0.51–0.57.

Sensor	DMC	UltraCamX	RMK
Base-to-height ratio	0.57	0.51	0.56
Matching success rate (%)	99.9	97.6	99.9
σ_r before filtering (cm)	2.35	2.05	2.86
σ_r after filtering (cm)	2.16	1.97	2.68
σ_I (pix)	0.15	0.12	0.19

in the same range as for the digital systems. Again, σ_r represents the standard deviation of object-point-to-plane residuals and σ_I accuracies in image space derived by error propagation of σ_r after removing outliers. The point densities of surfaces derived by UltraCam imagery are little lower than for DMC imagery. This is because of some low-textured areas in the UltraCam images where the matching fails. The standard deviations of point-to-plane residuals of the RMK data are slightly higher and a value of 2.9 cm is obtained, which corresponds to a matching accuracy of 0.19 pixels. It has to be mentioned, that the DMC and RMK images were recorded almost simultaneously at identical atmospheric and illumination conditions during the same flight. In contrast, the UltraCam flights were captured two months later. Thus, differences of the illumination that influenced the matching quality could not be avoided. This e.g. happened in areas of little texture such as the sports field. It must also be kept in mind that errors based on a comparison to approximating planes as presented in Tabs. 1 and 2 mainly represent the internal matching precision. A comparison to the lidar data shows that the differences between the respective planes are in the range of errors induced by the bundle block adjustment. For the datasets of the DGPF projects the vertical errors are in the order of 1/2 GSD (Jacobsen et al. 2010).

3.1.2 Comparison with Tests using Commercial Software Systems

For comparison, point clouds of the test area 'sports ground' were also computed using the commercial software MATCH-T DSM, which solves the correspondence problem using feature-based and least-squares matching techniques (Lemaire 2008). Similar to MATCH-T DSM, which uses all available images for object point generation, our SGM results are based on all available stereo pairs of one image strip. Depending on the available overlap

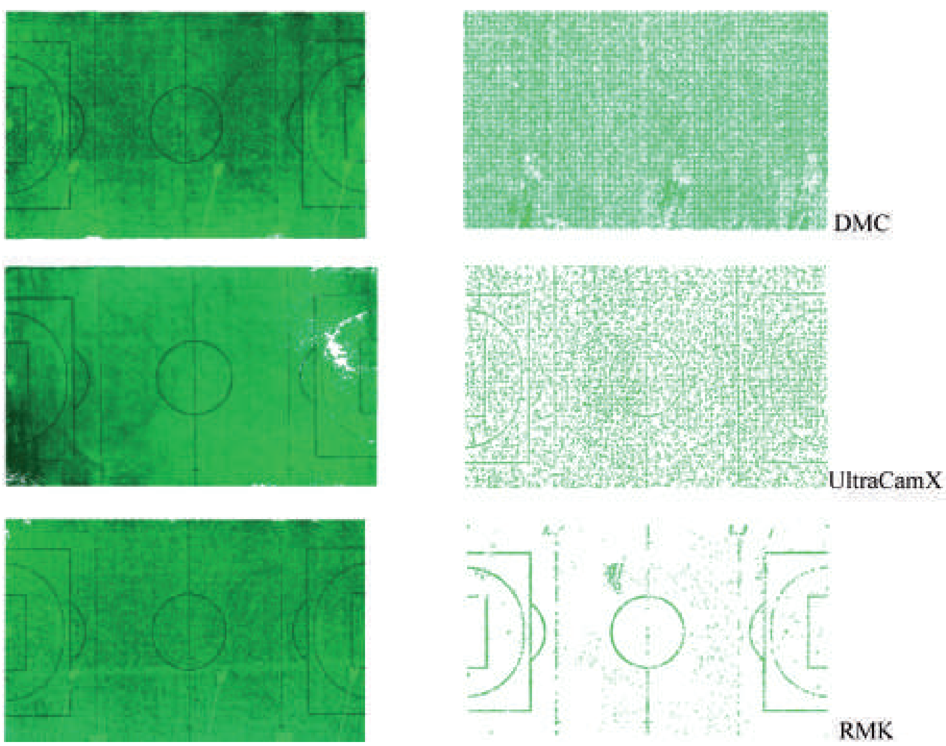


Fig. 4: Reconstructed point clouds for the test area 'sports ground' from SGM (left) and MATCH-T (right) using multiple stereo configurations.

for the respective datasets, this provided two image pairs for the DMC, five image pairs for the UltraCamX and one image pair for the RMK. SGM was again parameterized by the standard values. Object points were triangulated from all available measurements and accuracies were evaluated by comparing the point-to-plane residuals after removing the outliers.

Fig. 4 shows the resulting point clouds from SGM (left) and MATCH-T DSM (right). Whereas the matching success rate of MATCH-T significantly decreases for RMK imagery, SGM seems to be pretty robust against the reduced signal-to-noise ratio of the analogous camera system. It provides a rather complete reconstruction and a standard deviation of 3.0 cm for the respective point clouds. Despite of the homogenous texture our algorithm yields point densities of 135.4 pts/m² for the UltraCamX imagery. A standard deviation of 2.3 cm was achieved, which is slightly worse than the results of the stereo evaluation in section 3.1.1. This is caused by using stereo models with smaller base-to-height ratios for the object point triangulation. Again, some low textured areas could not be matched. In comparison, MATCH-T DSM provides point densities of 16.0 pts/m² with a standard deviation of 7.1 cm. The highest point densities for both methods were obtained for the DMC imagery. MATCH-T yields a reconstruction at a point density of 23.9 pts/m² with a standard deviation of 3.4 cm. Again, SGM provided a complete reconstruction with 141.6 pts/m² and an accuracy of 2.1 cm. One has to keep in mind that due to MATCH-T's black-box character, parameterization is limited, and therefore, the software might not perform at its optimum level. Furthermore, due to different weather and illumination conditions during acquisition of the DMC and UltraCamX imagery, this does not allow for a comparison of the different digital camera systems, but demonstrates the capability of the matching algorithms at areas of very homogenous texture.

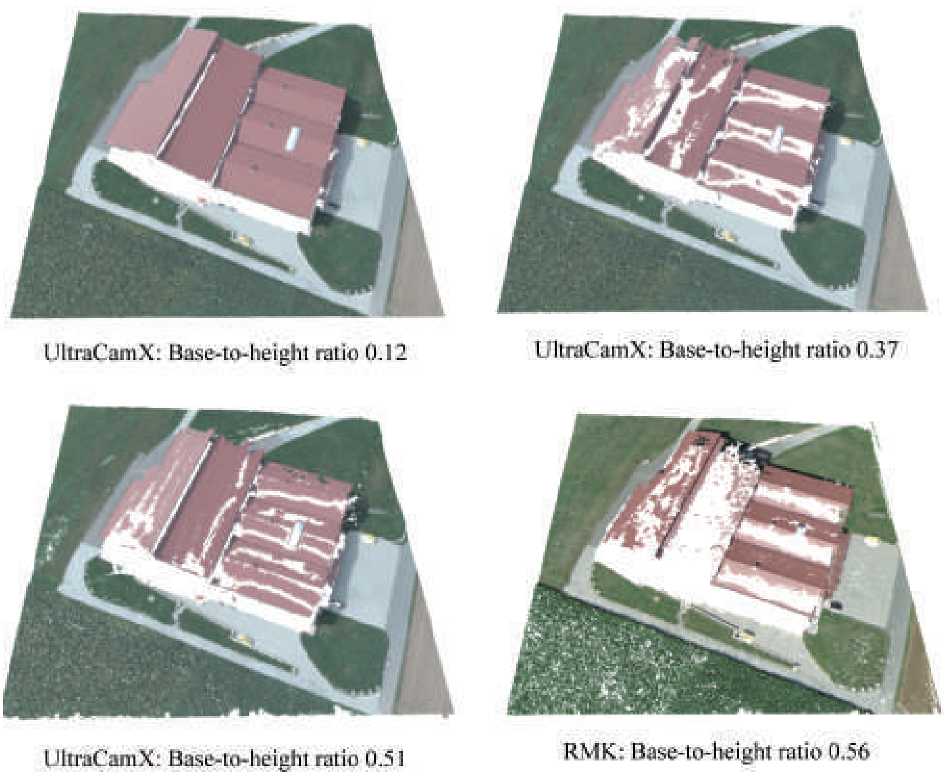


Fig. 5: Reconstruction of more complex surfaces, GSD 8 cm.

Besides rather homogenous intensities, which aggravate matching problems, the sports ground is a very special test scenario since its simple planar geometry can be modelled perfectly by the smoothness constraints. As a consequence, disparity estimations might be in large parts dominated by the penalty terms and not by distinctive cost minima. This might result in misleading values for sub-pixel accuracies. In order to demonstrate matching results for non-planar surface, Fig. 5 depicts a 3D point cloud at a building which was already examined in Haala (2009). The stereo models were generated using UltraCamX image pairs at base-to-height ratios of 0.125, 0.25 and 0.5, respectively. As can be seen in Fig. 5, image pairs with shorter baselines are better suited for stereo matching compared to large baseline imagery. Greater distances between the camera stations result in larger variation in the image content. As it can be seen in Fig. 5, this results in more mismatches and thus lower point densities, especially at complex geometries like buildings. Fig. 5 additionally shows a result from the RMK image pair at the base-to-height ratio of 0.56. Whereas tests on the sports ground resulted in comparable point densities for digital and analogous camera systems, in common 3D scenarios the impact of the worse signal-to-noise ratio is immediately noticeable.

3.2 *Influence of Varying Base-to-Height Ratios – Test at Complex Areas*

In order to further investigate the benefits of multiray photogrammetry, a test flight with 80 % forward and 70 % side lap was used. The aerial imagery was captured by Vexcel's UltraCamXp at a height above ground of 1600 m with a GSD of approximately 10 cm. Fig. 6 shows a test area selected from this block which is used in the following investigations. It includes several potentially problematic regions for image matching, marked by solid rectangles. These areas contain low texture (yellow), varying shadows due to illumination changes (purple), periodic patterns (red and green), vegetation of varying appearance in different images (orange), smaller structures as cars (blue) and a planar area (dashed rectangle). In contrast to the investigations in section 3.1, matching is not restricted to the planar region. Still, no assumption on underlying object geometry is required for error analysis and filtering. In contrast, point accuracies and the ratio of successfully matched points are determined using the redundancy from multiple stereo image pairs as already described section 2.3.

As shown in Fig. 7, two image pairs with different base-to-height ratios were combined

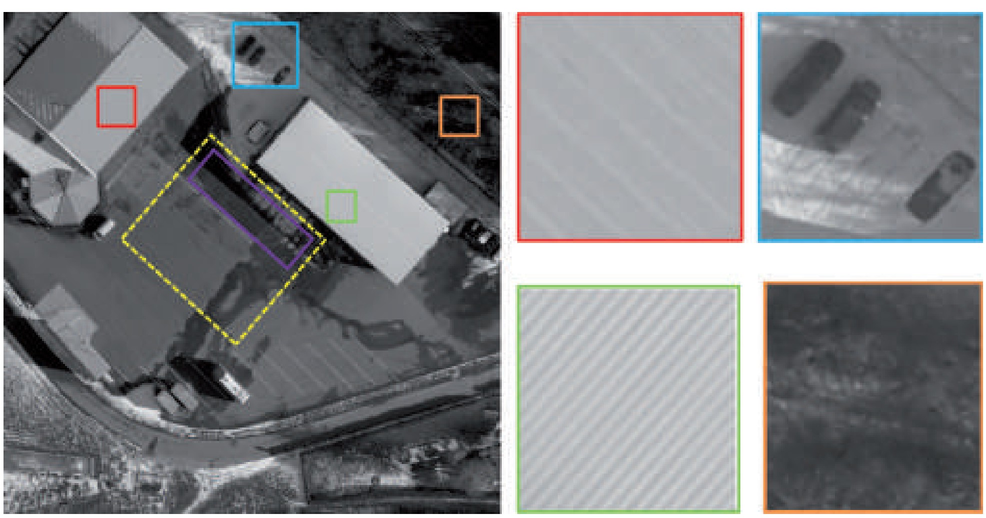


Fig. 6: Test area with rectangles defining problematic areas for image matching.

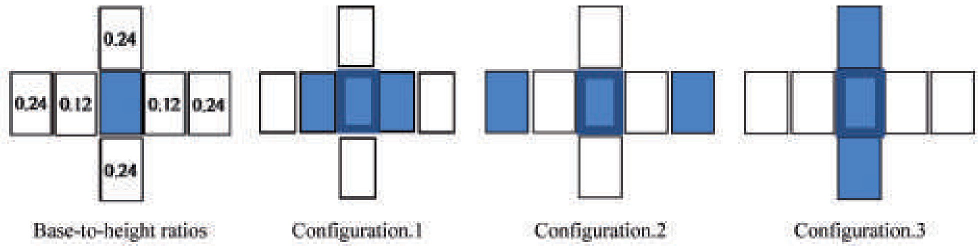


Fig. 7: Base-to-height ratios and configurations for multi-stereo matching.

in a first investigation. For stereo measurement, configuration 1 uses the direct neighbours in flight direction. The corresponding base-to-height ratio of 0.12 for these image pairs is also represented in the left image of Fig. 7. In configuration 2, the base-length of the respective stereo pairs increases by a factor of two. While configuration 1 and 2 are based on images of the same strip, configuration 3 uses search images from the upper and lower neighbouring strips. As represented in Fig. 7 (left), configurations 2 and 3 feature stereo pairs of the same base-to-height ratios of 0.24.

The resulting point clouds from multiple image matching using the configurations 1–3 are shown in Fig. 8. During processing, outliers were eliminated using the standard deviation of each object point $\mathbf{x}^i$. As discussed in section 2.3 the respective point accuracy can be computed in spatial intersection of the three image rays, which resulted from matching the base image with two match images, respectively. To eliminate erroneous matches the vertical standard deviation σ_{Z_i} obtained from the covariance matrix was used. For this purpose, a threshold is defined using the standard deviation $3\sigma_{\sigma Zall}$. This value is computed from the error σ_{Z_i} of all 3D point measurements in the test area as described in section 2.3. In this test, a valid 3D point thus requires two successful matches. These successful matches are represented by the colour coded point clouds in Fig. 8. Areas with no valid point measurements occur as black regions.

Similar to the example already given in Fig. 5, the number of successfully matched points decreases if the base-to-height ratio is increased. For the combination of the stereo pairs with base-to-height ratio of 0.12 (configuration 1), successful matches were generated for 81.4 % of all points. This corresponds to an average point density of 86 pts/m^2 for the available imagery. The percentage of successful matches decreases to 72.8 % for the base-to-height-ratio of 0.25 in configuration 2. Compared to percentages of successfully matched points in section 3.1, the respective values are lower in this experiment. However, here a point requires valid matches in two ste-

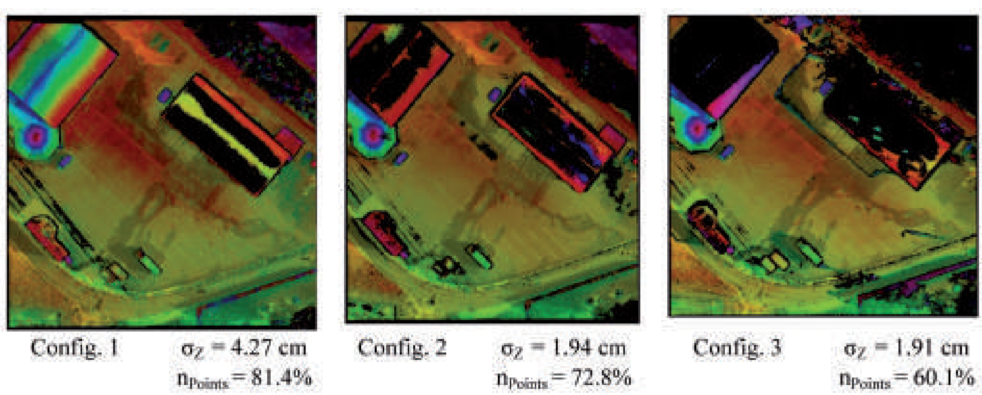


Fig. 8: Matching results for configurations 1–3.

reo pairs. Furthermore, measurements in section 3.1.1 were restricted to a planar region, whereas now complex 3D geometry and problematic areas for image matching are present. The distribution of no-point areas in Fig. 8 shows problems for all configurations at the repetitive pattern of the building roof marked by the green rectangle in Fig. 6. For the larger base-to-height ratios of configuration 2 and 3, additional problems occur especially at the bushes and trees marked by the orange rectangle in Fig. 6. Such objects are difficult for stereo matching because their appearance in the aerial images is especially sensitive to viewpoint changes. The number of matched points is even lower for configuration 3, which has the same base-to-height ratio as configuration 2, but uses images from different image strips. There, the larger differences between the times of exposure results in shadow movement, which again affects the matching process in a negative way. Fig. 8 gives an average vertical accuracy σ_Z = 4.3 cm for points from configuration 1. In contrast, σ_Z is better than 2 cm for configurations 2 and 3 due to the improved geometric configuration. It must be kept in mind that these values just represent the internal precision from the intersection of image rays. However, even if the additional error from bundle block adjustment is considered, the geometric quality of the generated point clouds is remarkable.

A further increase of both accuracy and density or completeness of 3D point clouds is feasible if multi-stereo matching is extended to even more image pairs. This is demonstrated in Fig. 9, which displays results of four and six image pairs. For the left example, which defines configuration 4, five images of one strip are used to form four stereo image pairs. The second example given by configuration 5 additionally includes two neighbouring images from the upper and lower image strips to generate two more stereo pairs. Compared to configuration 2 and 3 in Fig. 8 the vertical errors σ_Z slightly increase after the additional imagery is included. This holds true for both configurations 4 and 5. This is due to the fact that an increasing number of image rays for point determination gives a better evidence on remaining errors in image orientation from bundle block adjustment. However, the use of additional stereo pairs considerably increases the percentage of successfully measured 3D points. These values increase to 89.5 % and 91.5 % for configuration 4 and 5, respectively. Although the increase of successfully matched points between configurations 4 and 5 is rather low, the additional computation is still useful. The two additional stereo pairs in configuration 5 provide rays from images of different strips. Thus they enhance the point clouds mainly in areas that were previously occluded, such as facades. Furthermore, point determination at the building roofs, which is aggravated by the periodic patterns is improved.

4 Conclusions

In our investigations, SGM proved to be a robust and easy-to-parameterize matching algorithm. Best matching results were obtained for stereo images with short baselines, good

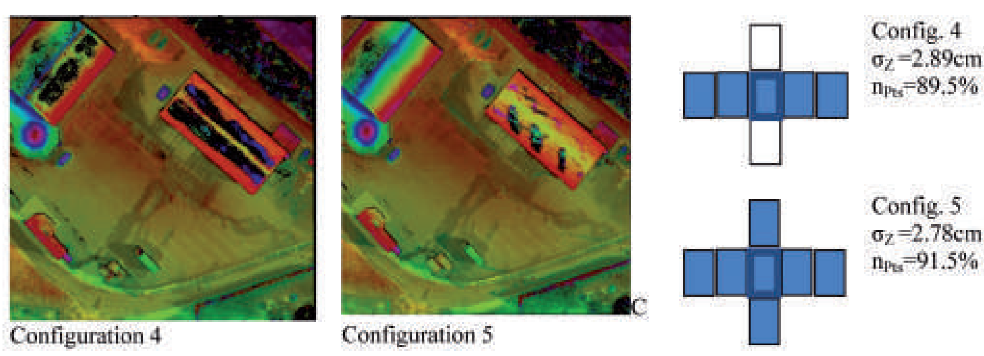


Fig. 9: Matching results for configuration 4 and 5 with for four and six image pairs.

texture and small signal-to-noise ratios. Obviously, matching performance depends on the image content, since homogenous intensities and large signal-to-noise ratios can limit the quality of the generated 3D point clouds. Still, matching accuracies better than 0.2 pixels at high point densities were feasible even for areas with very little texture. Potentially, the matching accuracy and reliability decreases for large base-to-height ratios due to changes in perspective and illumination. Nevertheless, the beneficial geometric properties for larger baselines during point determination by spatial intersection at least partially compensate the reduced matching accuracy. Mismatches, which occur more frequently for large baseline images, can be eliminated efficiently during the implemented multi-stereo matching. However, this outlier filtering reduces the density of the generated point clouds. For this reason, the combination of several stereo image pairs for multiple parallax estimation is especially beneficial. The combination of multiple measurements in triangulation increases the accuracy of the generated 3D point clouds, whereas their completeness can be increased by multi-stereo matching of imagery with varying perspective. Even more important, the redundancy available from the combination of stereo matches from different image pairs allows a very efficient accuracy analysis. This enables an efficient elimination of erroneous matches and results in a considerable reliability of the 3D points at vertical accuracies well at the sub-pixel level. Furthermore, multiple viewpoints also provide a good coverage in complex urban areas, which further increases the already large number of potential applications for multi-ray photogrammetry from dense image matching.

References

Braun, J., 2003: Aspects on True-Orthophoto Production. – Photogrammetric Week **'03**, Herbert Wichmann Verlag, Heidelberg.

Cramer, M., 2010: The DGPF-Test on Digital Airborne Camera Evaluation – Overview and Test Design. – PFG **2010** (2): 75–84.

Grün, A. & Baltsavias, E., 1988: Geometrically Constrained Multiphoto Matching. – Photogrammetric Engineering and Remote Sensing **54** (5): 633–641.

Haala, N., Hastedt, H., Wolff, K., Ressl, C. & Baltrusch, S., 2010: Digital Photogrammetric Camera Evaluation – Generation of Digital Elevation Models. – PFG **2010** (2): 99–115.

Haala, N., 2011: Multiray Photogrammetry and Dense Image Matching. – Photogrammetric Week **2011:** 185–195, Wichmann.

Hartley, R.I. & Zisserman, A., 2004: Multiple View Geometry in Computer Vision. – Cambridge University Press, Cambridge, Chap. **12**.

Hirschmüller, H., 2008: Stereo Processing by Semi-Global Matching and Mutual Information. – IEEE Transactions on Pattern Analysis and Machine Intelligence **30** (2): 328–341.

Hirschmüller, H. & Scharstein, D., 2009: Evaluation of Stereo Matching Costs on Images with Radiometric Differences. – IEEE Transactions on Pattern Analysis and Machine Intelligence **31** (9): 1582–1599.

Jacobsen, K., Cramer, M., Ladstädter, R., Ressl, C. & Spreckels, V., 2010: DGPF-Project: Evaluation of Digital photogrammetric Camera Systems – Geometric Performance. – PFG **2010** (2): 85–98.

Leberl, F., Irschara, A., Pock, T., Meixner, P., Gruber, M., Scholz, S. & Wiechert, A., 2010: Point Clouds: Lidar versus 3D Vision. – Photogrammetric Engineering and Remote Sensing, **76** (10): 1123–1134.

Haala, N., 2009: Comeback of Digital Image Matching. – Photogrammetric Week **2009:** 289–301, Wichmann.

Lemaire, C., 2008: Aspects of the DSM Production with High Resolution Images. – IAPRS **XXXVII** (B4): 1143–1146.

Seitz, S., Curless, B., Diebel, J., Scharstein, D. & Szeliski, R., 2006: A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms. – IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'**2006**).

Address of the Authors:

Apl. Prof. Dr. Norbert Haala, Dipl.-Ing. Mathias Rothermel, Institut für Photogrammetrie, Universität Stuttgart, Geschwister-Scholl-Str. 42D, D-70174 Stuttgart, Tel.: +49-711-685-83383, {norbert.haala} {mathias.rothermel}@ifp.uni-stuttgart.de

Manuskript eingereicht: Oktober 2011
Angenommen: März 2012

PFG 2012 / 4, 0345–0358
Stuttgart, August 2012

Article

Accurate Reconstruction of Near-Epipolar Line Segments from Stereo Aerial Images

ALI ÖZGÜN OK, Mersin, Turkey, JAN DIRK WEGNER, CHRISTIAN HEIPKE, FRANZ ROTTENSTEINER, UWE SÖRGEL, Hannover & VEDAT TOPRAK, Ankara, Turkey

Keywords: line matching, line reconstruction, near epipolar line segments, projective geometry, stereo aerial images

Summary: In this study, a new 3D reconstruction approach for line segments from stereo images is proposed which covers the case that the image lines are nearly-aligned (≤ 10º) with the epipolar line. The method manipulates the redundancy inherent in line pair-relations to generate artificial 3D points and utilizes those entities during the estimation process to improve the reconstruction of the line segments. The best points for the reconstruction are selected based on a newly proposed weight function. To test the performance of the proposed approach, we selected three test patches over a built-up area of the city of Vaihingen, Germany. The proposed approach produced a highly promising reconstruction performance for line segments that are nearly aligned with the epipolar line.

Zusammenfassung: In diesem Beitrag wird eine neue Methode zur 3D Rekonstruktion von Geraden aus Stereobildern vorgeschlagen, welche in den Bildern beinahe parallel (≤ 10º) zur Kernlinie liegen. Die Redundanz, welche sich aus paarweisen Relationen von Linien ergibt, wird ausgenützt, um künstliche 3D Punkte zu erzeugen, die dann in die Schätzung der Parameter der 3D Linien als Beobachtungen eingehen und die Ergebnisse der Rekonstruktion verbessern. Die am besten für die Rekonstruktion geeigneten Punkte werden auf Basis einer neu entwickelten Gewichtsfunktion ausgewählt. Um die Leistungsfähigkeit des Verfahrens zu testen, wurden drei Testgebiete in einem bebauten Gebiet der Stadt Vaihingen (Baden-Württemberg) ausgewählt. Das Verfahren lieferte viel versprechende Ergebnisse für die Rekonstruktion von 3D Linien, die in den Bildern beinahe parallel zu den Epipolarlinien liegen.

1 Introduction

The matching and reconstruction of line segments in ultra-high resolution (3–10 cm) stereo aerial images is a very challenging task due to various reasons, e.g. substantial change in viewpoints, inconsistency of line endpoint locations, lack of rich textures in local line neighbourhood, repetitive patterns, etc. Up to now, a significant number of research papers have been devoted to stereo line matching, e.g. SCHMID & ZISSERMAN (1997), BAILLARD & DISSARD (2000), SCHOLZE et al. (2000), ZHANG & BALTSAVIAS (2000), SUVEG & VOSSELMAN (2004), BAY et al. (2005), WANG et al. (2009), OK et al. (2010a, b). However, ultimately the problem has not been solved. In particular, in stereo geometry, the reconstruction of line segments that are nearly parallel to the epipolar lines is a challenging problem. The preferred way to alleviate this problem is to acquire more images, in particular with at least 60 % side lap. This type of overlapping image acquisition ensures the availability of at least one image in the cross-strip direction; thus, the problem of nearly collinear image perspective centers is mitigated, which comes along with better intersection angles of the projection planes. On the other hand, utilizing additional images in the cross-strip direction leads to new problems. Since those images are acquired in adjacent strips, there is a time delay between those images, which causes a difference in the illumination; thus, the moving

DOI: 10.1127/1432-8364/2012/0122 1432-8364/12/0122 $ 3.50

shadow phenomenon occurs. As a result, the overall geometric quality of the final reconstruction is negatively affected (Haala 2009). Moreover, from an economical point of view, any increase of side-overlap during image acquisition results in longer flight times and increases costs. Therefore, from an operational perspective, the number of strips must be minimized and common side overlaps of about 30 % are still preferred and frequently used in many countries. For instance, the surveying authorities of the German Federal State of Lower Saxony still regularly acquire aerial images with a side lap of 30 % (LGLN 2012). For those reasons, accurate reconstruction of line segments from aerial images acquired in a stereo configuration has a major importance and in this paper, our focus is on this field of research.

Reconstructed 3D line segments can be used for applications such as the reconstruction of buildings (Scholze et al. 2002) and roads (Zhang 2004). Theoretically, object lines corresponding to slope changes or height discontinuities could also be detected in a dense DSM as it can be generated by modern dense matching techniques (Hirschmüller 2008). However, this is not the case for lines that are just defined by their radiometric differences from their surroundings; such lines are very useful for road extraction (Zhang 2004). Furthermore, dense matchers have problems with object planes that have no texture and that are not fronto-parallel, in particular if wide baselines are involved, because the smoothness term used in such methods will cause neighbouring pixels having similar grey values to have the same depth (Bulatov 2011). These problems and the fact that 3D lines provide a sparse representation of the information most relevant for the reconstruction of man-made objects motivate our interest in 3D line reconstruction.

The general strategy to solve line matching problems in stereo vision has relied on various descriptors specialized for one-to-one line matching. However, the integration of the line-to-line relations in the matching process does not only expose new constraints to improve the matching performance (Ok et al. 2010a, b), but also provides new opportunities for the reconstruction stage. In Ok et al. (2010a), we presented an approach for pair-wise matching of line segments from stereo aerial images, an approach that was extended to deal with repetitive linear patterns in Ok et al. (2010b). In this paper, we want to focus on the reconstruction of the matched segments. Although the method of direct construction (Heuel & Förstner 2001, Heuel 2001) gives satisfactory reconstruction results for lines that are not aligned with the epipolar line, a dramatic decrease in terms of accuracy for lines that are nearly or exactly aligned ($\leq 10°$) with the epipolar line is inevitable (Fig. 1). This is due to the fact that if the angles of lines in image space get closer to the epipolar direction, the two projection planes generated from line segments become similar and in the worst case (exact alignment) become identical. For those cases, the direct construction of 3D lines from the intersection of planes may fail or at least produce inaccurate results. For that reason, Zhang (2005) proposed an alternative way which relies on free-form line structures. The major difficulty with the free-form structures is that the problematic line(s) must have an edge connection with the other neighbouring lines. However, this is rarely the case for most of the line segments. Therefore, in this paper, we propose a new reconstruction method that relies on the line-to-line relations developed in the pair-wise matching approach. The main idea is to manipulate the redundancy inherent in pair-relations to generate artificial 3D points from available pair matches and utilize those points during the estimation process to perform accurate reconstruction of the matched segments. However, since we do not know whether the two lines in a pair really intersect in object space or not, we select the point entities best suited for the estimation process by means of a new weight function which is composed of three main terms that are computed in a pair-wise manner: Euclidean distance, epipolar constraint and intersection angle in image space. For each problematic matching case ($\leq 10°$), we automatically select the appropriate artificial 3D points and use the selected points in the estimation process along with the projection planes of the line segments. Thus, we can also reconstruct these problematic line segments with promising final accuracies.

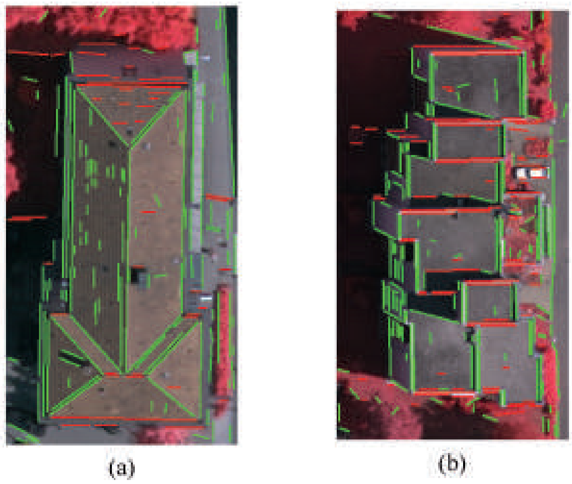


Fig. 1: The line segments extracted from two test patches. The red segments indicate the lines that are found to be nearly aligned (≤ 10°) with the epipolar line (flight direction).

The remainder of this paper is organized as follows. We shortly summarize the used stereo line matching approach in section 2. In section 3, the details of the proposed reconstruction approach are described. We present the test images and the results in section 4. Finally, in section 5 we give concluding remarks and make suggestions for future work.

2 Matching of Line Segments

For the matching of line features, we recently proposed a new relational approach in which the line correspondences between the aerial stereo images are established using pair-wise relations. In this paper, we only briefly summarize the matching algorithm and refer the reader to the references for further details (Ok et al. 2010a, b). The algorithm consists of three fundamental steps: (i) 2D line extraction, (ii) stereo matching of the extracted lines with a pair-wise relational approach, and (iii) iterative pair-based post-processing.

In the first step, the multispectral information existing in aerial images is fully utilized in order to maximize the performance of line detection. To accurately describe the straight edge segments, a principal component analysis technique was adapted and the parameters of the straight line segments are determined using orthogonal regression. In addition, during the raster-to-vector conversion, we allow curved structures to be approximated by piecewise linear representations. In the second step, a new pair-wise stereo line matching approach is used to establish the pair-wise line correspondences between the stereo images. The approach initially generates reference line pairs in the base image and collects all potential matching candidate pairs from the search image with the aid of a-priori known image-to-image geometry. Next, the number of matching candidate pairs of each line pair in the base image is significantly reduced after imposing a weighted pair-wise matching similarity score computed over a total of eight pair-wise constraints (an epipolar, three geometric, two photometric, a correlation and a spatiogram constraint). In the last step, an iterative pair-based post-processing algorithm is utilized. For each line in the base image, the best corresponding line in the search image is found in an iterative disambiguation process in which the matching inconsistencies are further eliminated using nearest/next distance ratios and a final similarity voting scheme (Ok et al. 2010a, b).

3 Reconstruction of Line Segments

The reconstruction process starts with a test which determines the angle difference (0°–90º) between the line segments and the related epipolar line. Based on our experiences, the line segments that have angle differences of less than 10º are highly susceptible to produce inaccurate reconstruction results. Fig. 2 clarifies this fact. In the figure, the RMS distances of direct reconstruction results from lidar data are plotted as a function of the angle differences for all line segments correctly matched in test patch 1 used in our experiments (see the details of the comparison method in section 4). As can be seen from Fig. 2a, the best reconstruction performance is achieved for the line segments that have angle differences of 60º or more (≈ 1 pixel). However, the line segments that have angle differences of 10º or less provided extremely large overall RMS distances. Therefore, during the test, we label the line segments as not-aligned if the angle difference is larger than 10º. The reconstruction of all those matched line segments is performed by intersecting the projection planes (Fig. 3a), $\mathbf{A}^1(l_1)$ and $\mathbf{A}^2(l_2)$. The intersection is carried out with the method of direct construction (Heuel & Förstner 2001, Heuel 2001), which is an algebraic approach used for the construction of the 3D lines directly from given plane entities.

The underlying approach for the reconstruction of line segments that are nearly-aligned (≤ 10º) with the epipolar line is to manipulate the redundancy inherent in pair-relations (Ok et al. 2010b) to generate artificial 3D points ($\mathbf{X}_i$) and utilize those points in the estimation process. In this way, the neighbouring line segments that have line pair-connections with the problematic segment contribute to the reconstruction (Fig. 3b). As we shall see in the subsequent section, all generated points lie exactly on the problematic line segments in image space and, thus, also belong to the projection planes. Therefore, each artificial 3D point $\mathbf{X}_i$ must be generated beforehand, and the final estimation will consider these points jointly with the projection planes ($\mathbf{A}_i$). In this section, we follow the representation of the homogeneous uncertain vectors given in (Heuel & Förstner 2001).

3.1 *The Generation of Uncertain Artificial 3D Point Entities*

Formally the uncertain homogeneous vectors are denoted as $(\mathbf{x}, \boldsymbol{\Sigma}_{xx})$, where $\boldsymbol{\Sigma}_{xx}$ is the covariance matrix of the homogeneous vector $\mathbf{x} = [u_x, v_x, w_x]^T$. A 2D uncertain line $(\mathbf{l}, \boldsymbol{\Sigma}_{ll})$ in image space can be generated by joining the endpoints of the line segment, $(\mathbf{x}, \boldsymbol{\Sigma}_{xx})$ and $(\mathbf{y}, \boldsymbol{\Sigma}_{yy})$:

$$(\mathbf{l}, \boldsymbol{\Sigma}_{ll}) = \left(\mathbf{x} \times \mathbf{y},\, S(\mathbf{y})\boldsymbol{\Sigma}_{xx}S(\mathbf{y})^T + S(\mathbf{x})\boldsymbol{\Sigma}_{yy}S(\mathbf{x})^T\right)$$

$$S(\mathbf{x}) = \begin{bmatrix} 0 & -w_X & v_X \\ w_X & 0 & -u_X \\ -v_X & u_X & 0 \end{bmatrix} \qquad (1)$$

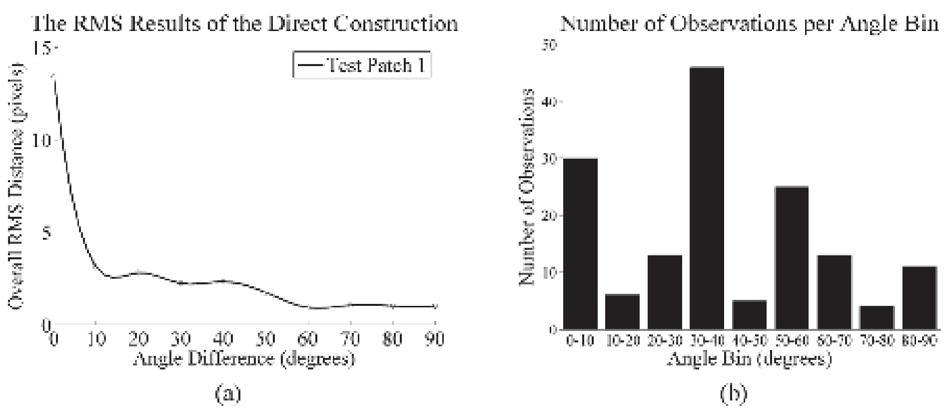


Fig. 2: (a) The performance of the direct construction method for various angle differences (0°–90°), and (b) the number of observations of line versus plane comparison used to generate RMS results in Fig. 2a.

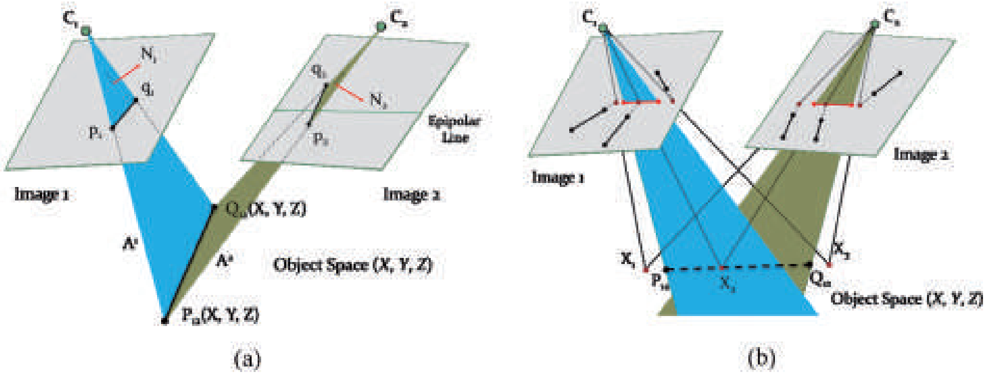


Fig. 3: The reconstruction of the line segments (a) that are not aligned with the epipolar line, and (b) that are nearly-aligned with the epipolar line. In (b), a dashed line style is used to illustrate that the line segment in object space is between the two projection planes.

where the skew-symmetric matrices S(**x**) and S(**y**) are the matrix representations of the vector cross product with the points **x** and **y**, respectively. Similar to (1), the intersection point **x** of two lines **l** and **m** can be computed as

$$(\mathbf{x}, \boldsymbol{\Sigma}_{xx}) = \left(\mathbf{l} \times \mathbf{m},\, S(\mathbf{l})\boldsymbol{\Sigma}_{mm}S(\mathbf{l})^{T} + S(\mathbf{m})\boldsymbol{\Sigma}_{ll}S(\mathbf{m})^{T}\right) \quad (2)$$

with the related skew-symmetric matrices, S(**l**) and S(**m**). In our case, initially, the covariance matrices of line endpoints are computed from edges that form the line segment in the way described in (Madsen & Christensen 1995). Next, the lines and their uncertainties are computed from those endpoints using (1) (Fig. 4a). Thereafter, for each problematic segment, the neighbouring line segments that have a pair-connection with the problematic segment are collected from the available pairwise matches, and their intersection points are computed using (2). Fig. 4b illustrates an example of a generated artificial intersection point with its elliptical confidence region.

The estimation of the artificial 3D points $\mathbf{X}_i$ from 2D correspondences cannot be performed using direct construction, because the

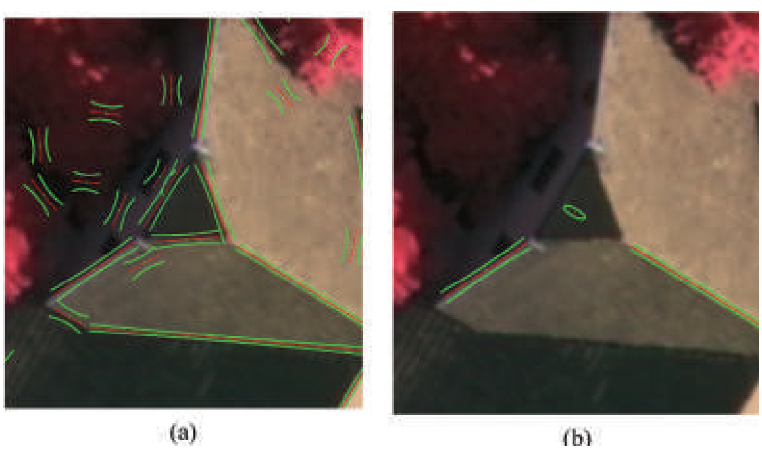


Fig. 4: (a) Confidence regions of extracted line segments, (b) an example of the generated artificial intersection point and its elliptical confidence region. Entities and confidence regions are shown in red and green colours, respectively. Confidence regions are 15 times exaggerated for visualization.

projecting rays emerging from 2D points rarely intersect in object space. Therefore, we estimate the 3D points using the iterative linear Gauss-Helmert model with constraints (McGlone et al. 2004, Förstner 2005). We have six observations per point (two homogeneous 3D vectors $\mathbf{x}_{ij}$ per image point, where j is the image index) and four unknowns, i.e. the homogeneous coordinates of the 3D point. Thus, referring to the Gauss-Helmert model, we have the vector γ of the observations and the vector β of the unknown parameters:

$$\gamma_{6\times1} = \begin{bmatrix} \mathbf{x}_{i1} \\ \mathbf{x}_{i2} \end{bmatrix} \qquad \beta_{4\times1} = \mathbf{X}_i \tag{3}$$

The information that a point $\mathbf{X}_i$ has to be on the image ray emanating from point $\mathbf{x}_{ij}$ in image j can be used to formulate three constraint equations $S(\mathbf{x}_{ij})\, P_j\, \mathbf{X}_i = 0$, where P_j is 3x4 projection matrix for points on image j. Of these three equations, only two are independent. In this study, we describe the independent constraints of $S(\mathbf{x})_{3x3}$ with reduced skew-symmetric matrices $S^r(\mathbf{x})_{2x3}$. Two independent constraints can be selected by choosing those two which belong to the rows of $S(\mathbf{x})$ containing the absolute largest element of the vector $\mathbf{x}$, at the same time being the largest element in $S(\mathbf{x})$ (McGlone et al. 2004). Thus, there are two independent constraints $\boldsymbol{g}$ for each observed image point $\mathbf{x}_{ij}$ and the unknown parameters ($\mathbf{X}_i$). In addition, we have a single length constraint h on the unknown parameters due to the homogeneity of $\mathbf{X}_i$. Thus, the relations of the Gauss-Helmert model for the estimation of the 3D point entities can be written as:

$$\boldsymbol{g}(\boldsymbol{\beta},\boldsymbol{\gamma})_{4\times1} = \begin{bmatrix} S^r(\mathbf{x}_{i1})P_1\mathbf{X}_i \\ S^r(\mathbf{x}_{i2})P_2\mathbf{X}_i \end{bmatrix} = 0 \tag{4}$$

$$h(\boldsymbol{\beta}) = \mathbf{X}_i^{\mathrm{T}}\mathbf{X}_i - 1 = 0$$

The initial approximate values of $\mathbf{X}_i$ for the iterative solution can be obtained from the singular value decomposition (SVD) solution (Heuel & Förstner 2001) which is a common method used to solve a set of homogeneous linear equations. Once the estimation is completed, the covariance matrices of the estimated 3D point entities can be computed from the inverted normal equation matrix.

3.2 *The Joint Estimation of the 3D Line Segments*

For the estimation procedure, we parameterize the 3D lines in Plücker representation $\boldsymbol{L}^T = (\boldsymbol{L}_h^T, \boldsymbol{L}_0^T) = (L_1, L_2, L_3, L_4, L_5, L_6)$ (McGlone et al. 2004) and utilize an iterative linear Gauss-Markoff model with constraints. The algebraic expressions of the form $\mathbf{g}_i\ (\beta;\ \gamma_i) = 0$ with respect to all possible observation (γ_i) and unknown (β) entities are developed and explicitly given in Förstner et al. (2000) and Heuel & Förstner (2001). In our case, we are searching for an unknown line $\mathbf{M}$ which primarily must lie in two planes; thus,

$$\mathbf{g}_1(\mathbf{M};\mathbf{A}_i)_{4x1} = \Pi^{\mathrm{T}}(\mathbf{A}_i)\mathbf{M} = 0 \tag{5}$$

where i is the plane index and Π^{T} is the homogeneous matrix representation:

$$\Pi^{\mathrm{T}}(\mathbf{A}_i) = \begin{bmatrix} A_{i0}\mathbf{I}_n & S^{\mathrm{T}}(\mathbf{A}_{ih}) \\ -\mathbf{A}_{ih}^{\mathrm{T}} & \mathbf{0}^{\mathrm{T}} \end{bmatrix} \tag{6}$$

In (6), $\mathbf{A}_{ih}$ and $\mathbf{A}_{i0}$ correspond to homogeneous and Euclidean components of the plane $\mathbf{A}_i$, and $\mathbf{I}_n$ is a 3 x 3 identity matrix. The projection planes $\mathbf{A}_i$ for each line $\mathbf{l}_i$ and the related uncertainties can be determined using the projection matrices P_j of the images:

$$\begin{aligned} \mathbf{A}_i &= P_j^{\mathrm{T}}\mathbf{l}_i \\ \Sigma_{A_iA_i} &= P_j\Sigma_{l_il_i}P_j^T + (I_4 \otimes \mathbf{l}_i)\Sigma_{P_jP_j}(I_4 \otimes \mathbf{l}_i^{\mathrm{T}}) \end{aligned} \tag{7}$$

where $\mathbf{I}_n$ represents n x n unit matrix and $\otimes$ denotes the Kronecker product. The uncertainty of each line is derived from (1) and for this study; we assume that the projection matrices are free of error. In addition to the projection planes, the unknown line should also contain the artificial 3D points $\mathbf{X}_i$ generated from the neighbouring line segments:

$$\mathbf{g}_2(\mathbf{M};\mathbf{X}_i) = \overline{\Pi}^{\mathrm{T}}(\mathbf{X}_i)\mathbf{M} = 0 \tag{8}$$

where i is the point index and $\overline{\Pi}^{\mathrm{T}}$ is the homogeneous matrix representation:

$$\overline{\Pi}^{\mathrm{T}}(\mathbf{X}_i) = \begin{bmatrix} -S(\mathbf{X}_{i0}) & X_{ih}\mathbf{I}_n \\ -\mathbf{X}_{i0}^{\mathrm{T}} & \mathbf{0}^{\mathrm{T}} \end{bmatrix} \tag{9}$$

In (9), X_{ih} and $\mathbf{X}_{i0}$ correspond to homogeneous and Euclidean components of the points $\mathbf{X}_i$ and $\mathbf{I}_n$ is a 3 x 3 identity matrix. However, since we do not exactly know whether the two lines in a pair really intersect in object space or not, before the estimation process we compute a weight for each artificial 3D point. The weights depend on three measures computed in a pair-wise manner, namely the minimum 2D Euclidean distance (d_{ij}) between two line segments (l_i and l_j, where i and j represents the line indices), the minimum angle (θ_{ij}) enclosed by line segments l_i and l_j, and the minimum orthogonal distance (d_{ij}^e) between the intersection points and related epipolar lines (l_{epi}) (Fig. 5a). The minimum 2D Euclidean distance (d_{ij}) between the two line segments is computed in a distinctive way in which the edge segments that form the line segments are taken into account (Fig. 5b). To do that, first, edge segments are labelled by 8-neighbourhood pixel connectivity, and after that, all edge-to-edge distances between line segments are evaluated. Finally, the distance which provides the minimum distance between the two edge segments is selected as the minimum 2D Euclidean distance. For example, in Fig. 5b, the distance between the line segments l_1 and l_2 is computed as zero since they belong to the same edge segment. In this way, we fully take into account the information of edge connectivity (or edge proximity) for the line segments during the computation of the minimum 2D distances in image space.

In principle, the reliability of a point increases if the distances (both d_{ij} and d_{ij}^e) computed are relatively short and decreases if the intersection angle is quite narrow (e.g., < 10°). Therefore, we designed the new cumulative weight function (W_i) as:

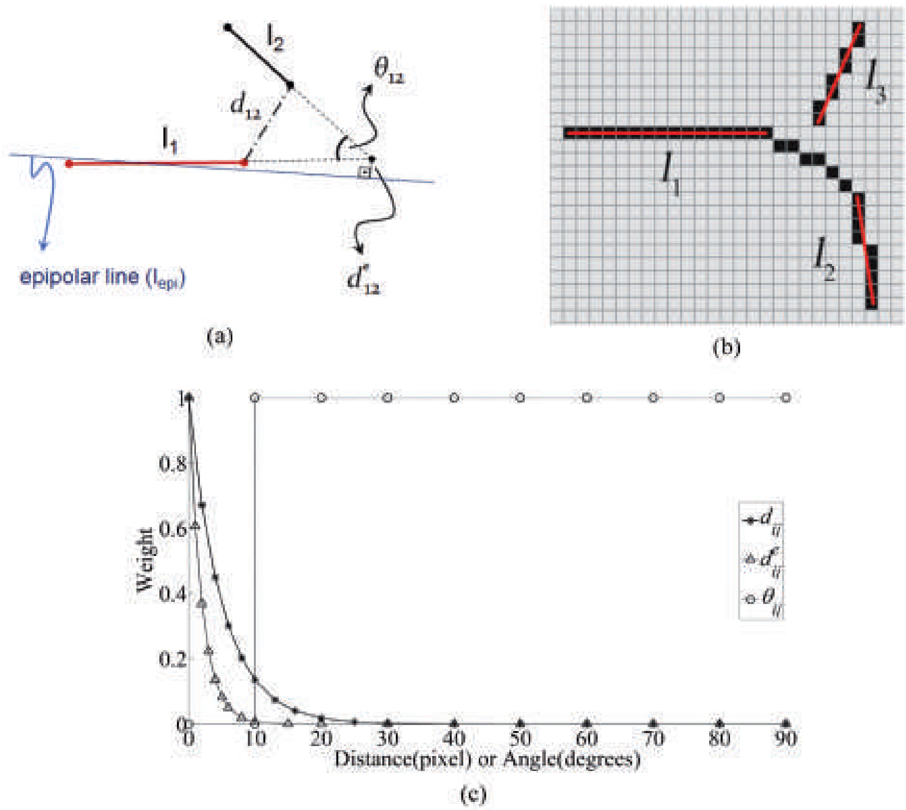


Fig. 5: (a) Three measures used in the weight function, (b) the computation of minimum 2D distances in image space, and (c) independent weight curves for each measure. The control parameters are utilized as $\sigma_1 = 5$ and $\sigma_2 = 2$.

$$W_i = t_{ij} . e^{-\left(\frac{\sigma_2 d_{ij} + \sigma_1 d_{ij}^e}{2\sigma_1\sigma_2}\right)}$$

$$t_{ij} = \begin{cases} 0 & \text{if } \theta_{ij} \leq 10^\circ \\ 1 & \text{if } \theta_{ij} > 10^\circ \end{cases} \quad (10)$$

where the parameters σ_1 and σ_2 control the weighting for the metrics d_{ij} and d_{ij}^e, respectively. Empirically, we found that the points which have distances of d_{ij} > 30 pixels or d_{ij}^e > 10 pixels are not reliable for the estimation process. Therefore, the control parameters σ_1 and σ_2 in (10) are selected as 5 and 2, respectively. Fig. 5c shows the weighting curves computed separately for each measure. As can be seen in Fig. 5c, the exponential function used for the distance d_{ij}^e rapidly decreases with an increase of the computed 2D Euclidean distances and reduces to a weight value of almost zero at a d_{ij}^e distance value of 10 pixels. Besides, the exponential weight function used for the distance d_{ij} reduces to almost zero weight value when the distance d_{ij} is 30 pixels. On the other hand, for the angle parameter θ_{ij}, a piecewise constant function with a single interval at 10º guarantees that the artificial 3D points generated from two neighbouring line segments are computed with intersection angles of larger than 10º.

For each nearly-aligned matching case (≤ 10º), the artificial 3D points $\mathbf{X}_i$ are collected and their weights W_i are determined using (10). However, it is not logical to integrate all observed point entities directly to the estimation process, since some of those points may be generated from wrong matches. Therefore, first, we eliminate all point entities that have weights less than a pre-defined threshold ($T_w \leq 0.05$). Thereafter, among the remaining points, only the points that have the highest weights are integrated to the estimation process along with the observed projection planes. However, the selection approach for the best point entities is not trivial, since the spatial distribution of the 3D points eventual-

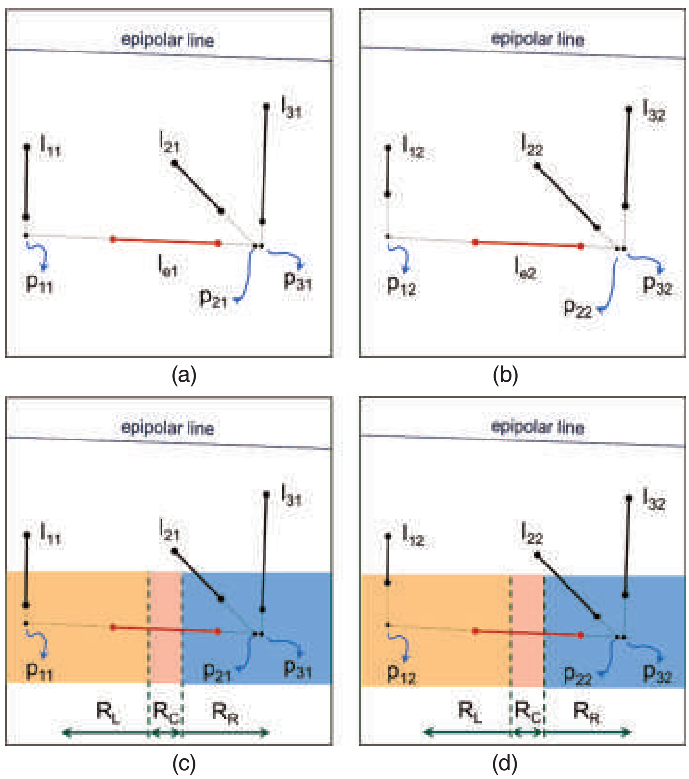


Fig. 6: Left (a, c) and right (b, d) stereo images. The directional regions utilized during the selection of the best weighted point entities.

ly determines the quality of the estimated 3D line. Fig. 6 clarifies this fact. In the figure, the line segments l_{e1} and l_{e2} are nearly aligned to the epipolar lines. After pair-wise matching, assume that l_{e1} and l_{e2} are correctly matched and have pair-relations with three neighbouring line segments (l_1, l_2, and l_3). We estimate the related 3D points ($\mathbf{X}_1$, $\mathbf{X}_2$, and $\mathbf{X}_3$) from the corresponding intersection points (p_{ij}) using (1)–(4) and compute the related weights (W_1, W_2 and W_3) for each point from (10). It is clear that the points $\mathbf{X}_2$ and $\mathbf{X}_3$ will get higher weights than $\mathbf{X}_1$, because, assuming no edge connectivity, the computed minimum 2D Euclidean distance (d_{ij}) measure for $\mathbf{X}_1$ is significantly larger than the other distances. Thus, in this case, the 3D line would be estimated from $\mathbf{X}_2$ and $\mathbf{X}_3$ along with the projection planes $\mathbf{A}_{e1}$ and $\mathbf{A}_{e2}$. However, since those entities are very close to each other and are located just on one side (to the right) of the lines l_{e1} and l_{e2}, the final position of the line segments in object space is highly sensitive to the small deviations between the entities $\mathbf{X}_2$ and $\mathbf{X}_3$, which may significantly reduce the final accuracy of the estimated 3D line. To solve this problem, we propose a region-based weighting of points (Fig. 6 c, d). We utilize two split points ($l/3$, $2l/3$) to divide the line into three directional regions, left (R_L) – centre (R_C) – right (R_R). Next, for each region independently, the point having the highest weight is searched and utilized in the estimation process. Thus, we guarantee that the point $\mathbf{X}_1$ generated by the intersection points p_{l1} and p_{l2} contributes to the final estimation.

For the estimation process, we form the point-line and plane-line incidence relations in 3D space and perform an iterative linear Gauss-Markoff model with constraints. For each case, we have 6 unknowns (Plücker coordinates) and 2 constraints (Plücker constraint and length) for each 3D line ($\mathbf{L}_i$). In addition to the two projection planes, the final observation number depends on the number of point entities $\mathbf{X}_i$ available for each region. Once again, the required initial values can be taken from the SVD solution, and the covariance matrices of the estimated 3D line entities can be computed from the inverted normal equation matrix of the Gauss-Markoff model.

4 Test Sites and Results

Three test patches (Fig. 7) are chosen from the DGPF test site in Vaihingen, Germany (Cramer 2010). For those test patches, stereo images were acquired by the DMC digital camera with 800 m flying height, which corresponds to a ground sample distance (GSD) of approximately 8 cm. The images were acquired with 70 % forward overlaps with a base-to-height (B/H) ratio of 0.28.

During the experiments, for all test patches, we applied a 50 m (≈ 162 pixels) search range difference (between the min. and max. heights) along the epipolar lines. In order to assess the performance of the line matching, the line matches were classified into three categories: True Positive (TP), False Positive (FP), and False Negative (FN), by comparing the automatically matched line segments with the manually generated reference line matching list. We term a line match as a True Positive if that match correctly corresponds to a matching relation in the reference list. On the other hand, a False Positive is a line match that does not correspond to any of the line matches in the reference list, and a False Negative is a line match that exists in the reference list but cannot be found by the automated approach.

Tab. 1: The matching performance of the pair-wise approach (TP = true positive, FP = false positive, FN = false negative).

Patch ID	Number of			Level of		
	TP	FP	FN	Correctness	Completeness	Quality
1	101	3	2	97 %	98 %	95 %
2	94	6	5	94 %	95 %	90 %
3	111	5	14	96 %	89 %	85 %

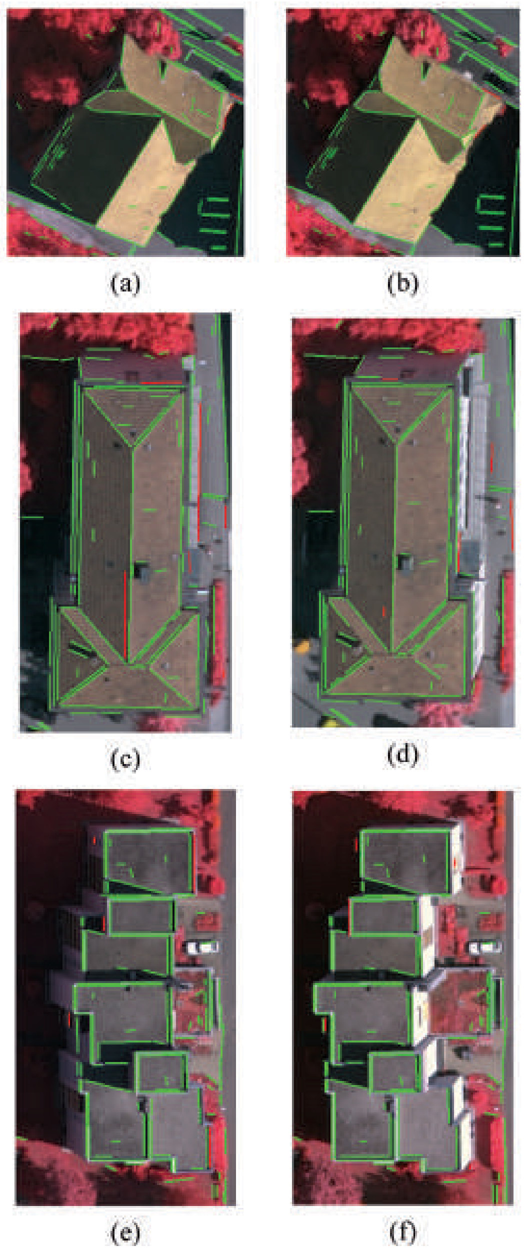


Fig. 7: Test patches. Left (a, c, e) and right (b, d, f) stereo images. Correct and false matches are shown in green and red colours, respectively.

In this study, we follow the well-known three metrics (correctness, completeness, and quality) (Rutzinger et al. 2009) to evaluate the quality of the line matching (Tab. 1). In all test patches, consistent correctness levels (> 94 %) are achieved. Similarly, the computed completeness levels for all patches are comparable. The highest number of FNs is found in the third test patch. This is due to the nature of the pair-wise approach: most of those FNs are located on the ground level and could not be paired correctly with other ground level line segments because of many obscured regions.

Despite the successful matching, the method of direct construction produced dramatic reconstruction problems for the lines that are nearly-aligned with the epipolar line (Fig. 8 a, d, g). It is clear from those figures that the reconstruction results of the line segments that are nearly-aligned with the epipolar line are extremely defective and irrelevant. On the other hand, our approach successfully recovered most of those problematic cases (Fig. 8 b, e, h). Although we believe that the level of improvement is visually apparent, we also evaluated the accuracy of the reconstructed line segments by comparing them to lidar data. The lidar data of the test site were captured with a Leica ALS50 system with an accuracy of 3–4 cm (Haala et al. 2010). In order to compare the reconstructed lines, we automatically extracted 3D planes from the point cloud in the vicinity of each line (Fig. 8 c, f, i). Depending on the type of the line, this plane reconstruction process resulted in one plane if the line corresponded to a step edge and in two planes if the line corresponded to the intersection of two planes. Thereafter, we determined the line's average orthogonal distance from its neighbouring planes and used these distances to compute the RMS average distance between the reconstructed lines and the lidar planes. The results for each patch are given in Tab. 2. As expected, the method of direct construction produced reasonable RMS distances (≈ 2 GSD) for the line segments that are not aligned with the epipolar line. On the other hand, large RMS distances (> 1 m) are inevitable for the nearly-aligned cases, which also reduce the overall performance considerably. In contrast to the results of the direct construction, for each test patch, our approach led to a considerable improvement of the RMS values after the reconstruction of the nearly-aligned line segments. Not surprisingly, in each case, this achievement also improves the overall RMS performance which is in the order of 2–3 GSD.

5 Conclusions

In this paper, we proposed a new reconstruction method for the line segments whose orientations in image space are found to be similar to the orientation of the epipolar line. The method manipulates the redundancy inherent in line-pair relations to generate artificial 3D points and utilize those entities in the estimation process to improve the accuracy of the reconstructed line segments. To test our approach, we selected three test patches over a built-up area of the city of Vaihingen, Germany. Based on the results, the proposed approach produced highly promising reconstruction results for the line segments that

Tab. 2: Comparison of the computed RMS distances of the Direct Construction method and the Proposed Approach. The numbers in brackets show the total number of planes used to compute the RMS distances.

Patch ID	Number of Planes Found		RMS Average Distance (m)					
			Direct Construction			Proposed Approach		
	Not Aligned	Nearly Aligned	Not Aligned	Nearly Aligned	Total	Not Aligned	Nearly Aligned	Total
1	153	30	0.152	1.041	0.495	0.152	0.357	0.204
2	85	27	0.173	4.451	2.240	0.173	0.196	0.179
3	119	32	0.194	6.278	3.768	0.194	0.459	0.275

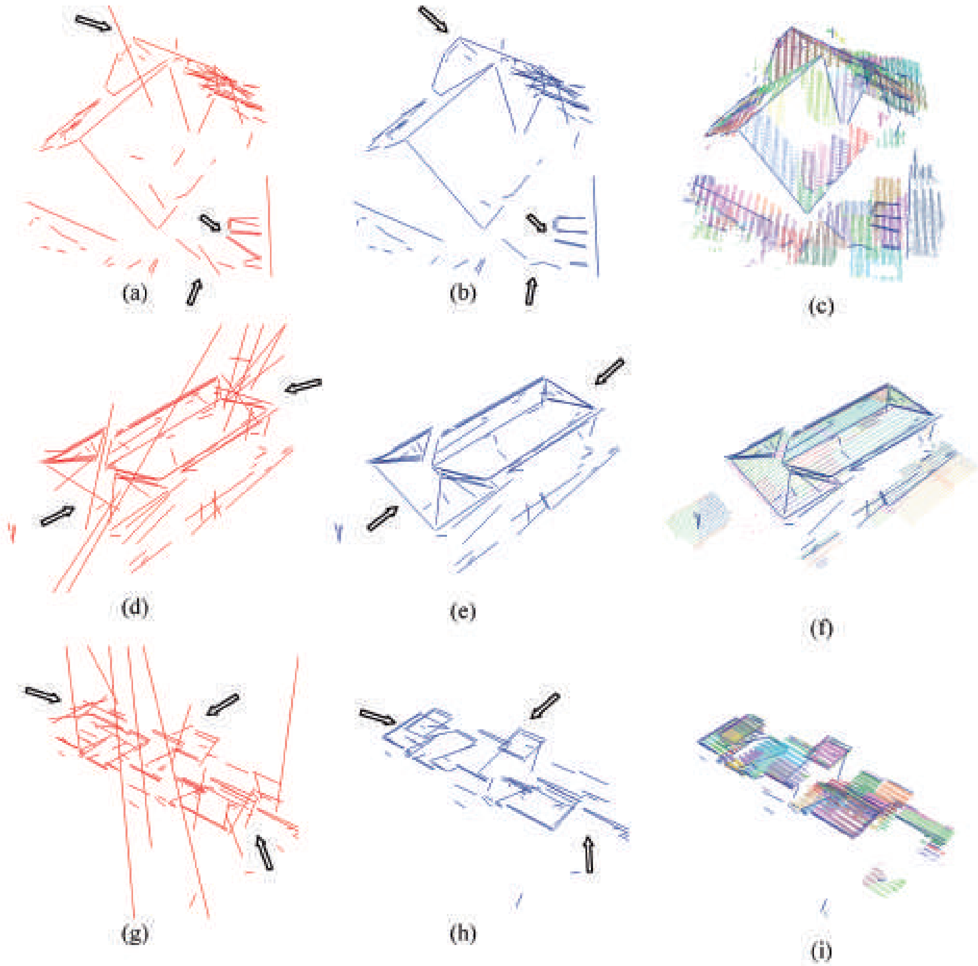


Fig. 8: 3D line segments generated with the method of direct construction (a, d, g) and our approach (b, e, h). Block arrows point to parts where some of the critical improvements are observed. The lidar point cloud overlaid with the lines reconstructed with our approach (c, f, i). Each colour in the point cloud belongs to the automatically extracted 3D planes in the vicinity of each line.

are nearly-aligned with the epipolar line. Indeed, the final improvement is considerable in terms of RMS distances, considering that the results of the previous approaches that just rely on direct construction consistently result in large RMS distances (> 1 m). Compared to the direct construction method, the gross errors are particularly reduced by the proposed approach, and an overall RMS performance of about 2–3 GSD is achieved. It is obvious that the selected 3D points $\mathbf{X}_i$ improve the final quality of the 3D line during the estimation process. In a worst-case scenario in which only a single 3D point entity is available, our method will not perform worse than the standard direct construction.

As a future work, it is a fact that the reconstruction of straight line segments which are nearly parallel to the epipolar line is also numerically unstable within a single image strip. Therefore, currently, we are extending the approach to deal with reconstruction drawbacks observed in image triplets that are acquired within a single strip, and the extended approach is expected to provide further improvements for the reconstruction of the near-epipolar line segments in aerial image triplets.

Acknowledgements

The Vaihingen dataset was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) (Cramer 2010): http://www.ifp.uni-stuttgart.de/dgpf/dkep-allg.html.

References

Baillard, C. & Dissard, O., 2000: A Stereo Matching Algorithm for Urban Digital Elevation Models. – Photogrammetric Engineering and Remote Sensing **66** (9): 1119–1128.

Bay, H., Ferrari, V. & Van Gool, L., 2005: Wide-baseline Stereo Matching with Line Segments. – IEEE Computer Society Conference on Computer Vision and Pattern Recognition **1:** 329–336.

Bulatov, D., 2011: Textured 3D Reconstruction of Urban Terrain from UAV-borne Video Sequences. – PhD thesis, Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover, Germany, **291**.

Cramer, M., 2010: The DGPF Test on Digital Aerial Camera Evaluation – Overview and Test Design. – PFG **2010** (2): 73–82.

Förstner, W., Brunn, A. & Heuel, S., 2000: Statistically Testing Uncertain Geometric Relations. – Sommer, G., Krüger, N. & Perwass, C. (eds.): Mustererkennung DAGM: 17–26, Springer.

Förstner, W., 2005: Uncertainty and Projective Geometry. – Bayro Corrochano, E. (ed.): Handbook of Geometric Computing: 493–535.

Haala, N., 2009: Come Back of Digital Image Matching. – Photogrammetric Week **2009:** 289–301, Wichmann Verlag, Heidelberg.

Haala, N., Hastedt, H., Wolf, K., Ressl, C. & Baltrusch, S., 2010: Digital Photogrammetric Camera Evaluation – Generation of Digital Elevation Models. – PFG **2010** (2): 99–116.

Heuel, S. & Förstner, W., 2001: Matching, Reconstructing and Grouping 3d Lines from Multiple Views Using Uncertain Projective Geometry. – IEEE Computer Vision and Pattern Recognition **2:** 517–524.

Heuel, S., 2001: Points, Lines and Planes and their Optimal Estimation. – Radig, B. & Florczyk, S. (eds.): DAGM, LNCS **2191:** 92–99, Springer, München.

Hirschmüller, H., 2008: Stereo Processing by Semi-global Matching and Mutual Information. – IEEE Transactions on Pattern Analysis and Machine Intelligence **30** (2): 328–341.

LGLN, 2012: Landesamt für Geoinformation und Landentwicklung Niedersachsen (LGLN), Digitale und Orientierte Luftbilder. http://www.lgn.niedersachsen.de/portal/live.php?navigation_id=11077&article_id=51723&_psmand=35 (30.1.2012).

Madsen, C.B. & Christensen, H.I., 1995: Modeling and Testing the Stability of Edge Segments: Length and Orientation. – Scandinavian Conference on Image Analysis: 1011–1019.

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

Ok, A.Ö., Wegner, J.D., Heipke, C., Rottensteiner, F., Sörgel, U. & Toprak, V., 2010a: A New Straight Line Reconstruction Methodology From Multi-Spectral Stereo Aerial Images. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **38** (3A): 25–30.

Ok, A.Ö., Wegner, J.D., Heipke, C., Rottensteiner, F., Sörgel, U. & Toprak, V., 2010b: A Stereo Line Matching Technique For Aerial Images Based on A Pair-Wise Relation Approach. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **38** (1/W17), Istanbul (on CD-ROM).

Rutzinger, M., Rottensteiner, F. & Pfeifer, N., 2009: A Comparison of Evaluation Techniques for Building Extraction from Airborne Laser Scanning. – IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **2** (1): 11–20.

Schmid, C. & Zisserman, A., 1997: Automatic Line Matching Across Views. – Computer Vision and Pattern Recognition: 666–671.

Scholze, S., Moons, T., Ade, F. & Van Gool, L., 2000: Exploiting Colour for Edge Extraction and Line Segment Stereo Matching. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences: 815–822.

Scholze, S., Moons, T. & Van Gool, L., 2002: A Generic 3D model for Automated Building Roof Reconstruction. – The International Archives of Photogrammetry and Remote Sensing **34** (5): 204–209.

Suveg, I. & Vosselman, G., 2004: Reconstruction of 3D Building Models from Aerial Images and Maps. – ISPRS Journal of Photogrammetry and Remote Sensing **58** (3-4): 202–224.

Wang, Z., Wu, F. & Hu, Z., 2009: MSLD: A Robust Descriptor for Line Matching. – Pattern Recognition **42** (5): 941–953.

Zhang, C. & Baltsavias, E.P., 2000: Edge Matching and 3D Road Reconstruction Using Knowledge-based Methods. – Schriftenreihe der Fachrichtung Geodäsie **10:** 251–265, Darmstadt.

Zhang, C., 2004: Towards an Operational System for Automated Updating of Road Databases by Integration of Imagery and Geodata. – ISPRS Journal of Photogrammetry and Remote Sensing **58** (3–4): 166–186.

Zhang, L., 2005: Automatic Digital Surface Model (DSM) Generation from Linear Array Images. – PhD Thesis, Swiss Institute of Technology Zurich.

Addresses of the Authors:

Dr. Ali Özgün Ok, Department of Civil Engineering, Mersin University, 33343, Mersin, Turkey, Tel.: +90-324-361-0001, Fax: +90-324-361-0015, e-mail: oozgun@mersin.edu.tr, ozguneo@gmail.com

Dr.-Ing. Jan Dirk Wegner, Prof. Dr.-Ing. Christian Heipke, PD Dr. Franz Rottensteiner & Prof. Dr.-Ing. Uwe Sörgel, Institut für Photogrammetrie und GeoInformation (IPI), Universität Hannover, Nienburger Str.1, 30167 Hannover, Germany, Tel.: +49-511-762-2482, Fax: +49-511-762-2483, e-mail: {wegner} {heipke} {rottensteiner} {soergel}@ipi.uni-hannover.de

Prof. Dr. Vedat Toprak, Geodetic and Geographic Information Technologies, Middle East Technical University, 06531, Ankara, Turkey, Tel.: +90-312-210-5748, Fax: +90-312-210-1263, e-mail: toprak@metu.edu.tr

Manuskript eingereicht: Oktober 2011
Angenommen: Mai 2012

PFG 2012 / 4, 0359–0370
Stuttgart, August 2012 **Article**

Change Detection in Peri-urban Areas Based on Contextual Classification

Txomin Hermosilla, José L. Gil-Yepes, Jorge A. Recio & Luis A. Ruiz, Valencia, Spain

Keywords: change detection, object-based classification, peri-urban areas, lidar

Summary: This paper presents a methodology for change detection in peri-urban areas using high spatial resolution image and lidar data, founded on object-based image classification and a comparison of the classification results from two epochs. The definition of the objects is based on cadastral boundaries obtained from a geospatial database. An exhaustive set of descriptive features is computed, characterising each object for both epochs regarding spectral, texture, geometrical, and three-dimensional (3D) aspects. In addition, contextual features describing the object at two levels are defined. Internal context features describe the relations between different land cover elements within the object, whereas external context features describe each object considering the common properties of neighbouring objects, usually coinciding in urban areas with an urban block. Both the classification and the change detection process are thoroughly evaluated, and the specific contribution of 3D features to the accuracy of the processes is analysed. The results show that 3D information enables to improve the classification results, remarkably increasing the accuracy values of certain classes, and allowing for an enhanced discrimination of building typologies. Moreover, the change detection efficiency is notably improved by a significant reduction of both commission and omission errors.

Zusammenfassung: Dieser Beitrag beschreibt eine Methodik zur Erkennung von Änderungen in Stadtrandgebieten mit Hilfe von räumlich hoch aufgelösten Bild- und Laserscannerdaten, die auf einer objektbasierten Klassifikation und einem anschließenden Vergleich der Klassifikationsergebnisse für zwei Epochen aufbaut. Die Definition der Objekte basiert auf Katastergrenzen aus einer räumlichen Datenbasis. Ein umfangreicher Satz von Merkmalen wird berechnet, der jedes Objekt aus beiden Epochen hinsichtlich seiner spektralen, texturellen, geometrischen und dreidimensionalen (3D) Aspekte charakterisiert. Zusätzlich werden Kontext-Merkmale, welche das Objekt auf zwei Ebenen beschreiben, definiert. Die internen Kontext-Merkmale beschreiben die Beziehungen zwischen verschiedenen Landbedeckungsarten innerhalb eines Objekts, während die externen Kontext-Merkmale jedes Objekt hinsichtlich der gemeinsamen Eigenschaften von benachbarten Objekten beschreiben; letztere sind für Objekte innerhalb eines Gebäudeblocks normalerweise ähnlich. Sowohl die Klassifikation als auch die Änderungsdetektion werden sorgfältig evaluiert, und der Einfluss der 3D Merkmale auf die Genauigkeit dieser Prozesse wird analysiert. Die Ergebnisse zeigen, dass die 3D Information dabei hilft, die Genauigkeit der Klassifikation zu steigern, wobei diese Steigerung für einige Klassen bemerkenswert ausfällt und insbesondere eine verbesserte Unterscheidung von Bebauungsarten erlaubt. Darüber hinaus wird die Effizienz der Änderungsdetektion deutlich erhöht, indem sowohl die Zahl der nicht erkannten Änderungen als auch jene der falschen Alarme signifikant reduziert wird.

1 Introduction

In many developed countries peri-urban areas are undergoing major transformations (Busck et al. 2006). These areas are under a considerable urban pressure that often entails their fast degradation and the abandonment of agriculture for a more profitable use of the land (Allen 2003, Huang et al. 2009). The high dynamics of peri-urban areas produce con-

DOI: 10.1127/1432-8364/2012/0123 1432-8364/12/0123 $ 3.00

tinuous changes of land cover and land use and, consequently, cartographic information is quickly outdated. Detailed and up-to-date cartographic and geographic information is needed for adequate urban management and planning.

Although urban growth has been traditionally monitored by using pixel-based classification techniques (Del Frate et al. 2005, Yuan et al. 2005, Zhou et al. 2008a) or standard change detection methodologies (Kwon et al. 2006, Xian & Homer 2010), many reliable change detection analyses in urban environments have been performed using object-based approaches. In an object-based approach, image analysis is performed by considering objects instead of pixels. An image object, or simply an object, is a group of pixels with common characteristics according to a segmentation criterion (Blaschke 2010). Several works have been presented using automatic segmentation methodologies applied to urban and peri-urban areas (Zhou et al. 2008b, Doxani et al. 2008, Lu et al. 2010, Myint et al. 2011). The segmentation method employed has a considerable impact on the extraction of descriptive features of objects, since the resultant object shapes and properties will differ depending on the algorithm used and the parameters selected (Hay & Castilla 2006).

Plot-based image analysis is a specific case of object-based image analysis that uses plot boundaries derived from land-use/land-cover (LU/LC) geospatial databases (i.e. cadastre, agricultural inventories). These boundaries enable the definition of semantic properties better than pixel aggregation algorithms. This methodology is suitable for environments such as urban, suburban and peri-urban areas, where landscape units present unambiguous boundaries. Plot-based approaches have also been employed for change detection in forest (Hall & Hay 2003) and in agricultural areas (Raza & Kainz 2002, Walter 2004, Gamanya et al. 2007), since they allow for an easy integration and relation of image-derived change information with LU/LC geospatial databases.

Buildings represent primitive entities of urban areas (Thomson & Béra 2008), and many change detection studies are focused on comparing detection or classification results corresponding to two different epochs (Hyyppä et al. 2007), or comparing the classification results from one date to historical databases (Vosselman et al. 2004, Vu et al. 2004, Rottensteiner 2007, Champion et al. 2010, Matikainen et al. 2010, Bouziani et al. 2010).

This paper evaluates a plot-based approach for change detection in peri-urban areas by comparing classification results obtained for data from two different epochs, taking into consideration a set of descriptive features extracted from high-resolution multispectral imagery. Additionally, contextual features were considered that correspond to two aggregation levels: internal and external. Lastly, the effect of the use of three-dimensional (3D) descriptive features derived from lidar data on the classification and change detection efficiency is evaluated.

2 Study Area and Data

The study area is located in the north of the city of Valencia, Spain (Fig. 1), having approximately 5,372 ha. The region has undergone important changes in LU/LC for the last decade, including the transformation and urbanisation of large agricultural areas formerly covered by orchards and horticulture crops, and the removal of some industrial and agricultural structures.

Fig. 1: Location of the study area in the north of the city of Valencia.

The objects were defined using the plot boundaries derived from cadastral maps produced by the Spanish General Directorate for Cadastre (Dirección General de Catastro), with a scale of 1:1,000 in urban areas and 1:2,000 in rural areas.

High-resolution multispectral imagery and lidar data were available for both epochs. For the first epoch, QuickBird imagery acquired in February 2004 was available, with 11 bits/pixel radiometric resolution and four spectral bands (visible and near infrared). The multispectral and panchromatic bands were merged using the substitution method based on principal components transformation, obtaining a final spatial resolution of 0.6 m/pixel. The image was georeferenced and orthorectified. Lidar data were acquired in December 2003 by an ALTM-2033 sensor, and a digital surface model (DSM) having a grid size of 1 m was computed. In addition, a manually edited digital terrain model (DTM) was available. The normalised digital surface model (nDSM) was generated as the difference between the DSM and the DTM.

Additional aerial images were acquired in August 2008 in the framework of the Spanish National Plan of Aerial Orthophotography (PNOA), with 0.5 m/pixel spatial resolution, 8 bits/pixel radiometric resolution and four spectral bands in the visible and near infrared domains. These images were orthorectified and georeferenced. Panchromatic and multispectral bands were merged, and mosaicking and radiometric adjustments were applied as a part of the PNOA programme. Lidar data were acquired in September 2009 using a RIEGL LMS-Q680 sensor with a nominal density of 0.5 points/m^2. The DTM was computed by means of the iterative algorithm described by Estornell et al. (2011) that selects minimum elevation points and eliminates points belonging to above-ground elements such as vegetation or buildings.

3 Methodology

Land use classification and change detection were carried out according to the following steps: class definition, sample selection, feature extraction, object classification at both epochs, change detection, and evaluation. A plot-based image classification approach was used, where the objects were defined by the cartographic boundaries of the plots in the cadastral database. The objects were exhaustively described by image-based features, geometrical features describing the shape of each object, and contextual features. Contextual features were defined at two levels: internal and external. Urban change detection – concerning building construction, destruction or use alteration – was performed by comparing the class assigned to each plot at both epochs.

3.1. *Definition of Classes and Sample Selection*

The definition of land-use classes was based on the land cover catalogue of the Land Cover and Land Use Information System of Spain (SIOSE). This database, corresponding to a scale of 1:25,000, was created to homogenize former LU/LC national databases. Based on this catalogue seven classes were defined, and some of them were subdivided into more specific classes for classification purposes. Thus, four building-related classes were defined: *historical*, *planned urban* (subdivided in *closed urban* and *open urban*), *industrial*, and *suburban housing* (consisting of *semidetached housing* and *detached housing*). The rest of the classes were *arable land* and *croplands* (internally divided into *cropland* and *arable land*, depending of the presence of vegetation), *citrus orchard*, and *bare soil*. Fig. 2 shows examples of the classes.

Training samples were collected by assigning classes to objects (cadastral plots) by visual photointerpretation, applying two main criteria: the representativeness of samples selected per class; and the homogeneity of the spatial distribution of samples in the study area. A restricted randomization scheme (Chatfield 1991) was applied consisting of a random sample selection followed by a monitored sample reallocation and selection in order to maintain a minimum number of samples per class, to obtain a representative sample of each class, and to represent the changes in the area. As a result, 1458 training samples were selected (see sample number distribution for both ep-

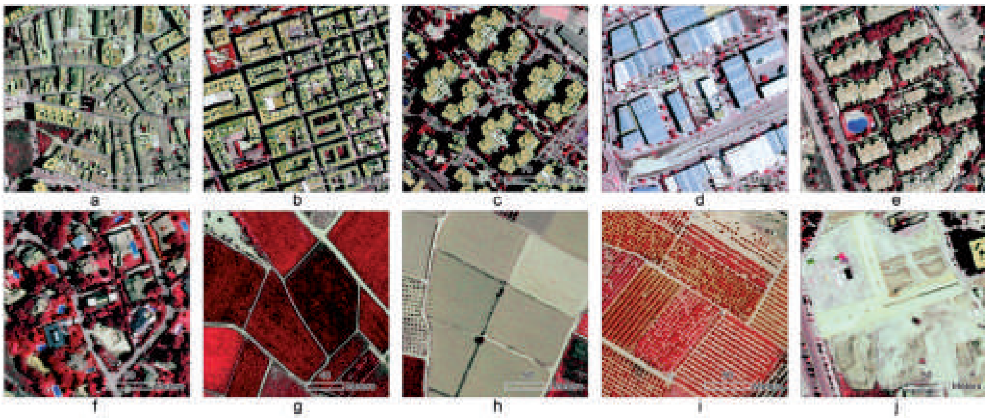


Fig. 2: Examples of the classes defined: historical (a), closed urban (b), open urban (c), industrial (d), semidetached housing (e), detached housing (f), cropland (g), arable land (h), citrus orchard (i), and bare soil (j).

ochs in Tab. 1). A total of 128 objects (8.8 %) of the samples corresponded to urban changes, concerning building construction, destruction, or use change.

3.2. *Feature Extraction*

Descriptive features were extracted attending to three different object aggregation levels: object-based, related to internal context and related to external context (Fig. 3). Object-based features describe the properties of each object (plot) considered as a single unit. They were computed using the object-based feature extraction software FETEX 2.0 (Ruiz et al. 2011). These descriptive features provide information about spectral, textural, geometrical, and 3D properties. Spectral features inform, through the intensity values of the pixels contained in the plots, about the objects' overall spectral behaviour in the different spectral bands used. Mean, standard deviation, as well as minimum and maximum of the intensity values of the pixels for all available bands and an NDVI image were computed for each object (NDVI = normalized difference vegetation index). Texture features quantify the spatial distribution of the intensity values in the analysed objects. The following descriptive features were derived: histogram kurtosis and skewness, descriptors derived from the grey level co-occurrence matrix (GLCM) proposed by Haralick et al. (1973), edgeness factor (Sutton & Hall 1972), and semivariogram-based descriptive features (Balaguer et al. 2010). Three-dimensional features were derived from the nDSM computed from lidar data, each object being characterised by the mean, standard deviation, and maximum values of the heights. Geometrical features describing the dimensions of the objects and their contour complexity were computed: area, perimeter, compactness (Bogaert et al. 2000), shape index and fractal dimension (Krummel et al. 1987, McGarigal & Marks 1995).

Some of the context-based features were based on the area covered by buildings and vegetation inside each object, since this distri-

Tab. 1: Number of samples per class used in both dates.

Classes	Dates	
	2004	2008
Historical	213	213
Urban	234	298
Industrial	105	98
Suburban housing	222	256
Arable land and croplands	277	259
Citrus crop	235	182
Bare soil	172	152
Total	1458	1458

bution is strongly related to the different urban typologies. Building and vegetation masks were obtained using an automatic building detection technique consisting of applying a multiple-threshold based approach, as described in Hermosilla et al. (2011). Buildings and vegetation correspond to the sub-objects inside the plots.

Internal-context features describe an object by characterising the internal sub-objects of the plots. Both 2D and 3D features describing the buildings inside each object were computed. The 2D features consist of the built-up area and the percentage of built-up areas in an object. The latter feature – usually referred to as building coverage ratio (BCR) or sealed surface (Yoshida & Omae 2005, Van de Voorde et al. 2009, Yu et al. 2010) – is computed as described in (1):

$$BCR = \frac{A_{Building}}{A_{Object}} \cdot 100 \qquad (1)$$

where $A_{Building}$ is the built-up area, and A_{Object} is the area of the considered object. Buildings were also characterised by a set of 3D features describing their heights, using the mean, standard deviation, and maximum values from the nDSM. Similarly to (1), the percentage of surface covered by vegetation within an object was defined, and additionally a set of statistical descriptors of height and status of vegetation sub-objects (mean and standard deviation of nDSM and NDVI, extracted only from groups of vegetation pixel) was computed.

Due to the hierarchical structure of urban landscapes, the analysis and exploration of their various aggregation levels result in a precise analysis of the relations of objects (Bruzzone & Carlin 2006). Features based on external context provide information about the properties of the urban blocks. An urban block is the area surrounded by public roads or streets and it may be subdivided into any number of plots. In the Spanish urban cadastral maps, streets are not represented by polygons. This enables to delimit the urban blocks by iteratively merging adjacent plots. Urban blocks significantly determine the appearance of urban environments, influencing spatial experience and defining local particularities related to a spatial identity (Laskari

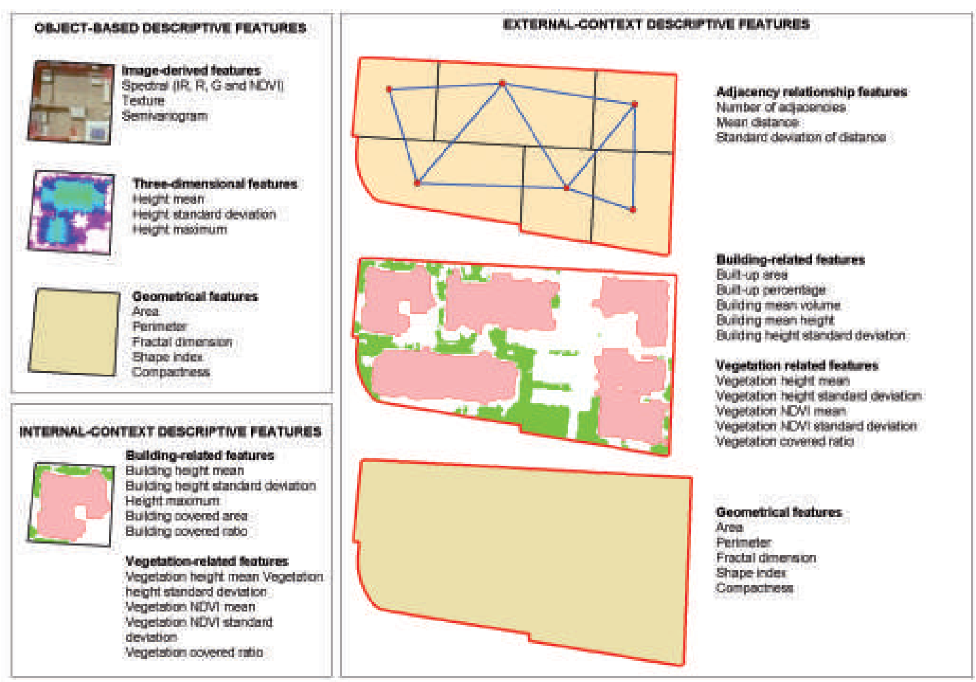


Fig. 3: List and graphic examples of the features computed.

et al. 2008). Urban block entities have been commonly employed in urban environments to define a higher hierarchical context (Bauer & Steinnocher 2001, Herold et al. 2003, Wijnant & Steenberghen 2004, Novack et al. 2010, Huck et al. 2011). Two-stage approximation methods are frequently employed. Initially, the main land-cover types are detected and this information is afterwards analysed in the spatial context defined by the urban blocks to determine land use. Hermosilla et al. (2012) showed that the specific addition of external context features derived from urban blocks increases the classification accuracy of different typologies of urban plots. Applying this methodology, all the plots contained in the same urban block are described with the same external context features. Thus, external context is described by considering the spatial relations of adjacent objects by means of building-based, vegetation-based, geometrical and adjacency features. Adjacency between objects was characterised based on the graph theory (Laurini & Thompson 1992) by using the number of neighbours with surrounding objects, as well as the mean and standard deviations of the centroid distances between adjacent objects. The shape, size, and number of buildings per block are often related to their socio-economic function and determine the area and volume for an urban block. Therefore, the land use of an urban block may be indicated by the quantitative observations related to the buildings present in it (Yoshida & Omae 2005). Thus, urban blocks were also characterised by the built-up area and the built-up percentage. The height distribution of the buildings contained in an urban block was described using the height mean and standard deviation values obtained from the nDSM. Features related to the volumetric information of buildings were also considered. Thus, the mean volume was computed as the total volume of buildings divided by the number of buildings contained in the block. In a similar manner to the internal context features, vegetation distribution was characterised using the vegetation cover ratio, as well as the mean and standard deviation values of the nDSM and NDVI obtained only from the vegetation area masked within the super-object. Finally, the geometrical properties of the urban blocks were described using the features area, perimeter, compactness, shape index, and fractal dimension.

3.3. *Classification, Accuracy Assessment and Change Detection*

In order to evaluate and quantify the effect of the 3D features in the change detection process, two classifications were performed per epoch: with and without using lidar-derived features. Classification was carried out using decision trees constructed with the C5.0 algorithm (Quinlan 1993), combined with the boosting technique. This algorithm divides the sample set by using mutually exclusive conditions, until homogeneous subgroups are generated, i.e., all the elements in a subgroup belong to the same class, or a stopping condition is satisfied. The stopping condition defined in this work is the minimum number of training cases that must follow each node. It was fixed to 5 cases, constraining the degree to which the initial tree fits the training data.

In order to maximize the efficiency of the number of available samples, the accuracy of the classification was assessed using leave-one-out cross-validation. This method uses a single observation from the original sample set as validation data, and the remaining observations as training data, iterating until each observation in the sample set is used for validation once. The evaluation of the classification was based on the analysis of the confusion matrix, which compares the class assigned to each sample to the reference class obtained by photointerpretation. From the confusion matrix, the user's and producer's accuracies per class were computed, measuring the commission and omission errors, respectively (Congalton 1991).

Changes were detected by comparing the classes assigned to each plot in the classifications of both epochs. The change detection process was focused on urban changes concerning building construction, destruction or use change. Changes produced between agricultural classes were not considered in this study. Although many land-use transitions are theoretically possible, there are limitations by law and by nature which may

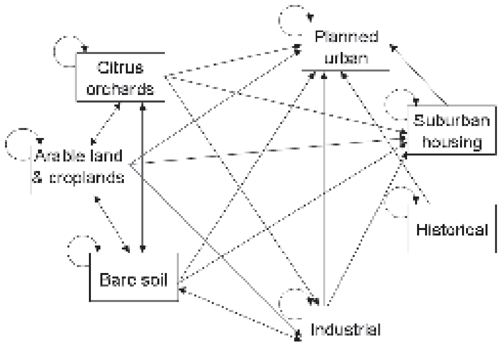


Fig. 4: Land-use transition diagram, showing the possible changes and hiding the restricted land-use transitions.

be used to improve the change detection process (Pakzad 2002). Following these criteria, a knowledge-based land-use transition diagram was designed to formulate the possible land-use changes that are likely to occur in the study area (Fig. 4), restricting unlikely land-use changes. Therefore, when a change that does not fulfil the land-use transition diagram is detected, it is directly removed and considered as no change. These unlikely transitions would necessarily require the revision by a human operator before being fully accepted or rejected in a LU/LC geospatial database updating process.

The LU/LC differences found between the classification results of both epochs enable the detection of the actual LU/LC changes occurring in the territory, but also reveal classification errors. The result of the change detection process is shown in the matrix of Tab. 2. The change detection efficiency is defined by adding the correctly detected changes to the correctly detected unchanged objects. Detectable errors are given when unchanged plots are wrongly detected as changed, whereas undetectable errors are produced when plots where changes occurred are detected as unchanged. The combination of detected changes and detectable errors composes the number of plots that would be manually reviewed in an updating process.

4 Results and Discussion

Classification overall accuracy values obtained for both epochs with and without considering 3D data are shown in Tab. 3. The classification performance without considering lidar data is approximately 90 %, while the subsequent addition of 3D descriptors results in an increase of the accuracy of about 5 %.

Analyzing the specific per-class accuracy values shown in Fig. 5a, the *bare soil* class shows the lowest user's and producer's accuracies without considering 3D data. This can be attributed to the high variability of this class, which includes a diversity of agricultural and non-agricultural plots, such as abandoned agricultural or un-built areas. In the same sense, the class *industrial* also has low accuracy values, especially when compared to the other building-related classes. Due to the similar spectral and textural properties of *industrial* and *bare soil*, these classes were occasionally confused, notably affecting their respective errors. For the rest of the classes, high user's and producer's accuracies were obtained for both epochs.

Although the inclusion of 3D information from lidar data has a limited effect on the overall accuracy (about 5 %), some classes noticeably improve their user's and producer's accuracies (Fig. 5b). Especially remarkable is

Tab. 2: Distribution of cases and errors in the change detection assessment.

		Classification	
		Unchanged	Changed
Reference	Unchanged	**Coincidences**	**Detectable errors**
	Changed	**Undetectable errors**	**Detected changes**

Tab. 3: Classification overall accuracy values achieved for both epochs with and without considering 3D features, for a total of seven LU/LC classes (as in Fig. 5).

Date	Overall accuracy	
	Without lidar	With lidar
2004	90.3 %	94.5 %
2008	89.2 %	94.5 %

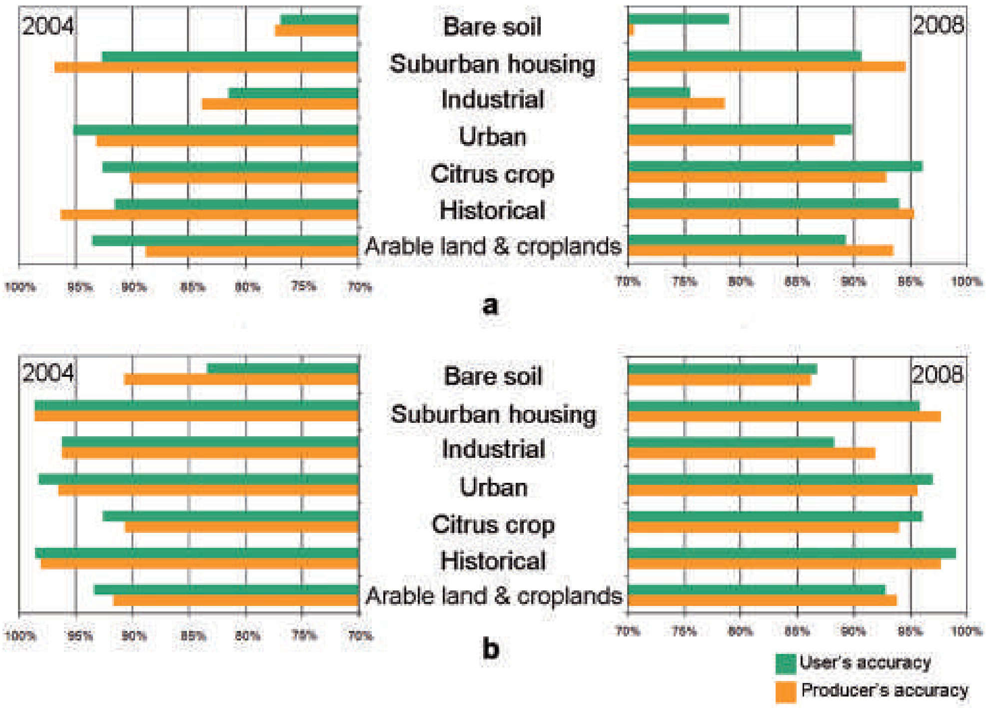


Fig. 5: Per-class user's and producer's accuracies in both dates without (a) and with three-dimensional data (b).

Tab. 4: Change detection assessment concerning urban-related changes considering the land-use transition diagram shown in Fig. 4.

		Without lidar data		With lidar data	
		Unchanged	Changed	Unchanged	Changed
Reference	Unchanged	**87.8 %**	**3.7 %**	**90.6 %**	**0.9 %**
	Changed	**1.9 %**	**6.7 %**	**0.4 %**	**8.1 %**

the case of class *industrial*, where misclassifications are significantly reduced when using 3D information. *Bare soil* has a more modest improvement of the accuracy values, due to confusion with *arable land* and *croplands*. The accuracies of *historical*, *urban* and *suburban housing* classes increase when 3D information is added, even though their values were already particularly high, producing more balanced user's and producer's accuracies per class. Finally, *citrus* accuracies do not increase when 3D information is added, probably because of the correct characterization provided by the texture and semivariogram-based features used.

The change detection results are shown in Tab. 4. When 3D data are not considered, the detection efficiency is 94.5 %, obtained as a result of adding correctly detected changes plus correctly detected unchanged objects. The percentage of detectable errors is 3.7 %, and the percentage of plots that would be manually revised to be confirmed as actual changes is 10.4 %. There is a significant rate of undetectable errors (1.9 %), mostly due to misclassifications of *bare soil* for one of the epochs.

Additionally, buildings under construction, which were considered as *bare soil* samples, are finally classified as one of the building-related classes. Moreover, the shadows of neighbour buildings affect the classification negatively, increasing the number of undetectable errors.

When 3D features derived from lidar data are used, both detectable and undetectable errors are reduced by approximately 75 %, enabling the change detection efficiency to achieve up to 98.7 %. Thus, the number of plots to be manually reviewed using photointerpretation and field visits would decrease from 10.4 % to 9 %, mostly corresponding to plots with actual changes. Height information has been shown to be critical to satisfactorily discriminate between building typologies and buildings from *bare soil*, correcting most of the detectable and undetectable errors, and improving the performance of the change detection process. Fig. 6 shows an image detail of the study area and the change detection represented for the selected samples with and without considering 3D data, where the significant reduction of undetectable errors and the increase of detected changes are perceptible.

For comparative purposes, Tab. 5 shows the change detection assessment results when the land-use transition diagram (Fig. 4) is not considered. Contrasting these results with the results presented in Tab. 4 it is noticeable that the restriction of unlikely changes enables to reduce the detectable errors by 35 % to 45 % with and without lidar data, respectively, and to increase the coincidences. Undetectable errors practically remain invariable, and the amount of detected changes is slightly reduced.

5 Conclusions

In this paper, an object-based classification methodology for building change detection in peri-urban areas is presented and evaluated. Objects are defined by using cadastral plot boundaries, allowing the direct relation of the information derived from imagery to the information present in LU/LC databases. A comprehensive set of descriptive features

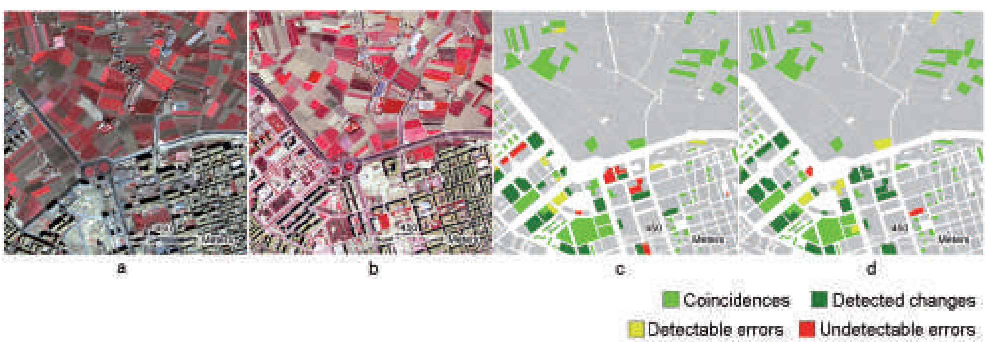


Fig. 6: Detail of study area in colour infrared composition for years 2004 (a) and 2008 (b); and maps showing change detection results of samples without considering (c) and considering (d) 3D features.

Tab. 5: Change detection assessment concerning urban-related changes without considering the land-use transition diagram.

		Without lidar data		With lidar data	
		Unchanged	Changed	Unchanged	Changed
Reference	Unchanged	**84.8 %**	**6.7 %**	**89.8 %**	**1.4 %**
	Changed	**1.9 %**	**6.9 %**	**0.5 %**	**8.3 %**

characterising each object was extracted at both epochs (2004 and 2008). These features described the particular properties of the objects – spectral, textural, geometrical, and three-dimensional –, and their context at two levels, internal – related to those elements inside the plot –, and external – related to the common properties of plots sharing a superior aggregation level. The specific effect of 3D features on the efficiency of the classification and change detection processes has also been evaluated. In our experiments, this information was derived from lidar data, however, similar results can be expected when using stereo images as the source of 3D information.

The results showed that, although 3D information only produces a slight improvement of the overall classification accuracy, it notably increases the specific accuracy values of some classes, allowing for a particularly enhanced discrimination of building typologies and bare soil. In the change detection process, the use of 3D information produces a notable reduction of both detectable and undetectable errors, remarkably diminishing the number of plots that would be reviewed by human operators in a changing detection procedure, in order to confirm the veracity of the detected changes. As a consequence, the overall change detection efficiency considerably improves (from 94.5 % to 98.7 %).

Although this change detection methodology has been tested in a peri-urban area, it could be adapted for being applied to different scenarios, by selecting a specific set of representative samples of those areas and their particular object class catalogues. In future work, the reduction of sample collection tasks could be further studied by using only samples collected on the initial date, or even directly obtained from the original LU/LC geospatial database to be updated, assuming that the proportion of changes is low compared to the unchanged areas.

In summary a methodology based on automated descriptive feature extraction from 3D data, high spatial resolution imagery and context information using an object-based approach is presented, which may be used to increase the efficiency and reduce photointerpretation and field tasks in many LU/LC geospatial database change detection procedures.

Acknowledgements

The authors appreciate the financial support provided by the Spanish Ministerio de Ciencia e Innovación and FEDER in the framework of the projects CGL2009-14220 and CGL2010-19591/BTE, and by the Spanish Instituto Geográfico Nacional.

References

Allen, A., 2003: Environmental Planning and Management of the Peri-urban Interface: Perspectives on an Emerging Field. – Environmental Planning and Management **15** (1): 134–147.

Balaguer, A., Ruiz, L.A., Hermosilla, T. & Recio, J.A., 2010: Definition of a Comprehensive Set of Texture Semivariogram Features and their Evaluation for Object-oriented Image Classification. – Computers and Geosciences **36** (2): 231–240.

Bauer, T. & Steinnocher, K., 2001: Per-parcel Land Use Classification in Urban Areas Applying a Rule-based Technique. – GEOBIT/GIS **6:** 24–27.

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

Bogaert, J., Rousseau, R., Hecke, P.V. & Impens, I., 2000: Alternative Area-perimeter Ratios for Measurement of 2D Shape Compactness of Habitats. – Applied Mathematics and Computation **111** (1): 71–85.

Bouziani, M., Goïta, K. & Dong-Chen, H., 2010: Automatic Change Detection of Buildings in Urban Environment from Very High Spatial Resolution Images Using Existing Geodatabase and Prior Knowledge. – ISPRS Journal of Photogrammetry and Remote Sensing **65** (1): 143–153.

Busck, A.G., Kristensen, S.P., Præstholm, S., Reenberg, A. & Primdahl, J., 2006: Land System Changes in the Context of Urbanisation: Examples from the Peri-urban Area of Greater Copenhagen. – Danish Journal of Geography **106** (2): 21–34.

Bruzzone, L. & Carlin, L., 2006: A Multilevel Context-based System for Classification of Very High Spatial Resolution Images. – IEEE Transactions on Geoscience Remote Sensing **44** (9): 2587–2600.

Champion, N., Boldo, D., Pierrot-Deseilligny, M. & Stamon, G., 2010: 2D Building Change Detection from High Resolution Satellite Imagery: A Two-step Hierarchical Method Based on 3D Invariant Primitives. – Pattern Recognition Letters **31** (10): 1138–1147.

Chatfield, C., 1991: Avoiding Statistical Pitfalls. – Statistical Science **6** (3): 240–252.
Congalton, R., 1991: A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. – Remote Sensing of Environment **37** (1): 35–46.
Del Frate, F., Schiavon, G. & Solimini, C., 2005: Change Detection in Urban Areas with Quick-Bird Imagery and Neural Networks Algorithms. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (8/W27): 2 pp.
Doxani, G., Siachalou, S. & Tsakiri-Strati, M., 2008: An Object-oriented Approach to Urban Land Cover Change Detection. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **37** (B7): 1655–1660.
Estornell, J., Ruiz, L.A., Velázquez-Martí, B. & Hermosilla, T., 2011: Analysis of the Factors Affecting LiDAR DTM Accuracy in a Steep Shrub Area. – International Journal of Digital Earth **4** (6): 521–538.
Gamanya, R., Maeyer, P.D. & Dapper, M.D., 2007: An Automated Satellite Image Classification Design Using Object-oriented Segmentation Algorithms: A Move Towards Standardization. – Expert Systems with Applications **32** (2): 616–624.
Hall, O. & HAY, G.J., 2003: A Multiscale Object-specific Approach to Digital Change Detection. – International Journal of Applied Earth Observation and Geoinformation **4** (4): 311–327.
Hay, G.J. & Castilla, G., 2006: Object-based Image Analysis: Strengths, Weaknesses, Opportunities and Threats (SWOT). – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science **36** (4/C42): 3 pp.
Haralick, R.M., Shanmugan, K. & Dinstein, I., 1973: Texture Features for Image Classification. – IEEE Transactions on Systems, Man and Cybernetics **3:** 610–621.
Hermosilla, T., Ruiz, L.A., Recio, J.A. & Estornell, J., 2011: Evaluation of Automatic Building Detection Approaches Combining High Resolution Images and LiDAR Data. – Remote Sensing **3** (6): 1188–1210.
Hermosilla, T., Ruiz, L.A., Recio, J.A. & Cambra-López, M., 2012: Assessing contextual descriptive features for plot-based classification of urban areas. – Landscape and urban Planning **106** (1): 124–137.
Herold, M., Liu, X. & Clarke, K.C., 2003: Spatial Metrics and Image Texture for Mapping Urban Land Use. – Photogrammetric Engineering and Remote Sensing **69** (9): 991–1001.
Huang, S.L., Wang, S.H. & Budd, W.W., 2009: Sprawl in Taipei's Peri-urban Zone: Responses to Spatial Planning and Implications for Adapting Global Environmental Change. – Landscape and Urban Planning **90:** 20–32.
Huck, A., Hese, S. & Banzhaf, E., 2011: Delineating Parameters for Object-based Urban Structure Mapping in Santiago de Chile Using Quick-Bird Data. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **38** (4/W19): 6 pp.
Hyyppä, J., Matikainen, L., Kaartinen, H., Xiaowei, Y., Hyyppä, H. & Rönnholm, P., 2007: Improving Automation in Map Updating Based on National Laser Scanning, Classification Trees, Object-based Change Detection and 3D Object Reconstruction. – Urban Remote Sensing Joint Event: 1–10, Paris, France.
Krummel, J.R., Gardner, R.H., Sugihara, G., O'Neill, V. & Coleman, P.R., 1987: Landscape Patterns in a Disturbed Environment. – OIKOS **48** (3): 321–324.
Kwon, S., Shin, S. & Kim, K., 2006: A Development of Change Detection System for Urban Monitoring. – IEEE Geoscience and Remote Sensing Symposium, IGARSS **06:** 1508–1511.
Laskari, S., Hanna, S. & Derix, C., 2008: Urban Identity Through Quantifiable Spatial Attributes: Coherence and Dispersion of Local Identity Through the Automated Comparative Analysis of Building Block Plans. – **Third** International Conference on Design Computing and Cognition: 615–634, Dordrecht, The Netherlands.
Laurini, R. & Thompson, D., 1992: Fundamentals of Spatial Information (5). – Academic Press, London, UK.
Lu, D., Hetrick, S., Moran, E. & Li, G., 2010: Detection of Urban Expansion in an Urban-rural Landscape with Multitemporal QuickBird Images. – Journal of Applied Remote Sensing **4** (1): 17 pp.
Matikainen, L., Hyyppä, J., Ahokas, E., Markelin, L. & Kaartinen, H., 2010: Automatic Detection of Buildings and Changes in Buildings for Updating of Maps. – Remote Sensing **2:** 1217–1248.
McGarigal, K. & Marks, B.J., 1995: FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure. – General Technical Report PNW-GTR-**351**: 122 pp., U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR, USA.
Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S. & Weng, Q., 2011: Per-pixel vs. Object-based Classification of Urban Land Cover Extraction Using High Spatial Resolution Imagery. – Remote Sensing of Environment **115** (5): 1145–1161.

Novack, T., Kux, H.J.H., Feitosa, R.Q. & Costa, G.A., 2010: Per Block Urban Land Use Interpretation Using Optical VHR Data and the Knowledge-based System Interimage. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **38** (4/C7): 6 pp.

Pakzad, K., 2002: Knowledge Based Multitemporal Interpretation. – The International Archives of Photogrammetry and Remote Sensing **34** (3A): 234–239.

Quinlan, J.R., 1993: C4.5: Programs for Machine Learning. – Morgan Kaufmann Publisher, San Francisco, CA, USA.

Raza, A. & Kainz, W., 2002: An Object-oriented Approach for Modeling Urban Land-use Changes. – URISA Journal **14** (1): 37–55.

Rottensteiner, F., 2007: Building Change Detection from Digital Surface Models and Multispectral Images. – The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (3/W49B): 145–150.

Ruiz, L.A., Recio, J.A., Fernández-Sarría, A. & Hermosilla, T., 2011: A Feature Extraction Software Tool for Agricultural Object-based Image Analysis. – Computers and Electronics in Agriculture **76** (2): 284–296.

Sutton, R.N. & Hall, E.L., 1972: Texture Measures for Automatic Classification of Pulmonary Disease. – IEEE Transactions on Computers **21** (7): 667–676.

Thomson, M.K. & Béra, R., 2008: A Methodology for Inferring Higher Level Semantic Information from Spatial Databases. – Geographical Information Science Research UK 16th Annual Conference, GISRUK **08:** 268–274, Manchester, United Kingdom.

Van de Voorde, T., Van der Kwast, J., Engelen, G., Binard, M., Cornet, Y. & Canters, F., 2009: Quantifying Intra-urban Morphology of the Greater Dublin Area With Spatial Metrics Derived from Medium Resolution Remote Sensing Data. – IEEE Geoscience and Remote Sensing Society, 7th International Urban Remote Sensing Joint Event **2009:** 1–7, Shanghai, China.

Vosselman, G., Gorte, B. & Sithole, G., 2004: Change Detection for Updating Medium Scale Maps Using Laser Altimetry. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **35** (B3): 207–212.

Vu, T.T., Matsuoka, M. & Yamazaki, F., 2004: LIDAR-based Change Detection of Buildings in Dense Urban Areas. – IEEE International Conference, Geoscience and Remote Sensing Symposium, IGARSS **04:** 3413–3416.

Walter, V., 2004: Object-based Classification of Remote Sensing Data for Change Detection. – ISPRS Journal of Photogrammetry and Remote Sensing **58** (3–4): 225–238.

Wijnant, J. & Steenberghen, T., 2004: Per-parcel Classification of Urban Ikonos Imagery. – 7th AGILE conference on geographic information science: 447–455, Heraklion, Greece.

Xian, G. & Homer, C., 2010: Updating the 2001 National Land Cover Database Impervious Surface Products to 2006 Using Landsat Imagery Change Detection Methods. – Remote Sensing of Environment **114** (8): 1676–1686.

Yoshida, H. & Omae, M., 2005: An Approach for Analysis of Urban Morphology: Methods to Derive Morphological Properties of City Blocks by Using an Urban Landscape Model and Their Interpretations. – Computers, Environment and Urban Systems **29** (2): 223–247.

Yu, B., Liu, H., Wu, J., Hu, Y. & Zhang, L., 2010: Automated Derivation of Urban Building Density Information Using Airborne LiDAR Data and Object-based Method. – Landscape and Urban Planning **98** (3–4): 210–219.

Yuan, F., Sawaya, K.E., Loeffelholz, B.C. & Bauer, M.E., 2005: Land Cover Classification and Change Analysis of the Twin Cities (Minnesota) Metropolitan Area by Multitemporal Landsat Remote Sensing. – Remote Sensing of Environment **98** (2–3): 317–328.

Zhou, W., Troy, A. & Grove, M., 2008a: A Comparison of Object-based with Pixel-based Land Cover Change Detection in the Baltimore Metropolitan Area Using Multitemporal High Resolution Remote Sensing Data. – IEEE International Conference, Geoscience and Remote Sensing Symposium, IGARSS **08:** 683–686.

Zhou, W., Troy, A. & Grove, M., 2008b: Object-based Land Cover Classification and Change Analysis in the Baltimore Metropolitan Area Using Multitemporal High Resolution Remote Sensing Data. – Sensors **8** (3): 1613–1636.

Address of the Authors:

Dr. Txomin Hermosilla, José L. Gil-Yepes, Dr. Jorge A. Recio, Dr. Luis A. Ruiz, Universitat Politècnica de València, Geo-Environmental Cartography and Remote Sensing Research Group, Camino de Vera s/n, 46022 Valencia, Spain, Tel.: +34-963877000 ext. 75576, 75576, 79553, 75536, e-mail: txohergo@topo.upv.es, jogiye@upvnet.upv.es, jrecio@cgf.upv.es, laruiz@cgf.upv.es

Manuskript eingereicht: August 2011
Angenommen: Januar 2012

PFG 2012 / 4, 0371 – 0384
Stuttgart, August 2012

Article

Kernel Composition with the one-against-one Cascade for Integrating External Knowledge into SVM Classification

Andreas Ch. Braun, Uwe Weidner, Boris Jutzi & Stefan Hinz, Karlsruhe

Keywords: data fusion, knowledge integration, SVM, hyperspectral, airborne laserscanning

Summary: This work focuses on two main questions. How can data fusion be performed before SVM (support vector machine) classification? And secondly: how can the one-against-one cascade be exploited to use information selectively thus integrating human knowledge? Kernel composition represents a specialized method for fusing data on the feature level. Its main advantage is given by the fact that it reduces the Hughes phenomenon (performance decrease due to high dimensionality) because it abstains from raising dimensionality in the feature space. Since the paper focuses on hyperspectral data, a specialized kernel based on the spectral angle is employed and evaluated. Two application schemes are presented. At first, hyperspectral data are fused with laserscanning data by taking into account explicit knowledge on roof geometries. Secondly, a spectral-spatial framework for hyperspectral data is presented which integrates implicit knowledge on the relevance of spatial context into classification. Both approaches are promising as they obtain higher classification accuracies when integrating external knowledge. The innovation of the contribution is that data fusion with a second source of data via kernel composition is combined with a modification of the one-against-one cascade which allows integration of human knowledge.

Zusammenfassung: *Verknüpfung von Kernfunktionen mit der eins-gegen-eins Kaskade für die Einbindung von Wissen in die SVM Klassifizierung.* Dieser Beitrag vertieft zwei Hauptfragen. Wie kann die Datenfusion für die SVM Klassifizierung vorgenommen werden? Und zweitens: wie kann die eins-gegen-eins Kaskade genutzt werden, um Information selektiv zu nutzen und menschliches Wissen einzubringen? Die Verknüpfung von Kernfunktionen stellt eine spezielle Methode der Datenfusion für kernbasierte Klassifikatoren wie Stützvektormaschinen (support vector machines, SVM) dar. Ihr Hauptvorteil besteht darin, dass dadurch das Hughes Phänomen (Performanzverlust durch hohe Dimensionalität) reduziert wird, indem sie es vermeidet, die Dimensionalität des Merkmalsraums zu erhöhen. Da sich der Beitrag mit hyperspektralen Daten beschäftigt, wird eine spezielle Kernfunktion, die auf dem spektralen Winkel basiert, eingesetzt und bewertet. Zwei Anwendungsschemata werden vorgestellt. Zuerst werden Hyperspektraldaten mit Laserscanningdaten fusioniert, wobei explizites Wissen über Dachgeometrien genutzt wird. Danach wird ein spektral-räumlicher Klassifizierungsansatz vorgestellt, welcher implizites Wissen über die Relevanz des räumlichen Kontextes in die Klassifizierung einbringt. Beide Ansätze sind vielversprechend, da sie höhere Klassifizierungsgenauigkeiten erzielen, wenn Wissen genutzt wird. Die Innovation des Beitrages ist, dass eine zweite Datenquelle über die Verknüpfung von Kernfunktionen kombiniert wird mit einer Modifikation der eins-gegen-eins Kaskade, die es erlaubt, Wissen zu integrieren.

1 Introduction

For many applications in remote sensing, it is beneficial to use multiple data sources (Amarsaikhan & Douglas 2004). By rendering classification results more reliably, a better interpretation of the resulting maps is ensured, which finally can help to make better decisions. Various approaches for multisource classification have been pub-

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0124 1432-8364/12/0124 $ 3.50

lished based on state-of-the-art classifiers like Markov random fields (Solberg et al. 1996), neural networks (Paola & Schowengerdt 1995), fuzzy classifiers (Binaghi et al. 1997), or combined classifiers (Huang & Lees 2004). Many approaches based on support vector machines (SVM) have been published as well. SVMs are a group of supervised learning methods that can be applied to classification or regression (Burges 1998). Watanachaturaporn et al. (2008) use an approach based on feature concatenation. Halldorsson et al. (2003) propose a multisource framework for SVM which modifies the RBF (radial basis function) kernel by using the distances of data points in various sources as features. Waske et al. (2007) classify two data sources separately. Then, the outputs of the SVM decision functions are used as new features which are again classified. Approaches for the fusion of hyperspectral with laserscanning data are presented by e.g. Jones et al. (2010) and Voss & Sugumaran (2008). A crucial prerequisite of all of these classification approaches is data fusion. For kernel based classifiers (like SVMs) a specialized method is offered by kernel composition. It exploits the fact, that kernel functions can be combined (e.g. by addition, multiplication, or weighting) to form new kernels. Thus, different kernels computed on different data sources can be combined. The first main focus of this contribution is to exemplify multisource classification by data fusion via kernel composition using SVMs. A second focus is also discussed. Adapting the SVM from binary to multiclass classification is frequently done via the one-against-one (OaO) cascade. Instead of considering all of the n classes in one step, the OaO cascade considers only two classes at a time. For each of the $\boldsymbol{n} \cdot (\boldsymbol{n} - 1)/2$ combinations of classes, an individual SVM is trained. Each SVM assigns a class label to each point. Afterwards, the final class label for each point is the label that has been most frequently assigned by the individual SVMs (Braun et al. 2010). Since the cascade considers only two of the n classes, it can be exploited to use additional sources of information selectively, i.e. only when the human operator considers it discriminative for the distinction of two particular classes. Thus, external knowledge on the classes available to the operator can be considered.

Both aspects are demonstrated and discussed on two application schemes. The first is multisource classification of hyperspectral and laserscanning data based on our previous work in Braun et al. (2011). The second is spectral-spatial classification on a hyperspectral benchmark dataset. Many researchers have made much effort on integrating the spatial context of pixel neighbourhoods into classification (Plaza et al. 2009). These spectral-spatial approaches slightly differ from the multisource case (fusion of hyperspectral and laserscanning data) since the spatial data source is derived from the spectral data source. However, an important issue which both have in common is the necessity to fuse two different sources of information. As a consequence, both multisource and spectral-spatial approaches need methodologies for data fusion. The main objectives of this paper are therefore data fusion and the way in which the OaO cascade can be used to integrate knowledge into classification. In the first application scheme, the knowledge is explicit knowledge of the operator about the shape of roofs. In the second scheme, it is implicit knowledge on the separability of classes. The OaO cascade allows using this knowledge in SVM classification. The remainder of the paper is organized as follows. Section 2 presents a mathematical introduction to the methods used. Section 3 outlines the data preprocessing, data fusion and classification of an urban dataset composed of hyperspectral and laserscanning data. In section 4, the transferability of the proposed framework will be shown for a spectral-spatial classification approach. In section 5, both classification approaches are discussed and compared in a synoptic manner, while section 6 concludes the paper.

2 Mathematical Foundations

This section provides briefly the mathematical foundations of the methods used. Kernel composition utilizes the Mercer property that kernels can be combined by addition, multiplication or ratio formation. SVMs ensure generalization by accepting a small amount of errors on the training data. This paper uses the ν-SVM by Schölkopf et al. (2000). The choice of data preprocessing and error handling strategy used herein is based on Braun et al. (2011), where a more thorough discussion on the different types can be found.

2.1 *Kernels and the Support Vector Machine*

Given a dataset X with n data points, kernel matrices $K(x_i,x_j)$ are the result of kernel functions applied over all n^2 pairs of data tuples. The outcome of a kernel function $K(x_i,x_j) = f\delta(x_i,x_j)$ is a similarity measure for the two training data x_i and x_j depending on some distance metric δ. φ is usuallythe Euclidean distance (Mercier & Lennon 2003). However, kernel functions can be modified e.g. by introducing different similarity measures (Amari & Wu 1999). To model complex distributions of the training data in the feature space, $f\varphi$ is usually some non-linear function. The most frequently applied family of non-linear functions are Gaussian radial basis functions (RBF) (Mercier & Lennon 2003, Amari & Wu 1999). The closer two points in the feature space are, the higher the resulting kernel value is. Given these facts, the kernel matrix simply represents the similarity between the points of the training dataset. To understand how the kernel matrix is used in SVM classification, it is helpful not to look at the primal, but the dual formulation of the SVM problem. The dual problem is given by (1).

$$\text{maximize: } \sum_i \lambda_i - \frac{1}{2}\sum_i \sum_j \lambda_i \lambda_j y_i y_j K(x_i, x_j) \qquad (1)$$

with

$\lambda_i > 0 \ \forall$ support vectors,
$\lambda_i = 0 \ \forall$ other points.

The Lagrange multipliers λ_i are only greater than zero for the support vectors (SVs) (see constraint of (1)). Hence, only training data which are SVs contribute to the solution of (1). For all cases where at least one of the points is not a support vector, $\lambda_i \cdot \lambda_j = 0$. Thus, the double sum term of (1) is also zero. Hence, if at least one of the points is not a support vector, the couple of points does not contribute to maximizing (1). The class labels y_i are in $\{-1,1\}$. Since the second part of (1) is subtracted, only points with different class labels can maximize the term. Since their product $y_i \cdot y_j = -1$, the double sum term is made positive and the double sum of this combination of points can contribute to maximizing (1). The problem is usually optimized by an optimization procedure. This procedure sequentially assigns λ_i to the training data. It converges if points are chosen as SVs which have different class labels but are found close to each other in the feature space. Such points will yield a high value in the kernel matrix $K(x_i,x_j)$. Their double sum of (1) will thus yield a high value and contribute strongly to maximizing (1). To conclude, it can be said that the similarity values of the kernel matrix are used for finding the best suited training points as SVs. By not letting non-SV points (which have $\lambda_i = 0$) influence (1), a sparse solution is found which only depends on the SVs (which have $\lambda_i > 0$).

2.2 *Spectral Angle Kernels*

In Braun et al. (2011) the linear kernel is used predominantly. It represents a dot product of the features and, thus, does not induce a high dimensional reproducing kernel Hilbert space (RKHS). However, the input feature space is already high-dimensional for hyperspectral datasets. Chances are that a good separation is achieved without transformation. Statistical learning theory teaches that unnecessarily complex functions raise the upper bound on the generalization error, a principle similar to Occam's razor. Therefore, it is beneficial to evaluate more simple functions – like the linear kernel – before trying more complex ones. Another advantage of using the linear kernel for kernel composition is its simple computation. Since the resulting kernel matrices are huge, computational load should be kept as low as possible. Good results were achieved by using this kernel function. On the other hand, the spectral angle (3) is considered a highly valuable measure of similarity for hyperspectral data due to their high dimensionality. One main advantage is that the spectral angle is insensitive with respect to differences in illumination. Thus, brighter and darker parts of the same material are more likely to be assigned to the same class if such a distance measure is used. Mercier & Lennon (2003) and Honeine & Richard (2010) propose a methodology for integrating the spectral angle into SVM classification. By simply replacing the $\|x_i - x_j\|$ in the RBF (2) with the spectral angle (3), they obtain a spectral kernel which combines the illumination insensitivity of the spectral angle distance with the discriminative power of kernel-based classification:

$$K_{RBF} = \exp\left(\frac{-\| x_i - x_j \|}{2\gamma}\right) \quad (2)$$

$$\alpha(x_i, x_j) = \arccos\left(\frac{x_i \cdot x_j}{\| x_i \| \times \| x_j \|}\right) \quad (3)$$

2.3 *Kernel Composition for Data Fusion*

For both application schemes presented in this paper, two information domains are available that have to be fused. Kernel matrices based on kernel functions are the representation of the similarity among input data and are used to identify the support vectors. According to Mercer's theorem, kernel matrices can be combined. Thus, after computing one kernel on the domain A and one kernel on the domain B, data fusion can be performed for instance by adding the kernels from A and B. Various methods for kernel composition are available (Camps-Valls et al. 2005). Herein, four composed kernels will be exemplified: the direct summation kernel (4), the weighted summation kernel (5), the product kernel (6) and the cross-information kernel (7). The direct summation kernel is a simple addition of two kernels. The weighted summation kernel introduces weighting factors τ_1 and τ_2. By setting these parameters, the user can define to which extent the information in each domain is considered as significant for the classification problem. For instance, if the user considers the first domain to be more relevant for the classification problem, he or she could set $\lambda_1 = 0.8$ and $\lambda_2 = 1 - \lambda_1$. Thus, the similarity of two features in the first domain influences their kernel values much more than their similarity in the second domain. The decision whether a point belongs to a certain class would therefore strongly depend on the first domain, while the second domain would influence it much less. Under these circumstances, the second domain may only be decisive if a point yields very comparable decision values for two classes based on the first domain. As mentioned above, this option is not available for concatenating features. (7) is called the cross-information kernel. It consist of four single kernels where the last two K_{AB} and K_{BA} allow the incorporation of the mutual information between the data sources A and B (e.g. differences between the values of both data sources for a particular data point). The most important advantage of kernel composition over concatenation is that the Hughes phenomenon is circumvented. The Hughes phenomenon is an accuracy decrease which occurs when the dimensionality of the feature space rises (Hughes 1968). Given a certain number of training samples, the predictive power of classification algorithms tends to decline as the dimensionality increases, e.g. because the feature space becomes more and more empty and statistics in the feature space cannot be estimated properly any more. For a thorough discussion see Landgrebe (1997). Concatenation fuses data in the feature space, thus raising dimensionality. Kernel composition fuses data in the RKHS. Since this space is high-dimensional (possibly infinite-dimensional), introducing extra features affects the solvability of the separation problem to a much smaller extent. On the other hand, it could be argued that concatenation has one advantage over kernel composition. If the data are not separable in all single kernel RKHS, there is a possibility that the separability in the composite RKHS is reduced, thus making concatenation the more favourable method in this case. For this reason, both approaches presented herein use kernel composition only if separability is assumed in the individual kernel spaces. If a certain information domain and its respective kernel are not considered representative for the separation, they are not used.

$$K_C(x_i^C, x_j^C) = K_A(x_i^A, x_j^A) + K_B(x_i^B, x_j^B) \quad (4)$$

$$K_C(x_i^C, x_j^C) = \mu K_A(x_i^A, x_j^A) + (1-\mu) K_B(x_i^B, x_j^B) \quad (5)$$

$$K_C(x_i^C, x_j^C) = K_A(x_i^A, x_j^A) \times K_B(x_i^B, x_j^B) \quad (6)$$

$$K_C(x_i^C, x_j^C) = K_A + K_B + K_{AB}(x_i^A, x_j^B) + K_{BA}(x_i^B, x_j^A) \quad (7)$$

2.4 *Exploitation of the one-against-one Cascade*

Since SVMs are binary classifiers, their original formulation is only able to solve classification problems distinguishing two classes. For our first application scheme, 14 subclasses need to be distinguished, though. Thus, the 14-class problem needs to be split up into various two class problems. The OaO strategy considers each of the 91 permutations of the classes separately – leading to 91 training and classification steps. For each step, a model is trained to separate a subset of two classes, considering e.g. the training pixels of class 6 and class 9. During classification, this SVM assigns either 6 or 9 as a label to each pixel. Each pixel is labelled by all 91 models and, thus, receives 91 labels, so that a 1×91 label vector v_i for each pixel is produced. The final class membership for the ith pixel is decided by $mode(v_i)$, i.e. the label most frequently assigned to the pixel. As mentioned above, two different information domains are available for each application scheme in this paper. Consider the urban image scenario in which different roof material classes are to be separated. While information of the airborne laserscanning (ALS) data can be helpful to distinguish between sloped brick roofs and flat gravel roofs, it will not facilitate the separation of two sloped roof classes significantly. At this point, the OaO cascade can be exploited. As it does not consider all 14 classes in one step, the cascade can be used to recognize whether the user considers geometry as significant for a given classification step or not. When separating two roof material classes with different geometries (sloped vs. flat) the spectral domain is fused with the geometric domain. In contrast, when separating similar roof geometries (sloped vs. sloped or flat vs. flat), the spectral domain is used exclusively. In this way, data known not to contain any new information can be omitted and the classifier training is focused on relevant information. Since this framework is very general, it can be applied to different kinds of data and different knowledge.

3 Fusion and Classification of Hyperspectral and ALS data

Within this section, improvements of our previous work on data fusion are presented. The application scheme is based on the fusion of hyperspectral and laserscanning data. The approach in Braun et al. (2011) is enhanced by using the spectral angle kernel for the hyperspectral data and a radial basis function kernel for the ALS data. Moreover, further kernel composition approaches are employed.

3.1 *Data Preprocessing*

An image from the city of Karlsruhe taken by the HyMap sensor in 2003 with a spatial resolution of *4×4 m²* and 126 spectral channels provides the hyperspectral information. A laserscan with *1×1 m²* resolution from 2002 delivers geometrical information. Apart from the first and last pulse information, the gradient and curvature of the first pulse are computed. This information is used to distinguish sloped roofs from flat roofs. Hence, six information channels are derived from the ALS data. The data of HyMap is resampled to the spatial resolution of the laser scan, using nearest neighbourhood interpolation. To reduce computational cost, and to allow for a later comparison with Lemp & Weidner (2004) and Braun et al. (2011), a *605×987* pixel subset is chosen which shows the campus of the Karlsruhe Institute of Technology (KIT). A z-transformation (i.e. normalization by mean and standard deviation) on each layer is performed to ensure comparability. The material classes to be distinguished are: brick (red in the classification results), copper (green), gravel (brown), slate (dark blue), zinc (light blue) and stone plates (grey). For each class, various training areas are defined. With respect to roof geometry, a sloped roof class and flat roof class are distinguished by integration of spatial information during the OaO cascade. However, the distinction is implicit in the sense that no classification map on sloped and flat roofs is produced. The relationship between the classes can be seen in Fig. 1. 200 points are randomly chosen from these areas for training and 100 for validation. Mean-shift segmentation is applied to the first pulse information. In order to be able to decide whether segments correspond to roofs or not, a 'roof mask' is calculated from laserscanning and hyperspectral data. First, a normalized digital surface model (nDSM) is computed by subtracting the terrain model from the surface model. All pixels having nDSM heights below two metres are considered to be non-roof pixels; this process will remove streets and other low man-made objects. Furthermore, the normalized difference vegetation index is computed on the hyperspectral data to mask out the remaining vegetated areas. Only segments being consistent with the roof mask thus created are considered in the subsequent processes. We start with a pixel-based classification. Afterwards, the results of the pixel-based classification are transferred to the roof segments. This is done by assigning each roof segment the class label most frequently assigned to the pixels it consists of.

3.2 *Selection of Single Kernels*

In a first step, we evaluated the capacities of single kernel functions on each domain. We trained a SVM using only one data domain to identify the best suited kernel for hyperspectral and ALS data separately. We evaluated the capacities of the linear, the polynomial, the radial basis function, the sigmoid and the spectral angle kernel by training, using five-fold cross validation to optimize ν and the respective kernel parameters (i.e. γ for RBF, sigmoid and spectral angle kernel, the degree for the polynomial and none for the linear kernel) simultaneously. Validation was performed on an independent set of validation data, entirely unknown to the classifier. The range of the

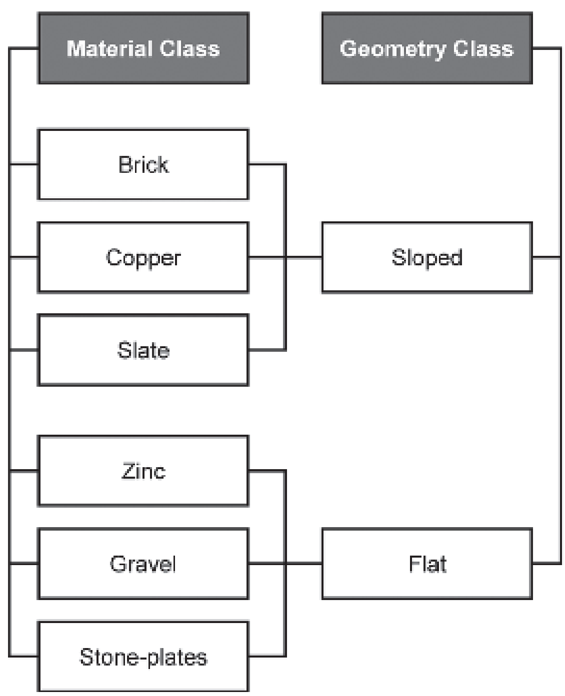


Fig. 1: Classes of the urban dataset and their relationship; left: material classes, right: geometry classes.

Tab. 1: Overall accuracy values using different kernel functions on a single data domain: urban dataset application scheme (RBF = radial basic function).

Kernel	Hyper-spectral	Laser-scanning
Linear	80.10 %	27.00 %
Polynomial	80.20 %	**55.90 %**
RBF	80.00 %	51.00 %
Sigmoid	53.90 %	25.30 %
Spectral angle	**84.80 %**	26.80 %

grid search for γ was $[2^{-15}, 2^{5}]$, ν was tuned in [0, 1], the polynomial degree was tuned in [1, 10]. For each tuning interval, 10 steps were evaluated (e.g. polynomial degree: 1,2,3,...,10).

As Tab. 1 reveals, the best suited kernel for the hyperspectral data is the spectral angle kernel. This finding empirically confirms the advantages which were theoretically described in section 2.2. The optimal kernel for laserscanning data is the polynomial kernel. The overall accuracy values yielded for the classification based on laserscanning data only were low, which is not surprising keeping in mind the features derived from laserscanning described above are merely enough to separate the spectrally different roof material classes. The sigmoid kernel performed poorly for both domains.

3.3 *Classification of Fused Dataset*

After identifying the most suitable single kernels, we performed kernel composition. Note that the cross-information kernel requires both input data domains to have the same dimensionality. Since the hyperspectral data consists of 126 features while the ALS data consists only of 6, this constraint is not fulfilled. Therefore, a principal component analysis (PCA) was computed on the hyperspectral data and only the first six principal components are used. The PCA is employed only for the cross-information kernel, because for the other kernel composition approaches the two data domains do not need to have the same dimensionality. In the case of the spectral angle kernel, feature reduction via PCA is not helpful. Since the spectral angle is especially designed to exploit the information of the numerous wavelengths in a hyperspectral dataset, it does not seem to be appropriate to reduce by discarding PCAs with lower eigenvalues. According to Braun et al. (2011), we performed data fusion by kernel composition only for the OaO steps where a sloped roof class is separated from a flat roof class. The selection scheme follows these short heuristics:

if geometry of roof_class 1 ≠ geometry of roof_class 2:

use hyperspectral and ALS data jointly via kernel composition

else: use hyperspectral data only via single kernel

After applying these heuristics, we checked for how many steps of the OaO cascade the second domain had been used. It had been used for 40 out of the 91 OaO steps (i.e. in ~44 % of the steps). As Tab. 2 reveals, all data fusion approaches yielded higher overall accuracy values than did the best single kernel. The best approach is the direct summation kernel which yielded 86.9 % overall accuracy, thus improving the best result published in former work by 0.3 %. An attempt to use both data sources at each step yielded 79.6 % overall accuracy, thus validating the assumption that selective usage is beneficial. These findings confirm the conclusions on selective usage thoroughly discussed in Braun et al. (2011). Running the algorithm concatenating the features yielded 84.1 % overall accuracy. The classification results of using only the hyperspectral domain are given in Fig. 2

Tab. 2: Overall accuracy values of different fusion approaches for urban dataset application scheme.

Classification approach	Overall accuracy
Only hyperspectral domain	84.80 %
Concatenation	84.10 %
Direct summation (selective) $\tau_1=1, \tau_2=1.$	**86.90 %**
Weighted summation (selective) $\tau_1=0.8, \tau_2=0.2.$	85.50 %
Cross-information kernel (selective)	85.60 %
Product kernel (non selective)	85.90 %

(3) and the result of the direct summation kernel is shown in Fig. 2 (4). Due to the quite high classification accuracies, both results are visually similar in their major part. Fig. 2 (2) visualizes the differences between the two results. The result based on hyperspectral information contains many roofs classified as stone-plate roofs (grey) or zinc roofs (bright blue) which in reality belong to other classes – mostly gravel (brown). The method based on the direct summation kernel of hyperspectral and ALS data classified the circular building in the north-east as brick (red), although it is a zinc roof. A McNemar test (Foody 2004) was used to check the significance of the improvement. The value of the McNemar test $|z|$ indicates the significance of differences between classification results. A difference is considered as significant if $|z| \geq 1.96$. The results based solely on hyperspectral data were tested against the proposed selective kernel composition approach using the McNemar test. The test yielded a value of $|z| = 18.9$. Thus, the advantage of the proposed selective kernel composition approach is highly significant.

In order to assess the performance of the two approaches to individual classes, producer's and user's accuracies were compared. As Tab. 3 reveals, the laserscanning data is helpful for the producer accuracies of gravel and slate. For copper and stone-plates, no differences were found, for brick and zinc, the laserscanning data impairs the producer accuracy slightly. The average improvement over all six classes is 1.81 percent points. The producer's accuracies of some classes of both sloped and flat roofs were improved, others slightly impaired. Hence, no clear trend with respect to roof geometry can be observed. Results for the user's accuracy are similar. Brick, zinc and stone-plates were improved by laserscanning, the copper class was not changed and gravel

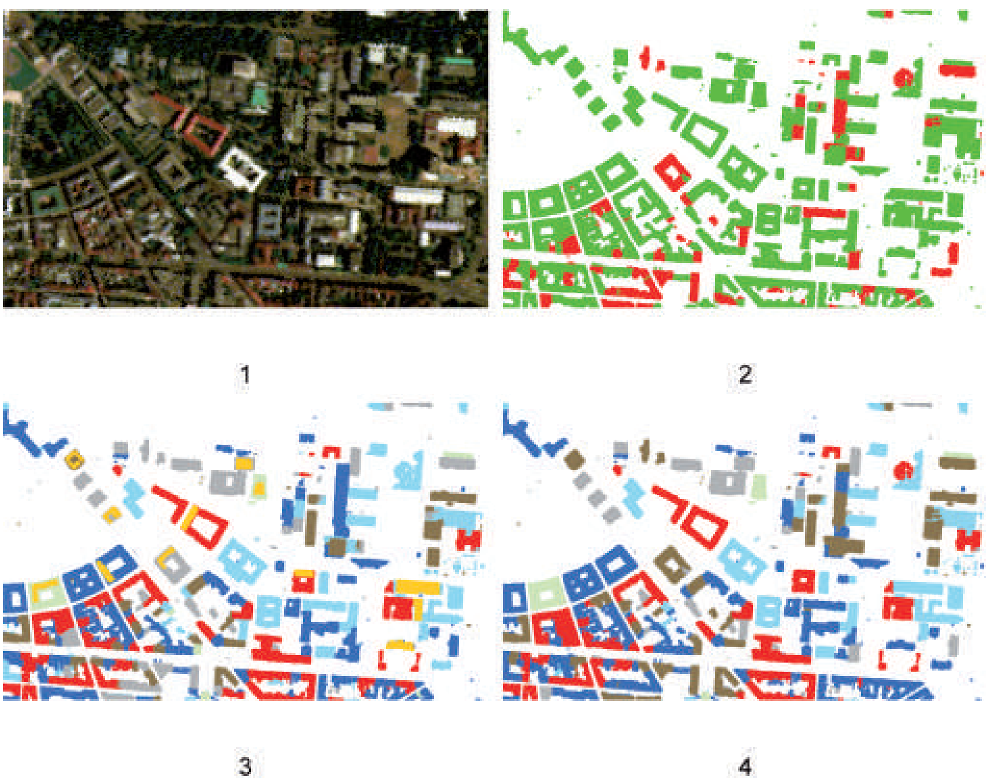


Fig. 2: Urban dataset, 1: True colour view of the hyperspectral data, 2: Agreement (green: equal labels, red: different labels), 3: Classification based on hyperspectral information only (orange fills: training areas), 4: Proposed direct summation approach (fusion of hyperspectral and ALS data).

Tab. 3: Producer's (PA) and user's (UA) values and their differences for the two approaches (HYP = hyperspectral data, LS = laserscanning data).

Class	PA (HYP/LS)	PA (HYP)	δ(PA)	UA (HYP/LS)	UA (HYP)	δ(UA)
Brick	85.17	87.70	-2.53	100.00	94.86	**5.14**
Copper	87.28	87.28	0.00	100.00	100.00	0.00
Gravel	100.00	89.73	**10.27**	62.30	66.69	-4.39
Slate	85.00	77.72	**7.28**	81.28	83.74	-2.46
Zinc	87.56	91.67	-4.11	97.38	88.77	**8.62**
Stone-plates	74.61	74.61	0.00	76.81	64.79	**12.02**
			Ø=1.81			Ø=3.15

and slate were slightly impaired. Again, no clear trend with respect to roof shape can be observed. The average improvement is 3.14 percent points.

4 Fusion and Classification of Spectral Spatial data

To show the transferability of the methodology developed above to a different application scheme, a similar framework of a different dataset is developed. This approach is based on a spectral-spatial classification of the well-known AVIRIS Indian Pines dataset provided in the MultiSpec toolbox (Biehl & Landgrebe 2002). The Indian Pines dataset consists of a *145x145* pixel image of an agricultural area taken by the AVIRIS sensor. This dataset is known as a difficult benchmark for hyperspectral classification due to the high spectral similarity of the classes included (different crops at an early phenological stage). It has 220 spectral bands and includes 16 land use classes. Many bands are quite noisy so some authors remove some bands from the dataset before classification. The dataset comes with a ground truth image in which around 49 % of the pixels are labelled. Nonetheless, some classes consist of very few labelled points (e.g. only 20 points). For this contribution, all 16 classes were used and no bands were removed in order to maintain a high dimensionality and the noise level typical for many hyperspectral sensors. The ground truth data were randomly split into training data (20 %) and evaluation data (80 %).

The spatial context is produced by a simple *3x3* median filter on each channel. The resulting feature vector is then fused with the spectral data. For each class, 20 % of the data labelled by ground truth are used for training and the rest for classification. Selective data fusion via kernel composition of spectral and spatial data by exploiting the OaO cascade is used. However, unlike in section 3, no explicit knowledge is available on whether to use spectral information for a certain classification step or not. Thus, this knowledge is created implicitly after two basic classifications. Again, we evaluated the capacities of the single kernels on the two data domains. The tuning of hyperparameters was performed identically to section 3.2. The polynomial kernel is identified to be ideal for the spectral domain and the RBF kernel to be ideal for the spatial domain (Tab. 4). The overall accuracies yielded were higher on the spatial domain for each kernel except for the sigmoid kernel whose performance is again poor. We believe that

Tab. 4: Overall accuracy values using different kernel functions on a single data domain: spectral-spatial application scheme.

Kernel	Spectral domain	Spatial domain
Linear	79.50 %	87.80 %
Polynomial	**81.10 %**	89.40 %
RBF	79.80 %	**90.10 %**
Sigmoid	38.70 %	38.40 %
Spectral angle	73.80 %	80.20 %

this is due to the lack of preprocessing steps of the spectral domain, which therefore is classified with all the noise it contains. For this reason, we decided to use the features derived by median filtering as our main domain. Unlike for the first classification approach, the spectral angle kernel did not produce the best results on the spectral domain. Having identified the best kernels, we classified the AVIRIS dataset using only the spectral information of the original channels first. Afterwards, a confusion matrix for the result based on spectral data only is computed (it is called **CMsc**). In a second step, we classified, but using the spectral and spatial information jointly at each step of the OaO cascade. Again, a confusion matrix based on the joint usage of spectral and spatial data is computed (it is called **CMscsp**). From these two confusion matrices, the knowledge of whether to use the spatial information is created by simple and straightforward heuristics:

if CMscsp(i,j) > CMsp(i,j) or CMscsp(j,i) > CMsp(j,i) with $i \neq j$:
use spectral and spatial information jointly via kernel composition
else: use spatial information only via single kernel

Given 16 classes, the OaO cascade consists of 120 classification steps. Applying this heuristic, the spectral information proved relevant for 49 steps. A third classification was performed, using the spectral and spatial information jointly for these 49 steps only. Various kernel composition approaches were evaluat-

Tab. 5: Overall accuracy values of different fusion approaches for spectral-spatial application scheme.

Classification approach	Overall accuracy
Only spatial domain	90.10 %
Direct summation (non selective) $\tau_1=1,\ \tau_2=1.$	83.70 %
Weighted summation (non selective) $\tau_1=0.6,\ \tau_2=0.4.$	84.00 %
Cross-information (non selective)	74.60 %
Product kernel (non selective)	90.70 %
Product kernel (selective)	**91.60 %**

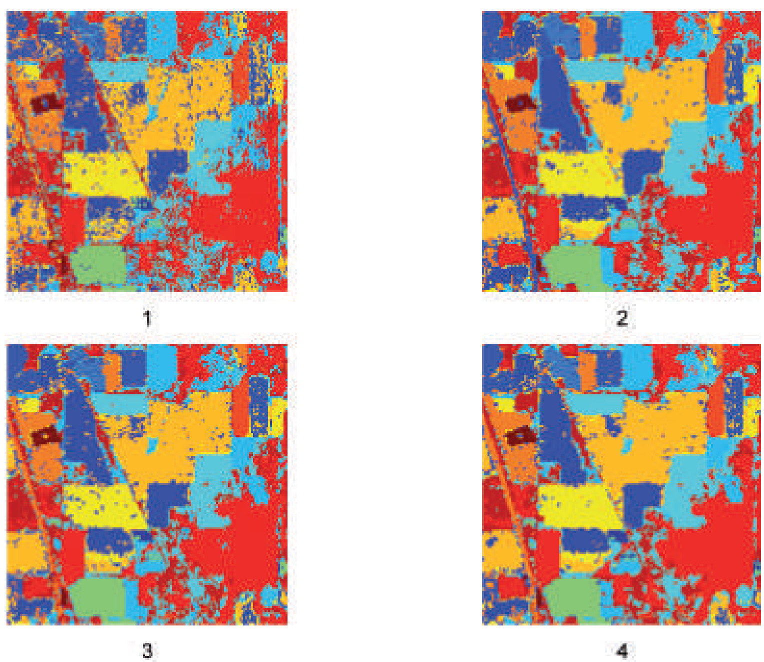


Fig. 3: Result for AVIRIS, 1: Spectral data, 2: Spatial data, 3: Spectral, spatial data fused at each step, 4: Spectral, spatial data with selective kernel composition.

ed. As can be seen in Tab. 5 all except one approach which use the two domains non-selectively produce worse results than using only the median filtering (spatial). The cross-information kernel performs worst of all approaches, confirming the findings of Camps-Valls et al. (2005, 2008) who state that more sophisticated kernels tend to produce worse results. Again, this finding underlines the care that should be taken when fusing data. However, the non-selective product kernel produces a slightly higher overall accuracy than using the approach based on one domain only. As can be seen, spectral information is helpful for most, but not for all the classes. Finally, when using a product kernel only for the 49 steps described above, an overall accuracy value of 91.6 % is obtained, which is the best value for all approaches presented for this dataset. The McNemar test is used to check the significance of the improvements yielded. The value for the test statistic when testing the selective product kernel against the single kernel (only spatial domain) is 2.9, thus, the improvement is significant. When testing the selective product kernel against the non-selective product kernel, a value for the test statistic of 2.7 is obtained, indicating that the improvement caused by selective usage was significant as well.

Fig. 3 shows a comparison of classification results. Fig. 3 (1) shows the result for a classification based on the spectral data only and Fig. 3 (2) shows the result obtained by using the spatial data. Fig. 3 (3) shows the result obtained by fusing spectral and spatial information at each OaO step. Fig. 3 (4) shows the result produced by selectively integrating the spatial information applying the heuristics given above. Note that the images based on using spatial context appear to be much smoother while the borders of different crop fields are preserved. The results are much less affected by salt-and-pepper classification noise. While using spatial information at each step already seems to raise classification performance, the selective usage of the spatial context seems to be even more suitable. The right image is even less affected by salt-and-pepper noise and the field borders are more clearly outlined (southeast quarter of the image).

5 Discussion

The results presented in Braun et al. (2011) could be enhanced by various means. First of all, the spectral angle kernel computed on the hyperspectral data improved the performance in terms of the best accuracy achieved by 4.7 % compared to the previous approach. As the spectral angle is a very well suited measure of similarity for hyperspectral data its integration into SVM combines the advantages of traditional classification approaches for hyperspectral data (the spectral angle mapper) with the advantages of kernel learning. Our results indicate that the spectral angle kernel should be taken into account for hyperspectral classification approaches. Rong et al. (2006) provide an insightful overview on selection of kernel functions. However, their paper does not include the spectral angle kernel. Fauvel et al. (2006) present a comparison taking into account the spectral angle kernel. Their results indicate a slight advantage of RBF over the spectral angle kernel. However, their dataset requires the distinction of classes which mainly differ in intensity (trees, meadows). Furthermore, the best classification accuracy achieved by an approach based on selective data fusion (Braun et al. 2011) could be raised to 86.9 % in this paper. Individual values of producer's and user's accuracies were improved for some classes and impaired for others. However, both producer's and user's accuracy were improved on average. No clear trend between improvements and roof geometries could be observed. This was to be expected, because each class had to be distinguished from classes having both the same and different roof shapes. The highest classification accuracy is achieved by selective data fusion based on the direct summation kernel. Just as Camps-Valls et al. (2005, 2008) we found that simple kernels tend to produce better results than more sophisticated ones. Results indicate that the single kernel type should be selected individually for each input domain. Although it is also possible to use the same kernel type for both domains, higher accuracy values and a more flexible approach are achieved by an individual selection. In our case, the spectral angle kernel and the polynomial kernel proved to be best suited for the fusion of hyperspectral and ALS data while the polynomial and the RBF kernel were optimal for the spectral-spatial approach. The result fails to outperform the overall accuracy (over 90 %) yielded by BÄHR et al. (2005) on the same dataset. However, he and his group used a rule-based object-oriented approach. In such a design, knowledge on the classification problem can be integrated in a much more flexible manner than in the machine learning approach implemented herein. Furthermore, rather complex knowledge can be easily integrated using more complex rules (e.g. combinations of thresholds). These factors seem to explain the advantage in overall accuracy by BÄHR et al. (2005) over the machine learning approach yielded herein. On the other hand, the object-oriented approach may require a thorough redesign of rules when applied to a different image scene. In contrast, the machine learning approach requires defining new training areas at most. Thus, despite its lower overall accuracy, it may be a valuable alternative to object-oriented designs since it can be applied to new images more easily.

To prove the transferability of our methodology to a different problem and to work towards a more general framework, we have presented a second application scheme. A spectral-spatial classification of AVIRIS Indian Pines reveals similar tendencies as the fusion of hyperspectral and laserscanning data. Of course, in this case the two domains are much more correlated than for the hyperspectral – ALS dataset. The median filter on the hyperspectral channels produces more accurate results than the original channels. Given the considerable noise level and the low separability of the classes in this dataset, smoothing the image by median filtering produces more robust results. The product kernel performs slightly better when applied in a non-selective framework and considerably better when applied in a selective manner. According to the McNemar test, these advantages are significant. The images appear visually much smoother and less affected by salt-and-pepper noise. In contrast to the hyperspectral-ALS approach, implicit knowledge is generated by comparing confusion matrices. These findings show the transferability of our framework to differently posed problems and constitutes an important step towards a general framework for knowledge integration. The results confirm that the approach is applicable to different data types and different types of knowledge. Using the second data domain in a naïve way does not necessarily increase classification accuracy; as some fusion approaches perform worse than the approach based on a single domain only, thorough consideration needs to be made not only on how to fuse the data but also on when to use the second information domain.

6 Conclusion

Kernel composition proves to be a potent method for data fusion when using SVM. Best results are achieved with the direct summation and the product kernel. The OaO cascade is a useful way to integrate the second information domain selectively. Knowledge is exploited to use information selectively and proves suitable to enhance classification results for both application schemes described in this paper. The spectral kernel is used for hyperspectral data since it integrates a well-established measure of similarity into SVM classification. Furthermore, an approach for spectral-spatial classification is presented for the well-known AVIRIS Indian Pines datasets. Spatial context is used for classification. Implicit knowledge on the relevance of the spatial context is integrated. The transfer from fusion of hyperspectral and ALS data based on given explicit knowledge to spectral-spatial classification based on implicit knowledge highlights the transferability of the proposed framework.

References

AMARI, S. & WU, S., 1999: Improving support vector machine classifiers by modifying kernel functions. – Neural Computation **12:** 783–789.

AMARSAIKHAN, D. & DOUGLAS, T., 2004: Data fusion and multisource image classification. – International Journal of Remote Sensing **25** (17): 3529–3539.

BÄHR, H.-P., LEMP, D. & WEIDNER, U., 2005: Hyperspectral Meets Laserscanning: Image Analysis of Roof Surfaces. – ISPRS Workshop High Resolution Earth Imaging for Geospatial Information **XXXVI** (I/W3): CD-ROM.

Biehl, L. & Landgrebe, D., 2002: MultiSpec – a tool for multispectral-hyperspectral image data analysis. – Computers and Geoscience **28** (10): 1153–1159.

Binaghi, E., Madella, P., Grazia Montesano, M. & Rampini, A., 1997: Fuzzy contextual classification of multisource remote sensing images. – IEEE Transactions on Geoscience and Remote Sensing **35** (2): 326–340.

Braun, A.C., Weidner, U. & Hinz, S., 2010: Support Vector Machines for Vegetation Classification – A Revision. – PFG **2010** (4): 273–281.

Braun, A.C., Weidner, U., Jutzi, B. & Hinz, S., 2011: Integrating model knowledge into SVM classification – Fusing hyperspectral and laserscanning data by kernel composition. – Heipke, C., Jacobsen, K., Rottensteiner, F., Müller, S. & Sörgel, U. (eds.): High-resolution earth imaging for geospatial information. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **38** (4/W19): CD-ROM.

Burges, C.J.C., 1998: A Tutorial on Support Vector Machines for Pattern Recognition. – Data Mining and Knowledge Discovery **2** (2): 121–167.

Camps-Valls, G., Gomez-Chova, L., Muñoz-Marí, J., Vila-Francés, J. & Calpe-Maravilla, J., 2005: Composite Kernels for Hyperspectral Image Classification. – IEEE Geoscience and Remote Sensing Letters **3** (1): 93–97.

Camps-Valls, G., Gomez-Chova, L., Muñoz-Marí, J., Rojo-Alvarez, J.L. & Martinez-Ramon, M., 2008: Kernel-Based Framework for Multitemporal and Multisource Remote Sensing Data Classification and Change Detection. – IEEE Transactions on Geoscience and Remote Sensing **46** (6): 1822–1835.

Fauvel, M., Chanussot, J. & Benediktsson, J.A., 2006: Evaluation of kernels for multiclass classification of hyperspectral remote sensing data. – IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) **2006**.

Foody, G.M., 2004: Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. – Photogrammetric Engineering and Remote Sensing **70** (5): 627–634.

Halldorsson, G.H., Benediktsson, J.A. & Sveinsson, J.R., 2003: Support vector machines in multisource classification. – IEEE International Geoscience and Remote Sensing Symposium (IGARSS) **2003**.

Honeine, P. & Richard, C., 2010: The angular kernel in machine learning for hyperspectral data classification. – 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS) **2010**.

Huang, Z. & Lees, B.G., 2004: Combining nonparametric models for multisource predictive forest mapping. – Photogrammetric Engineering and Remote Sensing **70** (4): 415–426.

Hughes, G., 1968: On the mean accuracy of statistical pattern recognizers. – IEEE Transactions on Information Theory **14** (1): 55–63.

Jones, T.G., Coops, N.C. & Sharma, T., 2010: Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. – Remote Sensing of Environment **114** (5): 2841–2852.

Landgrebe, D., 1997: On information extraction principles for hyperspectral data. – Purdue University, West Lafayette, IN, USA.

Lemp, D. & Weidner, U., 2004: Use of Hyperspectral and Laser Scanning Data for the Characterization of Surfaces in Urban Areas. – IAPRSIS **XXXV** (B/7): CD-ROM.

Mercer, J., 1909: Functions of positive and negative type, and their connection with the theory of integral equations. – Philosophical Transactions of the Royal Society of London, Series A, Containing Papers of a Mathematical or Physical Character **209:** 415–446.

Mercier, G. & Lennon, M., 2003: Support Vector Machines for Hyperspectral Image Classification with Spectral-based kernels. –IEEE International Geoscience and Remote Sensing Symposium (IGARSS'03) **2003**.

Paola, J.D. & Schowengerdt, R.A., 1995: A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery. – International Journal of Remote Sensing **16** (16): 3033–3058.

Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., Marconcini, M., Tilton, J.C. & Trianni, G., 2009: Recent advances in techniques for hyperspectral image processing. – Remote Sensing of Environment **113:** 110–122.

Rong, H., Zhang, G. & Jin, W., 2006: Selection of Kernel Functions and Parameters for Support Vector Machines in System Identification. – Journal of System Simulation **11:** 179–189.

Schölkopf, B., Smola, A., Williamson, R. & Bartlett, P., 2000: New Support Vector Algorithms. – Neural Computation **12:** 1207–1245.

Solberg, A.H.S., Taxt, T. & Jain, A.K., 1996: A Markov random field model for classification of multisource satellite imagery. – IEEE Transactions on Geoscience and Remote Sensing **34** (1): 100–113.

Voss, M. & Sugumaran, R., 2008: Seasonal effect on tree species classification in an urban environment using hyperspectral data, LiDAR, and an object-oriented approach. – Sensors **45** (12): 3020–3026.

Waske, B. & Benediktsson, J.A., 2007: Fusion of support vector machines for classification of multisensor data. – IEEE Transactions on Geoscience and Remote Sensing **8** (5): 3858–3866.

Watanachaturaporn, P., Arora, M.K. & Varshney, P.K., 2008: Multisource classification using support vector machines: An empirical comparison with decision tree and neural network classifiers. – Photogrammetric Engineering and Remote Sensing **74** (2): 239–246.

Address of the Authors:

Andreas Ch. Braun, Uwe Weidner, Boris Jutzi & Stefan Hinz, Karlsruher Institut für Technologie – KIT, Institut für Photogrammetrie und Fernerkundung, D-76131 Karlsruhe, Tel.: +49-721-608-4-2315, Fax: +49-721-608-4-8450, e-mail: {andreas.ch.braun} {uwe.weidner} {boris.jutzi} {stefan.hinz}@kit.edu

Manuskript eingereicht: Oktober 2011
Angenommen: März 2012

PFG 2012 / 4, 0385–0394
Stuttgart, August 2012 **Article**

Scientist's Idealism versus User's Realism on Radarsat-2 HR Stereo Capability without GCP: Two Cases over North and Arctic Sites in Canada

Thierry Toutin, Khalid Omari, Enrique Blondel, Ottawa, Daniel Clavet, Sherbrooke & Carla Vanessa Schmitt, Ottawa, Canada

Keywords: Radarsat-2, SAR, high-resolution, radargrammetry, DSM, Canadian Arctic

Summary: Digital surface models (DSMs) are extracted from high-resolution Radarsat-2 (R2) stereo images using our new hybrid radargrammetric modeling without ground control point developed at the Canada Centre for Remote Sensing. They are then evaluated over two Canadian northern and Arctic study sites: the first for the scientific validation and the second in the Arctic (steep relief and glaciated surfaces) for the operational evaluation. For the validation site, the bias and elevation linear errors with 68 percent confidence level (LE68) of R2 stereo-extracted DSM compared to lidar data were computed over bare surfaces: LE68 of 3.7 m and no bias were achieved. For the Arctic study site, a large negative bias of 18–20 m and LE68 of 21–30 m were computed versus ICESat data over and outside ice fields, respectively. In addition LE68 of 15 m with equivalent bias was obtained over ice fields with 0°–5° slopes, which generally occurred in ice fields. The large biases certainly suggest a bias in the R2 stereo-model and thus in the metadata used in our hybrid model computation.

Zusammenfassung: *Der Idealismus des Wissenschaftlers und die Realität des Nutzers hinsichtlich der Möglichkeit von Radarsat-2 zu hochaufgelösten Stereoaufnahmen ohne Passpunkte: Zwei Fallbeispiele für nördliche und arktische Gebiete in Kanada*. Digitale Oberflächenmodelle (DOM), die von hochaufgelösten Radarsat-2 (R2) Stereobildern mit Hilfe eines neuen am Kanadischen Fernerkundungszentrum entwickelten hybriden Radargrammetrie-Modells abgeleitet wurden, welches ohne Passpunkte auskommt, werden anhand von zwei Testgebieten in nördlichen bzw. arktischen Regionen Kanadas evaluiert. Die Evaluierung im ersten Testgebiet dient der wissenschaftlichen Validierung, während jene im arktischen Testgebiet (steiles Gelände, Eis als Untergrund) der Untersuchung operationeller Gesichtspunkte dient. Für die Validierung wurden im Gebiet der systematische Fehler und die linearen Höhenfehler des DOMs aus R2-Daten mit dem DOM aus Laserscanning-Daten auf vegetationslosem Boden verglichen und das 68 % Konfidenzniveau (LE68) bestimmt: LE68 betrug 3.7 m. Ein systematischer Fehler konnte nicht festgestellt werden. Für das arktische Testgebiet ergab sich an bzw. außerhalb von Eisfeldern ein großer negativer systematischer Fehler von 18–20 m und ein Wert des LE68 von 21–30 m im Vergleich zu den Daten von ICESat. Darüber hinaus ergaben sich ein Wert für LE68 von 15 m und ein äquivalenter systematischer Fehler über Eisfeldern, die mit 0°–5° geneigt waren, was eine typische Neigung von Eisfeldern darstellt. Die großen systematischen Fehler legen einen Zusammenhang mit systematischen Fehlern im R2 Stereomodell und somit in den Metadaten, die in unserem hybriden Modell verwendet wurden, nahe.

DOI: 10.1127/1432-8364/2012/0125 1432-8364/12/0125 $ 2.50

1 Introduction

Because Canadian Arctic suffers from poorly-known relief and the surface state of glaciated regions is rapidly evolving due to snowfall, snow transport by wind and/or surface melt, remote sensing data (synthetic aperture radar (SAR) or optical) used to retrieve the topography must be acquired with the shortest possible time interval to maximize their coherence or correlation. This is also important because the flow of the glacier (up to a few metres per day) can bias the topographic measurement during this time interval.

Stereo radargrammetry using SAR data, as developed since 1960's by La Prade (1963), Rosenfield (1968), La Prade & Leonardo (1969), Leberl (1972, 1978) and many other afterwards, can thus be an appropriate solution in Canadian Arctic, even if there was no previous radargrammetric experiment over glacierized region, to the best of our knowledge (Toutin 2011). Because the state-of-the-art of radargrammetry and DSM generation using SAR data can be found in many details in Leberl (1990) and Toutin & Gray (2000), it is thus more important to focus on the winning conditions of stereo-radargrammetry in the challenging conditions of a glacierized region. The multi-date across-track capability is not so problematic due to different SAR advantages specific to ice regions. First, the backscatter of SAR sensors is more dependent on the rugosity or the dielectric component, which enable more radiometric contrast and less depth penetration over ice fields covered with supraglacial debris, rock glaciers, moraines, etc. Second, the SAR sensors are operated in all-weather conditions and not dependent on the solar illumination conditions, which thus cancelled the large shadowed areas with optical data in high latitudes. Third, the convergence of heliosynchronous orbits to North/South poles combined with a large range of look angles (20°–60°) gives thus a strong advantage to drastically reduce the temporal variations to few days between the multi-date stereo-images acquired in the highest latitudes. Last but not least, the new satellite SAR sensors have now high-resolution (HR) capabilities (sub to few metres), and are dedicated toward operational applications.

In summary, the choice of SAR stereo-images should resolve the following issues:

- to have an operational SAR system for acquiring the planned data from shallow look angles at the best season corresponding to the end of the melt season with no snow or little soaked snow and supraglacial debris due to freezing/melt process for soaked ice/snow,
- to reduce temporal resolution and then temporal radiometric changes with the orbit convergence and a larger choice of appropriate look angles (from 35°– 50°),
- to reduce illumination and radiometric differences with images having close look angles,
- to increase the SAR backscatter on ice with shallow look angles due to higher surface roughness with supraglacial debris after the melt season,
- to increase the matching performance and reduce the mismatched areas due to less radiometric variations with small intersection angle, and
- to reduce the SAR depth penetration over the "wet or soaked" 1st-ice layer covered by supraglacial debris with shallow look angles (Dall et al. 2001).

Due to the remote and harsh environments of the Canadian Arctic, the 3D geometric processing of SAR images should require no ground control points (GCPs) collected by users for the operational applications. A new hybrid radargrammetric model (Toutin & Omari 2011) was recently developed for Radarsat-2 (R2) at the Canada Centre for Remote Sensing and thus used for the stereo-modeling and the generation of digital surface models (DSM) without GCP. This hybrid model combine the deterministic Toutin's model (Toutin & Chénier 2009), the rational function model (RFM) empirical model and the Radarsat-2 associated metadata. In order to first evaluate the performance of the stereo-radargrammetric process in a well-controlled research environment (scientist's idealism) and after in an operational environment (user's realism), two Canadian northern and Arctic study sites were used: the first one for the scientific validation and the second one for the operational evaluation in Arctic. This last experiment was the first application of stereo-radargrammetry on glacierized region.

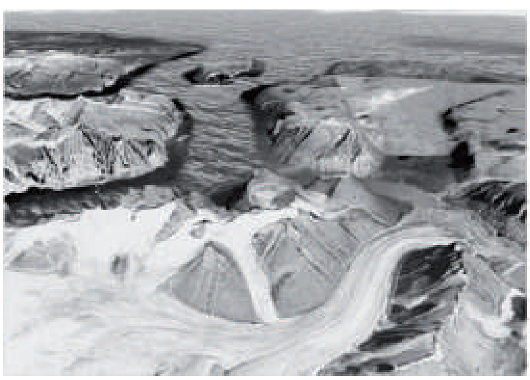
2 Field Study Sites

The first study site (validation site hereafter) is located north of Québec City, Québec, Canada and spans different environments: urban and residential, semi-rural and forested (Fig. 1).

Fig. 1: Radarset-2 (R2) ultra-fine mode (U2) image of the 1st study site. "Radarsat-2 Data ©MacDonald, Dettwiler and Associates Ltd. (2008) – All Rights Reserved" and Courtesy of CSA.

The elevation ranges almost from 10 m in the city in the southeast to around 1000 m in the Canadian Shield in the north. The northern part is a hilly to mountainous topography (5°–30° slopes) mainly covered with forests (deciduous, conifer and mixed) while the south part is a semi-flat topography (0°–5° slopes) with urban and residential areas. This site is to validate the quality output products in a well-controlled research environment: the scientist's idealism.

The second study site (operational site hereafter) located in the Baffin Island, Nunavut at approximately 70° 50' N and 71° 30' W (Fig. 2, left). There is no vegetation cover, except small plants. Even if there is some residual snow in the highest altitude on the Landsat ortho-image acquired at the beginning of the melt season in August (Fig. 2, right), the ice fields are similar to the exposed ice surfaces of Jakobshavn Isbrae, west Greenland, where a depth penetration of 1±2 m with C-band radar was found (Rignot et al. 2001). More than 80 % is covered by ice fields with cirque glaciers (permanent ice-covered mountains), outlets and valley glaciers and glaciers tongues surrounded by spectacular fjords with 70°–90° cliffs of 500–800 m height. Bare surface



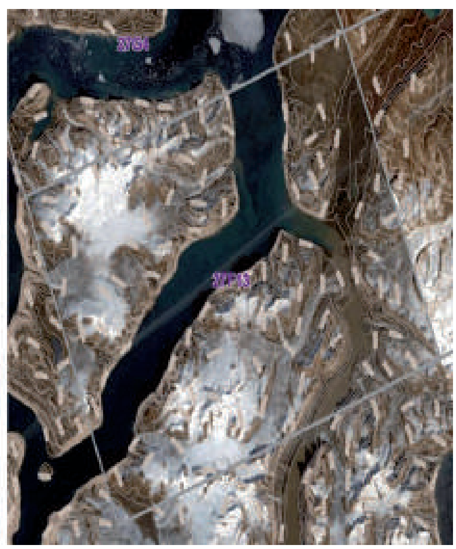


Fig. 2: Left: Perspective view from south to north of the second study site, generated with Google Earth using images from TerraMetrics and WorldView. © 2010 Google and Images ©2010 TerraMetrics and DigitalGlobe. Right: Ortho-rectified Landsat-7 with 1:50,000 map sheet grid. Some residual snow only appears in the highest altitudes at the beginning of the melt season.

mountains also with steep slopes surround the ice fields and glaciers. The valley glacier in the south-east is about 1-km wide with up-to-4° slopes surrounded by 600 m height bare surface and cirque glaciers. The elevation ranges from sea level to 1840 m and the slopes vary from 0°– 90° at fjord cliffs, illustrating a very challenging environment (in terms of land cover and relief). This site is to validate the quality output products in operational mapping environment: the user's realism.

3 Data Acquisition and Processing

For the validation site, the R2 SAR dataset included two stereo images (20 km by 20 km) acquired September 10 and 14, 2008 with the C-band ultra-fine (U) mode (1 by 1 look; 1.6–2.4 m by 3 m resolution) in VV polarization from descending orbits with view angles of 30.8°–32° (U2 Fig. 1) and 47.5°–48.3° (U25) at the near-far edges, respectively.

The reference cartographic data included ground points, mainly road intersections and electrical poles, collected from a differential global positioning system (DGPS) survey in November 2008 with 3-D ground accuracy of 10–20 cm. The collected points were used as independent check points (ICPs) to quantify/validate our new hybrid model accuracy. In addition, DSM (20 km by 20 km, 1 m spacing) were obtained from a first-echoed return lidar survey collected by GPR Consultants in October 2009 (Fig. 3), and covered the full R2 stereo pair. When compared to the previous DGPS, the positioning and altimetric accuracy (1 σ) of the lidar was computed to be better than 0.3 m and 0.1 m, respectively.

For the operational site, the R2 stereo images were acquired in 2009 from descending orbit with the ultra-fine mode (U mode with 3 m resolution) in HH polarization: U12 (Fig. 4) and U26 (20 km by 20 km; 1.6 m by 2.5–3.0 m pixel spacing) on September 28 with 38.83°–39.84° right look angles and September 9, 2009 with 48.12°–48.93° look angles, respectively. This acquisition period corresponded (i) to the end of the melt season and (ii) to -3°/+3° temperature range generating a freezing/melt process during the night/day, respectively. It results, especially with shallow angles that the 1st layer of the exposed "wet" ice fields covered by supraglacial debris and of the remaining soaked snow in the highest altitudes reduced the SAR depth penetration to its minimum (less than 1 m) due to a high water content and dielectric component (Dall et al. 2001, Gray 2011). The radiometry of these stereo data are, however, dominated by the geometric issues due to high relief with no

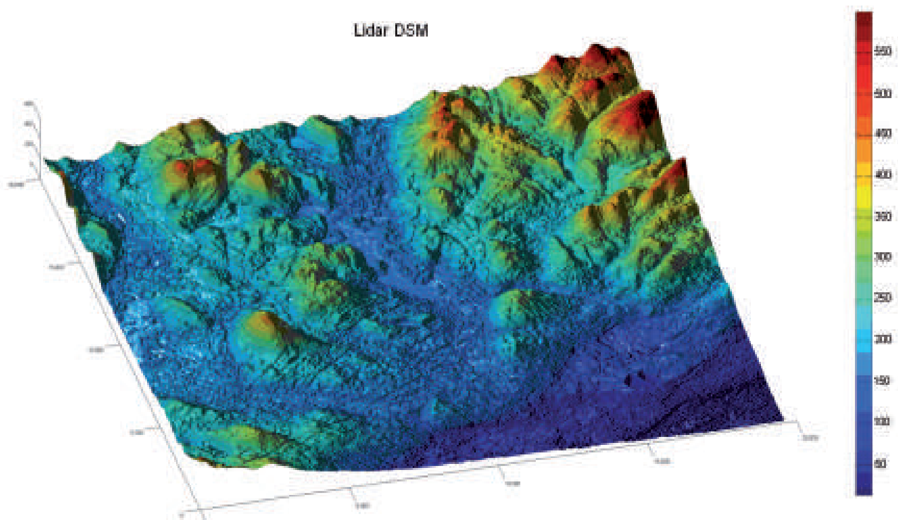


Fig. 3: Digital surface model (20 km by 20 km, 1-m spacing) from a lidar survey.

vegetation cover: more severe layover in U12 over the east-oriented slopes and more occluded areas over the west-oriented slopes in U26. In fact, the south-north curved land-water boundary of the left island represents the cliff layover over the ocean, and not the "smoother" coast line of a glacial-eroded fjord, as well as part of the low lands along this coastline cliff thus "disappeared" due to severe layover. On the other hand, it is almost impossible to discriminate the true water-land boundary for the opposite coastline cliffs due to the SAR shadow/occluded areas.

We can notice on Fig. 2 and Fig. 4 that all coastline cliffs and most cirque glaciers as having strong slopes (60°–90°) while the ice fields with their outlet glaciers have in general lower slopes (0°–20°). However, steep 20°–90° slopes also occurred in some ice fields. The lidar ICESat data (ascending and descending tracks) over the 27F13 map sheet was extracted from GLA14 product (L2 global land surface altimetry data) Release 28, 29 and 31 over the full life cycle of the mission 2003–2008. Conversion to Canadian reference systems was applied to translate ICESat GLAS data into orthometric heights. ICESat data points were spatially filtered: (1) horizontally to remove redundant values within 100-m radius, and (2) vertically within 1-arcsec

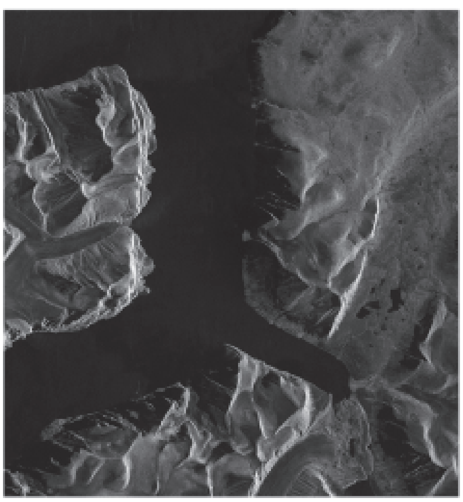


Fig. 4: R2 ultra-fine mode image U12 of the operational site. "Radarsat-2 Data ©MacDonald, Dettwiler and Associates Ltd. (2009) – All Rights Reserved" and Courtesy of CSA.

grid spacing (around 25 m at 70° latitude), to remove potential elevation anomalies resulting from clouds or valley fog.

The processing steps for DSM generation with HR SAR stereo-images were previously addressed and documented (Toutin 2010). The new hybrid Toutin's model developed for the radargrammetric processing of R2 at CCRS does not require any GCP collected by the user (Toutin & Omari 2011): it only uses the information in the metadata of the images for computing its parameters. The hybrid model has been proven to be 25-cm precise (Toutin & Chénier 2009), and the accuracy (1 σ) of the results in stereoscopy was better than one pixel with one-pixel biases in the three axes. The main processing steps, which used CCRS scientific software (steps 1–3 & 6) and commercial OrthoEngine^{SE} software of PCI Geomatics[1] (steps 4 & 5), are:

1. Acquisition and pre-processing of the SLC SAR images (speckle filtering with adaptive Frost SAR filter using 11 by 11 window) and metadata;
2. Collection of 60 ICPs from the DGPS survey, only for the validation site;
3. Computation of our hybrid models and their validation with ICPs (systematic and random errors) (Toutin & Omari 2010);
4. Elevation extraction using a 7-step hierarchical grey-level image matching (mean normalized cross-correlation method with sub-pixel computation of the maximum of the correlation coefficient) performed on the quasi-epipolar stereo-images, (Ostrowski & Cheng 2000). Different matching parameterization for both sites were tested to increase the matching success and to reduce elevation interpolation (Toutin 2010);
5. Edition of the quasi-epipolar DSM (blunders removal, gap interpolation, water bodies), and geocoding of the edited quasi-epipolar DSM.
6. Evaluating (systematic and random errors) the geocoded DSMs with the lidar elevation data (Toutin 2010).

[1] In 2012, all steps can be operationally performed within PCI Geomatics.

4 Results and Discussions

Results are related to the systematic and random errors (1) of the hybrid radargrammetric models computed over ICPs and (2) of the stereo-extracted DSMs computed over the lidar elevation data.

4.1 *Radargrammetric Hybrid Model Evaluation*

Because there was no control data in the operational site, Tab. 1 only summarizes the results of the radargrammetric modelling computation for the validation site and dataset previously described: the errors (bias and standard deviation (Std) in metres) computed over 60 DGPS ICPs providing independent and unbiased evaluations of the modelling accuracy. Biases of one pixel (or half SAR resolution) are obtained. Similarly, Std results in the order of one pixel are better for X-direction. It is certainly due to the better knowledge of the range direction than the azimuth direction corresponding to the satellite displacement. Both results in Z-direction are also very good versus the SAR resolution and the same-side stereo geometry. These results are comparable (10 % difference), but a little worse in the Y-direction, to the original radargrammetric model computed with user-collected GCPs (Toutin & Chénier 2009). On the other hand, the small lost in accuracy for the hybrid model is compensated by the gain of processing the stereo images without GCP.

4.2. *DSM Evaluation*

Validation site

Due to the few percents (less than 3 %) of sparse and small mismatched areas (mainly on water bodies), a basic cubic interpolation was performed to fill the gap without introducing significant additional errors. Visually, the DSM in the quasi-epipolar geometry before the geocoding (Fig. 5) is smooth and well describes the macro-topography and the macro linear trends with mountains and valleys, enhancing the structural geological framework in the northwest-southeast direction. The mountains and valleys are generally smooth, being a good representation of a Precambrian geomorphology (smoothed-glacial and eroded topography).

The quantitative evaluation was performed over the coverage of the lidar data and DSM, approximately 20 km by 20 km corresponding to 100 % of the stereo-pair and the elevation differences between the lidar and DSM are given in Tab. 2. However, the results computed over the full lidar/R2 overlap (Tab. 2, All surfaces) do not reflect the true DSM accuracy since other sources in the error budget comes from: (1) the footprint and penetration in the vegetated cover are different for both sensors (SAR and lidar); and (2) the compared stereo SAR and lidar points are not at the same elevation in the vegetated cover (70 % of lidar coverage). These errors are thus reflected in the -1.9 m bias due to more depth penetration of lidar in forested areas and 4.7 m LE68 due to differential depth penetration of lidar depending of tree species and densities. The negative bias is consistent with Z-bias error computed over ICPs (Tab. 1). To have the true elevation accuracy, the error evaluation was thus performed only on surfaces where the lidar and stereo SAR points were at the same ground elevation (Tab. 2, two last lines). The bare surfaces included the bare soils and the impervious areas: in the country (fields, highways, roads) and in the urban parts (streets, private/public parking and gardens, etc.) and are rep-

Tab. 1: Modelling results over 60 DGPS independent check points (ICPs) for the validation site: Bias and standard deviation (Std) in metres.

Bias (m)			Std (m)		
X	Y	Z	X	Y	Z
1.8	2.6	−2.7	0.93	1.33	2.34

Tab. 2: Differences between lidar data and R2 DSM over different surface types: Bias and LE68 in metres.

Surfaces	Bias (m)	LE68 (m)
All surfaces	−1.9	4.7
Bare surfaces	−0.9	3.7
Urban surfaces	−0.5	3.2

resentative of the full terrain relief because they occur not only on low lands and slopes but also in the high lands and slopes (mainly, in the northeast). The urban surfaces included the buildings only and occurred generally on low land and slopes in the south and few villages in the high lands and slopes in the north. No significant difference was noticed inside or outside the urban parts, when the matching was successful: almost no bias and LE68 of 3.7 m and 3.2 m are thus obtained for bare surfaces and urban surfaces, respectively. The

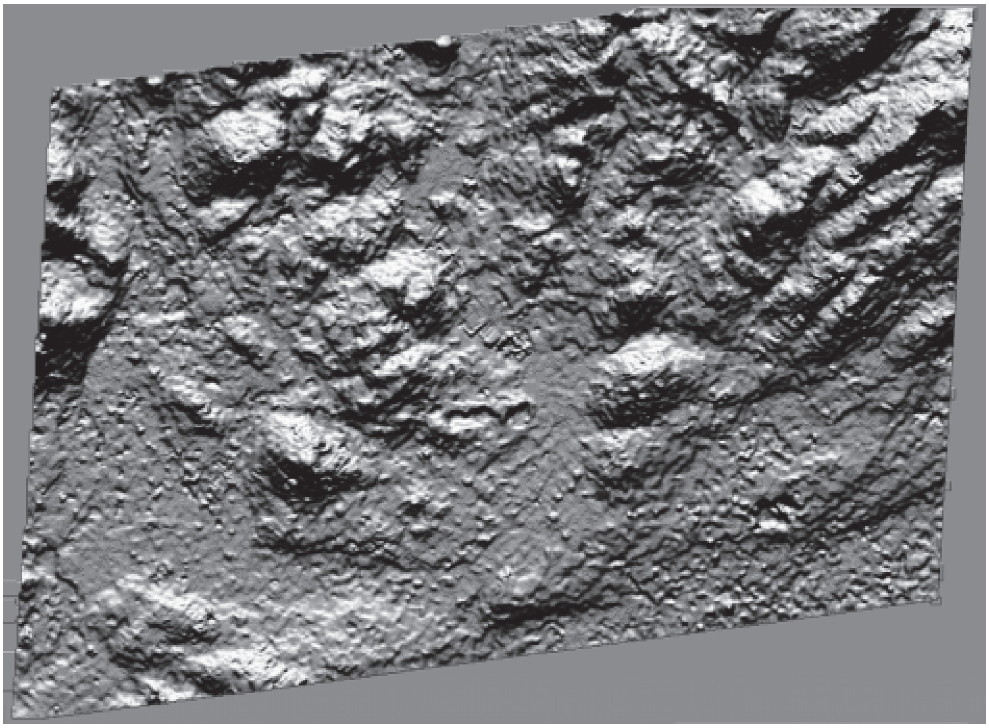


Fig. 5: Shaded-relief DSM (18 km by 16 km) of the validation site in the R2 quasi-epipolar geometry before the geocoding; the main city on low elevations and slopes is in the southeast part.

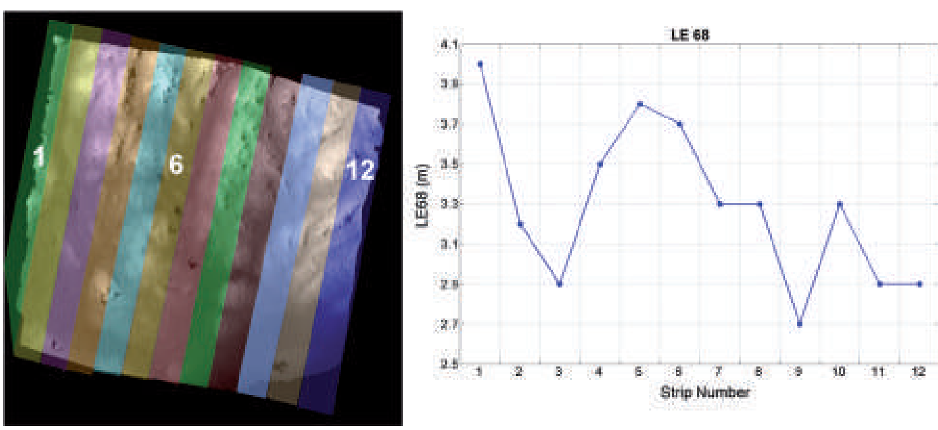


Fig. 6: Left: Geocoded DSM of the validation site with the 12-strip lidar overlaid: Strip 1 at far edge and Strip 12 at near edge. Right: LE68 (in metres) for each strip.

slightly better results over urban surfaces can be due to the fact that most of them generally occurred in low slopes.

The last accuracy evaluation was performed as a function of the SAR range. The lidar data was cut into 12 regular strips (except strip 1), more and less parallel to SAR image azimuth (Fig. 6, left). Strip 1 was at the far edge and strip 12 at the near edge. LE68 (in metres) was thus computed for each strip (Fig. 6, right) and shows there is no correlation between the elevation accuracy and the SAR range, which confirms that the radargrammetric model is not range or look-angle dependent. The result of strip 1 is less significant because there is not enough data when compared to other strips (less than 40 %).

Operational site

The R2 DEM is displayed in Fig. 7 with the ice field and glaciers boundaries (in red) and supraglacial debris and moraines boundaries (in blue). The DEM looks relatively smooth over the ice fields, even with the backscatter homogeneity in ice covered areas, mainly due to the choice of the matching parameterization. In addition, the planned R2 acquisition at the end of the melt season decreased the depth penetration (Dall et al. 2001), increased the roughness and thus the radiometric contrast. During this period, the wet/soaked ice due to freezing/melt process covered by supraglacial debris and dust offered its maximum degree of surface texture and a high backscatter coefficient with shallow SAR look angles. It is the other reason of few percent mismatched areas, which were filled in with the same interpolation without introducing significant additional errors in ice fields. Conversely, the DEM looks very strange along all coast cliffs displaying over 50° slopes and large geometric and radiometric differences between the two images. Instead of generating mismatched pixels, wrong pixels were correlated. The combination of these geometric and radiometric distortions, which only occurred in such a challenging Arctic study site, would certainly impede any image matching system and would have generate either mismatched or wrongly correlated pixels.

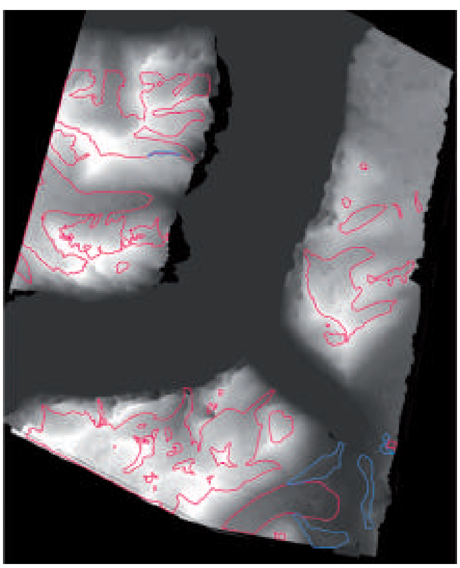


Fig. 7: R2 DSM of the operational site with ice field and glacier boundaries (in red) and supraglacial debris and moraine boundaries (in blue) overlaid.

The quantitative evaluation was performed with ICESat lidar. The height measurements of land surface will be the prime interest for DEM quality assessments (Zwally et al. 2002). While the total number of ICESat data from ascending and descending orbits is limited to thousand points (Fig. 8, blue and red footprints), it will be more limited because ICESat accuracy strongly degrades with slopes. Because the primary goal of ICESat is to measure inter-annual and long-term variations in the polar ice-sheet elevation and volume of Greenland and Antarctica, there were no absolute validation results over more than 5° slopes. Consequently as a function of the expected accuracy for R2 DSM, we only considered for R2 DSM evaluation the ICESat data on slopes less than 30° (Fig. 8, blue footprints). Tab. 3 gives the computed difference between ICESat data and R2 DSM outside (326 points) and over ice fields (387 points).

Because there was no significant bias in the validation site using the same processing, the two biases in this operational site suggest a systematic error in the stereo-model of R2 and thus in the metadata used for this stereo model computation and not from the stereo-radargrammetric processing. While uncertainty about the metadata reliability could

Tab. 3: Differences between ICESat blue footprints and R2 DSM over different surface types: Bias, LE68 and minimum/maximum (Min./Max.) in metres.

Surfaces	ICESat points	Bias (m)	LE68 (m)	Min./Max. (m)
Outside ice fields	326	−16	30	−85/97
Over ice fields	387	−18	21	−88/92

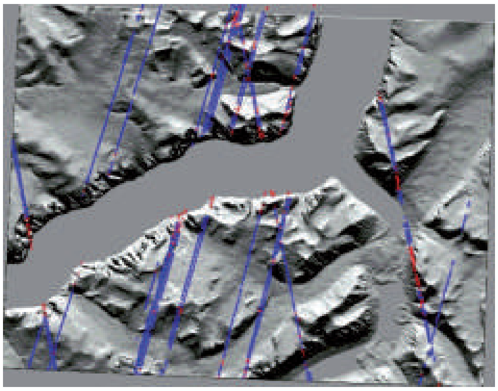


Fig. 8: Ascending and descending ICESat tracks overlaid over a shaded relief image: blue and red footprints are below and over 30° slopes, respectively.

impede the exploitation of the hybrid model with R2 data without GCPs, water bodies or the sea level, frequently occurring in ice regions, can be used for correcting the resulting Z-bias. Because the bias is negatively larger over the ice fields, it is good indication that there was (i) no significant or measurable elevation lowering or ice thinning between the two data acquisitions and (ii) there is no depth penetration. The LE68 over ice fields is a little better (21 m) than outside (30 m) due to less steep slopes over the ice field. In addition, most of ice bodies have low slope relief, other statistical results of elevation differences between ICESat and R2 DEM show that LE68 is strongly correlated with slopes: LE68 of 18 m with an equivalent bias was then computed over ice fields (103 ICESat points) within 0°–5° slopes.

Conclusions

A new hybrid radargrammetric model combining Toutin's deterministic and empirical models and R2 metadata was evaluated with R2 stereo data for DSM generation over two study sites in the north and Arctic of Canada without GCP. The validation site having accurate control data enabled DSM accuracy of 3.7 m (LE68) with no significant bias (-1 m) over bare surfaces to be computed. Even better LE68 (3.2 m) was obtained over buildings, due to their location generally in low slopes. It is certainly the scientist's idealism!

The operational site, a challenging environment with glaciated surfaces, fjords and steep relief, was used to evaluate the mapping potential of the method in the Canadian Arctic without control data. The R2 data acquisition was planned to reduce the SAR depth penetration to its minimum. In this remote and harsh environment, DSM LE68 of 21 m with large bias (-18 m) was achieved over ice field. This major part of this bias is certainly due to a systematic error in the metadata, which can easily be corrected with water bodies or the sea level, which frequently occurred in ice fields. However, the uncertainty of the metadata reliability could slightly impede the operational exploitation of this method due to a possible Z-bias to be corrected: it is maybe the user's realism!

While other methods using optical and SAR systems could achieve similar and even better results, this application demonstrated the capability of R2 to generate DSM with better than 21 m LE68 without collecting control data over ice bodies depending of the terrain slopes (0°–30°) or around 18 m over ice sheets (slopes less than 5°) when the radar penetration is small (less than few metres). This new method increases the applicability of R2 to remote and harsh environments even if there is a slight loss in accuracy when compared to the solution using control points. It is largely compensated by the gain of no control data.

Acknowledgements

The authors would like to thank Paul Briand and the Canadian Space Agency for supporting and financing this research under their SOAR and GRIP programs. They also thank Dr. Philip Cheng and PCI for the adaptation and integration of CCRS math model and algorithms in OrthoEngineSE of PCI Geomatica.

References

Dall, J., Madsen, S.N., Keller, K. & Fosberg, R., 2001: Topography and Penetration of the Greenland Ice Sheet Measured with Airborne SAR Interferometry. – Geophysical Research Letters **28** (09): 1703–1706.

Gray, A.L., 2011: personnel communication.

La Prade, G., 1963: An analytical and experimental study of stereo for radar. – Photogrammetric Engineering **29** (2): 294–300.

La Prade, G. & Leonardo, E., 1969: Elevations from radar imagery. – Photogrammetric Engineering **35** (42): 366–371.

Leberl, F.W., 1972: On model formation with remote sensing imagery. – Österreichische Zeitschrift für Vermessungswesen **2:** 43–61.

Leberl, F.W., 1978: Satellite radargrammetric. – 156 p., Deutsche Geodätische Kommission, Munich, Germany, Serie **C 239**.

Leberl, F.W., 1990: Radargrammetric image processing. – 595 p., Artech House, Norwood, USA.

Ostrowski, J.A. & Cheng, P., 2000: DEM extraction from stereo SAR satellite imagery. – IEEE Geoscience and Remote Sensing Symposium (IGARSS) **2000** (5): 2176–2178, Honolulu, Hawaii.

Rignot, E., Echelmeyer, K. & Krabill, W., 2001: Penetration depth of interferometric synthetic aperture radar signals in snow and ice. – Geophysical Research Letters **28** (18): 3501–3504, doi: 0094-8276/01/2000GL012484.

Rosenfield, G.H., 1968: Stereo radar techniques. – Photogrammetric Engineering **34:** 586–594.

Schwitter, M.P. & Raymond, C.F., 1993: Changes in the Longitudinal Profiles of Glaciers during Advance and Retreat. – Journal of Glaciology **39** (133): 582–590.

Toutin, T., 2000: Elevation Modeling from Satellite Data. – Meyers, R.A. (ed.): Encyclopedia of Analytical Chemistry: 8543–8572, Wiley Publisher.

Toutin, T., 2010: Impact of RADARSAT-2 SAR Ultrafine-Mode Parameters on Stereo-Radargrammetric DEMs. – IEEE Transactions on Geoscience and Remote Sensing **48** (10): 3816–3823.

Toutin, T., 2011: Digital elevation model generation over glacierized regions. – Singh, V.P., Singh, P. & Haritashya, U.K. (eds.): Encyclopedia of Snow, Ice and Glaciers: 202–213, Springer.

Toutin, T. & Chénier, R., 2009: 3-D Radargrammetric Modeling of RADARSAT-2 Ultrafine Mode: Preliminary Results of the Geometric Calibration. – IEEE Geoscience and Remote Sensing Letters **6** (2): 282–286 & **6** (3): 611–615.

Toutin, T. & Gray, A.L., 2000: State-of-the-art of extraction of elevation data using satellite SAR data. – ISPRS Journal of Photogrammetry and Remote Sensing **55** (1): 13–33.

Toutin, T. & Omari, K., 2011: A new hybrid modeling for geometric processing of Radarsat-2 data without user GCP. – Photogrammetric Engineering and Remote Sensing **77** (6): 601–608.

Zwally, H.J., Schutz, B., Abdalati, W., Abshire, J., Bentley, C., Brenner, A., Bufton, J., Dezio, J., Hancock, D., Harding, D., Herring, T., Minster, B., Quinn, K., Palm, S., Spinhirne, J. & Thomas, R., 2002: ICESat's laser measurements of polar ice, atmosphere, ocean, and land. – Journal of Geodynamics **34** (3–4): 405–445.

Addresses of the Authors:

Dr.-Ing. Thierry Toutin, Dr. Khalid Omari & Carla Vanessa Schmitt, Canada Centre for Remote Sensing, Natural Resources Canada, 588 Booth Street, Ottawa, Ontario, K1A 0Y7, Tel.: +1-613-947-1293, -1293, -1326, Fax: +1-613-947-1385, e-mail: {toutin} {komari} {caschmit}@nrcan.gc.ca

Enrique Blondel, Gismatix Inc, 1475 Cumberland Ridge Drive, Ottawa, Ontario, K4C 1E1, Canada, Tel.: +1-613-482-3381, Fax: +1-613-482-3381, e-mail: info@gismatix.ca.

Dr. Daniel Clavet, Centre for Topographic Information, 2144 King Street West, Sherbrooke, Quebec, J1J 2E8, Canada, Tel.: +1-819-564-5600, ext: 255, Fax: +1-819-564-5698, e-mail: dclavet@nrcan.gc.ca

Manuskript eingereicht: Oktober 2011
Angenommen: Februar 2012

PFG 2012 / 4, 0395 – 0408
Stuttgart, August 2012

Article

Maximum Likelihood Classification of Single High-resolution Polarimetric SAR Images in Urban Areas

Maryam Soheili Majd, Elisabeth Simonetto & Laurent Polidori, Le Mans, France

Keywords: PolSAR, urban area, supervised classification, maximum likelihood

Summary: In this work, our aim is to assess the potential of a single polarimetric radar image of high spatial resolution for the classification of urban areas. For that purpose, we concentrate on a fine definition of urban land cover types including ground classes corresponding to different roof types and we test several supervised classification algorithms. In particular, we deal with maximum likelihood classification using several polarimetric and textural indices. At first, we propose a state-of-the-art statistical analysis of polarimetric synthetic aperture radar (SAR) data to study the statistical behaviours of these indices. We consider the Gauss, log-normal, Beta 1, Weibull, Gamma, K, and Fisher statistical models and estimate their parameters using two methods: maximum likelihood estimation (MLE), and method of log-moment (MoLM). The Fisher probability density function (pdf) is able to properly model all the descriptors. Then, we propose to introduce this information in an adapted supervised classification scheme based on maximum likelihood and the Fisher pdf. We compare the classification results with the Wishart-based maximum likelihood algorithm, a Gaussian-based one and SVM (support vector machine). Our experiments are based on an image of a suburban area, acquired by the airborne RAMSES SAR sensor of ONERA, the French Aerospace Lab. The results highlight the potential of such data to discriminate urban land cover types, and the overall accuracy reaches 84 %. However, the results from the tested classification methods show a problematic confusion between roofs and trees. Some possible solutions are discussed at the end of this paper.

Zusammenfassung: Ziel dieser Arbeit ist es, das Potenzial eines einzelnen hoch aufgelösten polarimetrischen Radar-Bildes für die Klassifikation von urbanen Gebieten zu testen. Zu diesem Zweck zielen wir auf eine feine Unterscheidung von urbanen Landbedeckungsarten mit mehreren Objektklassen, die unter anderem auch verschiedenen Arten von Dächern entsprechen, und wir testen verschiedene überwachte Klassifikationsalgorithmen. Insbesondere verwenden wir die Maximum-Likelihood Klassifikation mit mehreren polarimetrischen und texturellen Indices. Zunächst schlagen wir eine Analyse der polarimetrischen Radar-Daten mit synthetischer Apertur (SAR) entsprechend des Standes der Forschung vor, um das statistische Verhalten dieser Indices zu untersuchen. Dabei berücksichtigen wir folgende statistischen Modelle: Gauss, log-normal, Beta 1, Weibull, Gamma, K, und Fisher. Wir schätzen die Parameter dieser Modelle mit Hilfe von zwei Methoden, nämlich Maximum Likelihood Schätzung (MLE) und Methode der log-Momente (MoLM). Die Wahrscheinlichkeitsdichte der Fisher Verteilung kann alle Deskriptoren ausreichend gut modellieren. Anschließend schlagen wir vor, diese Information in einer angepassten überwachten Klassifizierung auf Basis der Maximum Likelihood und der Fisher-Verteilung zu nutzen. Wir vergleichen die Klassifikationsergebnisse mit jenen aus einer Maximum-Likelihood Klassifikation auf Basis der Wishart- bzw. der Gauss-Verteilung sowie mit jenen aus SVM (Support Vector Machine). Unsere Experimente basieren auf einer Szene im Vorstadtbereich, die mit Hilfe des flugzeuggestützten RAMSES SAR Sensors von ONERA aufgenommen wurde. Unsere Ergebnisse zeigen das Potenzial solcher Daten für die Unterscheidung von urbanen Landbedeckungsarten auf. Die Gesamtgenauigkeit der Klassifikation liegt bei 84 %. Allerdings zeigen unsere Ergebnisse auch Probleme, z. B. die häufige Verwechslung von Dächern und Bäumen. Am Ende dieses Beitrages werden mögliche Lösungen für diese Probleme diskutiert.

DOI: 10.1127/1432-8364/2012/0126 1432-8364/12/0126 $ 3.25

1 Introduction

In this work, we want to assess the potential of a single polarimetric SAR dataset in a very high spatial resolution context to classify urban land cover classes: different types of roofs, lawn, bare soil, gardens, roads, highways, car parks, playgrounds, trees, shrubs, etc. For that purpose, we concentrate on eight ground classes: *highways* (including roads, parking lots, and concrete playgrounds), *trees* (including also shrubs), *lawn* (which is defined to include bare soil), three kinds of roofs, namely *flat roofs* (defined as non-metallic roofs covered by gravel or concrete that are usually flat), *sloped roofs* (usually covered by tiles) and *metallic roofs* (that can be flat and/or made of periodic structures (see Simonetto & Malak 2009 for examples), *radar shadows* and *bright pixels*.

Using single polarimetric SAR image sets for terrain classification has been a very active field of research work in recent years (Lee et al. 2001, Lardeux et al. 2007, Ainsworth et al. 2007, Lee & Pottier 2009). Some methods are based on the analysis of the physical scattering mechanisms using various polarimetric decomposition theorems, among them (Cloude & Pottier 1997, Lee et al. 2004, Lee et al. 1999, Pottier & Lee 1999). Other approaches are based on the statistical characteristics of the data or the coherency matrix and derive a decision criterion based on the complex Gaussian or Wishart distributions, as for instance recalled in Lee & Pottier (2009). In addition, the covariance matrix coupled with a Fisher distributed texture under the scalar product model assumption, can be modeled by a KummerU distribution (Bombrun & Beaulieu 2008). There are also classification algorithms based on image processing techniques, for instance approaches based on the Markovian theory (Tison et al. 2004, Frery et al. 2007, Moser et al. 2010), H/A/α unsupervised classification accompanied by decision trees (Simonetto & Malak 2009) or support vector machines (SVM) (Fukuda & Hirosawa 2001, Zhang et al. 2010).

Here, we assess the performance of different supervised classification methods: SVM and several maximum likelihood (ML) approaches. These processes are applied to a feature vector composed of amplitudes, co-polarisation ratio, de-polarization ratios, and other polarimetric and textural descriptors.

Initially, particular attention is given to several theoretical and heuristic models for the probability density functions (pdf) of SAR descriptors. This analysis is needed to propose an adapted criterion for a maximum likelihood classification algorithm. As in (Simonetto et al. 2002, Tison et al. 2004), the parameters of several parametric statistical distribution models (Gauss, Gamma, Beta 1, Weibull, Log-normal, K, Fisher) are estimated from real data according to different methods. Then the fitting of the estimated models is checked using the K-S (Kolmogorov-Smirnov) test, correlation coefficients and a visual analysis (Krylov et al. 2009). Secondly a supervised technique for classification is proposed where an adapted likelihood distance is introduced in a ML framework. In order to assess the performances, the results are compared with those obtained from SVM, ML Wishart classification, and the classical Gaussian-based ML algorithm. The experiments are carried out on a polarimetric image provided by the RAMSES SAR sensor of ONERA, the French Aerospace Lab, over a suburban area.

This paper is structured as follows. The test area and the acquired radar data are presented in section 2, which also includes a short description of the indices used in the feature vector. In section 3, we outline the principle of the statistical analysis and present its results. In section 4, we deal with supervised classification methods and the results achieved by them. Conclusions and perspectives are presented in section 5 of this paper.

2 Studied Area and Dataset

We investigate the potential of X-band fully polarimetric data for discriminating between the principal classes present over a site near Toulouse, France (Fig. 1 a, b). The data are from the ONERA airborne RAMSES (Radar Aéroporté Multi-Spectral d'Etude des Signatures) sensor (Dreuillet et al. 2006). The data is delivered in SLC (single look complex) format, and was acquired in 2006 with an off-nadir viewing angle of 60° and a pixel size of

35 cm in both azimuth and range directions. Because of the large viewing angle, roofs are well represented but not the facades. Besides, these pixel sizes give access to many urban objects and textures on roofs (Simonetto & Malak 2009).

Our main objective is to estimate different urban classes; in particular we are interested in discriminating three types of building roofs: *flat roofs*, *sloped roofs* and *metallic roofs*. *Flat roofs* can be made of different materials such as concrete or gravel, but not of metal. Indeed we do not separate different kinds of flat roofs based on roof material, because it is not possible to distinguish them by image interpretation using both optical and radar views and so it is not possible to build distinct reliable training samples for supervised classification. Indeed, some roofs with the same appearance in the radar image can appear with different colours in the optical view and inversely. In the optical view, they appear in different colours (red, brown and different gray levels) and in the radar data they also show a different appearance (for instance, showing more or less even-bounce or volume-like scattering mechanism). One could work with more roof classes if a ground truth were available to build the training samples. *Sloped roofs* are more likely to be covered by tiles, which are confirmed by their red colour observed in the optical image. The other classes are labelled *lawn*, including lawn and bare soil, *trees* including trees and shrubs, *highways* including roads and car parks, and also two classes, (radar) *shadows* and *bright pixels*. These two last classes do not correspond to any kind of urban objects but have turned out to be useful to limit the misclassification rate. The *bright pixels* are due to specular phenomena that occur on surfaces that are smooth in comparison to the radar wavelength. They consist of direct and multiple bounces, for instance between the building walls and the ground or between some objects located on a roof and the roof itself. This also happens on point-like objects such as street lights. According to the Rayleigh criterion more bright pixels are expected at larger viewing angles (Beckman & Spizzichino 1987).

We define two sets of knowledge: training samples as input in the supervised classification processes and control samples to assess the performance. Training and control data have been extracted manually by visual interpretation of optical views and the radar image (Fig. 1c, d). The training and control data are summarized in Tab. 1.

The fully polarimetric radar system records the complete characterization of the scattering field in all the configurations (HH, VV, HV, VH). Here, we suppose reciprocity with HV=VH. The recorded polarimetric data allow a better characterization of the surfaces based on the decomposition theorems. In this work, we use twenty descriptors summarized in Tabs. 2 and 3. In Tab. 2, S_{xy} represents the scattering coefficient of the targets, $\boldsymbol{x}$ the polarization of the incident field, $\boldsymbol{y}$ the polarization of the scattered field, λ_i denotes the i^{th} eigenvalue of the coherency matrix. In the $H / A / \alpha$ decomposition (Cloude & Pottier 1996), the angle α represents the main scattering mechanism (odd-bounce scattering, volume scattering or even-bounce scattering, etc). The

Tab. 1: Training and control samples (numbers of pixels).

Classes	**Training data**	**Control data**
Radar shadows	3596	2083
Highways (roads, parks…)	8538	8479
Lawns (lawns, bare soils)	17955	17929
Trees (trees, small shrubs)	4281	3719
Bright pixels	1990	766
Metallic roofs	546	423
Flat roofs	4958	2615
Sloped roofs	235	264

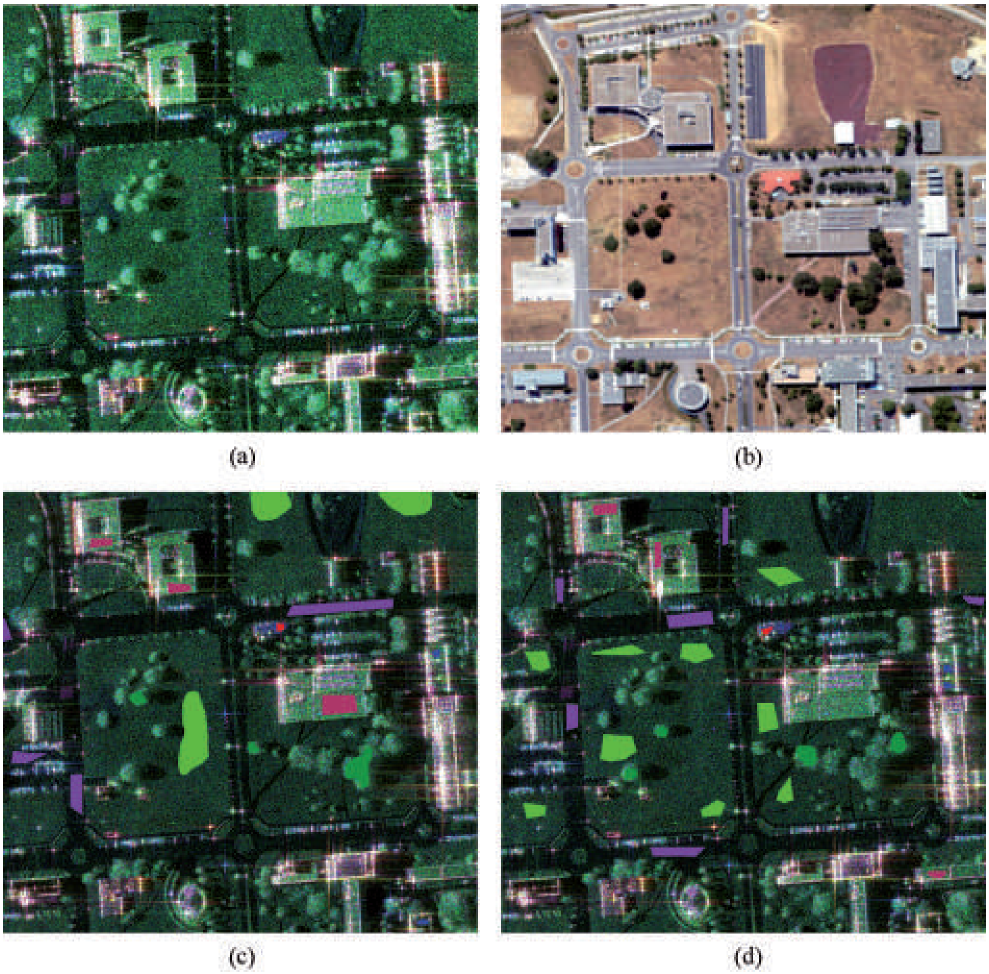


Fig. 1: (a) RAMSES image of the studied site (Toulouse, France), acquired in 2006, in Pauli colour-coded representation, (b) Aerial view from the French Geographical Institute (IGN) acquired in 2006, (c) Training samples superimposed on the RAMSES view (same colour legend as in Fig. 4), (d) Control samples superimposed on the view (same colour legend as in Fig. 4).

entropy H measures the degree of the randomness of the scattering process with the probabilities P_i .The anisotropy A gives the relative importance of the second and third eigenvalues (λ_2 and λ_3, respectively). We also use the polarization ratios (items 9, 10 and 11 in Tab. 2): the first one is the co-polarization ratio and measures the ratio of the co-polarized signals, S_{HH} and S_{VV}. The other two ratios are e-polarization ratios, $|S_{HV}|/|S_{VV}|$ and $|S_{HV}|/|S_{HH}|$.

As the texture information can be used for improving the classification results (Zhang et al. 2010), we also include it in our analysis: ENR (energy) gives a measure of textural uniformity, CON (contrast) explains the local variation of grey levels, and HOM (homogeneity) shows the local similarity of the grey values. These statistical parameters are extracted from the grey level co-occurrence matrix (GLCM) (Haralick 1979) and are summarized in Tab. 3. Each element of this matrix shows the relative frequencies $q(i,j)$ of two neighbouring values that are related to the distance d and angle θ (Zhang et al. 2010). We arbitrarily choose d=1, and $\theta=0°$. Textural parameters are computed on the three images:

Tab. 2: Polarimetric descriptors considered in this work. ⟨ ⟩ stands for the local mean.

No.	Feature	Expression
1	|Red Pauli|	$\lvert S_{HH} - S_{VV} \rvert / \sqrt{2}$
2	|Blue Pauli|	$\lvert S_{HH} + S_{VV} \rvert / \sqrt{2}$
3	Amplitude	$\lvert S_{HH} \rvert$
4	Amplitude	$\lvert S_{VV} \rvert$
5	Amplitude	$\lvert S_{HV} \rvert$
6	Entropy	$H=\sum_{i} P\log(P_i), P=\frac{\lambda_i}{\sum_{i=1}^{3}\lambda_i}$
7	Anisotropy	$A = (\lambda_2 - \lambda_3) / (\lambda_2 + \lambda_3)$
8	Alpha	$\alpha = \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3$
9	Co-polarisation ratio	$\langle \lvert S_{HH} \rvert \rangle / \langle \lvert S_{VV} \rvert \rangle$
10	De-polarisation ratio	$\langle \lvert S_{HV} \rvert \rangle / \langle \lvert S_{VV} \rvert \rangle$
11	De-polarisation ratio	$\langle \lvert S_{HV} \rvert \rangle / \langle \lvert S_{HH} \rvert \rangle$

T11, T22 and T33, which are the three diagonal components of coherency matrix T using a 5x5 neighbourhood.

Tab. 3: Mathematical expressions of statistical texture parameters from GLCM. q(i,j) : (i,j)th element in the co-occurrence matrix (GLCM = grey level co-occurrence matrix).

Texture parameter	Expression
ENR (energy)	$\sum_{i,j} q(i,j)^2$
CON (contrast)	$\sum_{i,j} \lvert i-j \rvert^2 q(i,j)$
HOM (homogeneity)	$\sum_{i,j} \frac{q(i,j)}{1+\lvert i-j \rvert}$

3 Statistical Analysis

The precise knowledge of the statistical properties of the SAR data plays a central role in SAR image processing and understanding (Gao 2010). The purpose of this section is to validate a classical statistical model for polarimetric or textural features from our data. This model will then be used in the classification scheme. Besides, this analysis allows us to check the possibility of discriminating each type of land cover according to its statistical behaviour. The process of parametric modelling consists of: (1) selecting several known statistical distribution models; (2) estimating the distribution parameters; (3) assessing the goodness-of-fit of the models.

We select the most widely used distributions in the literature, which are the Gauss, Gamma, Weibull, Beta 1, log-normal, K, and

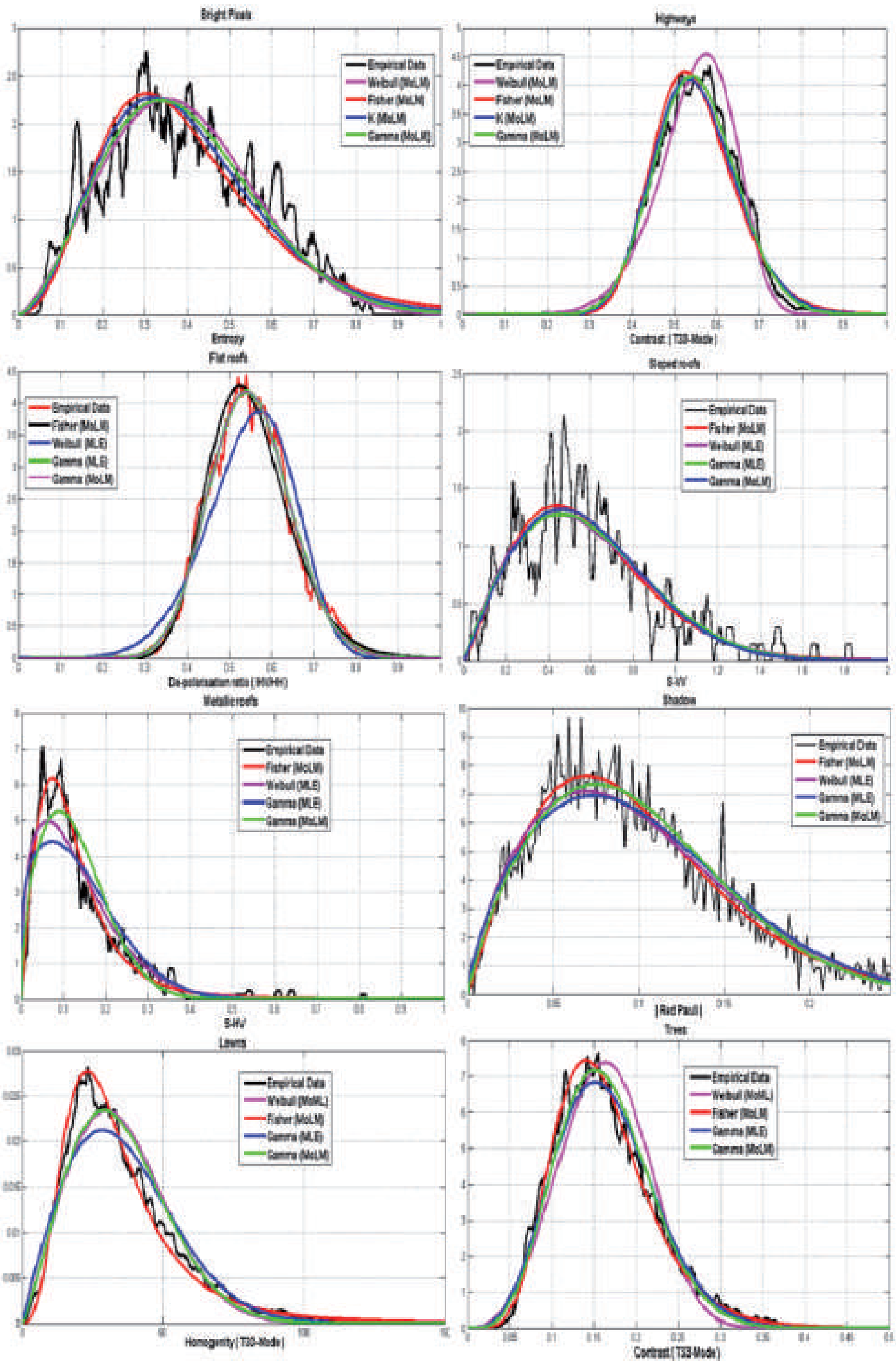


Fig. 2: Empirical data and estimated pdfs using MoLM and MLE for three kinds of roofs, bright pixels, highways, lawns, shadows and trees, and different descriptors (MoLM = method of log-moment, MLE = maximum likelihood estimation).

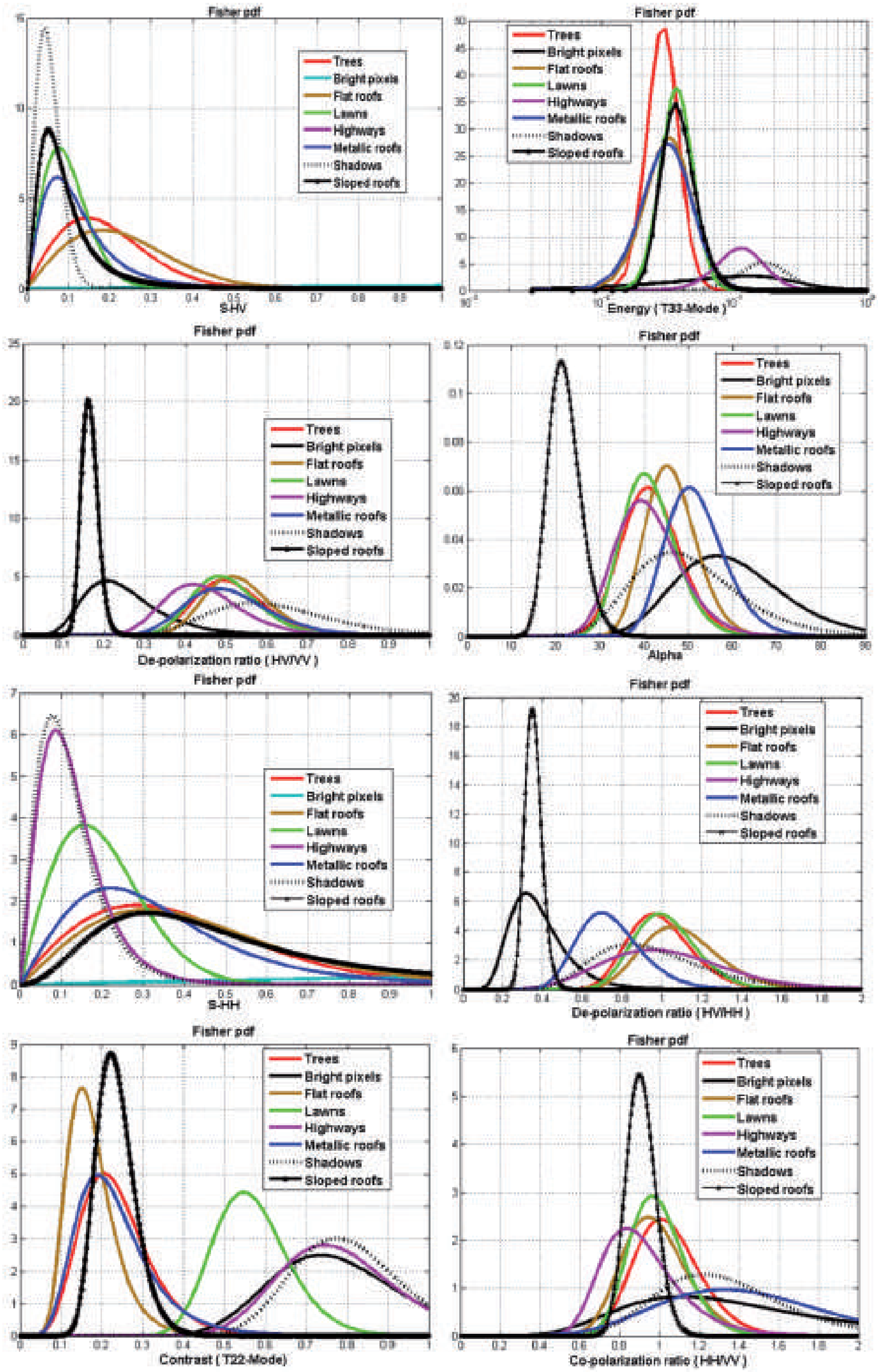


Fig. 3: Estimated Fisher distribution by MoLM for the eight classes and different descriptors.

Fisher distributions. To estimate their parameters, we use two methods: maximum likelihood estimation (MLE), and the method of log-moment (MoLM). We do not use the MLE for K and Fisher pdfs because of mathematical limitations; see Nicolas (2002) for more details. The MoLM is based on the characteristic function computed on logarithmic scale, from which we obtain parameter estimators from the log-cumulants (Nicolas 2002, Nicolas 2006). The formula and algorithm of the Fisher MoML parameter estimation can be found in Nicolas (2002, page 130). We expect the Fisher distribution to be the more appropriate as it has already been validated for high-resolution SAR amplitude statistics over urban regions (Nicolas 2006, Tison et al. 2004, Bombrun & Beaulieu 2008).

The GoF (goodness-of-fit) can be assessed by different ways. The K-S (Kolmogorov-Smirnov) hypothesis test involves comparing the empirical and modelled cumulative distribution functions (Feller 1948). In this paper, correlation coefficients between empirical and modelled histograms and visual analysis are also used (Krylov et al. 2009). During the visual analysis, we represent the empirical and estimated pdfs for each descriptor, each function, each parameter estimation method and each class.

Examples of empirical and modelled histograms are displayed in Fig. 2 for eight descriptors and eight classes. The quantitative results of GoF, obtained with the correlation coefficient and the K-S test for three kinds of roofs and three descriptors are presented in Tab. 4. These comparisons are performed for all descriptors over the eight classes, with all possible parametric functions and both parameter estimation methods.

It turns out that the Fisher pdf is the best model according to the K-S distance or the

Tab. 4: Goodness-of-fit: Correlation coefficients and the K-S (Kolmogorov-Smirnov) distance for three descriptors over three kinds of roofs and for different parametric modelling (MoLM = method of log-moment, MLE = maximum likelihood estimation).

	Sloped roofs and $\|S_{VV}\|$				**Metallic roofs and $\|S_{HV}\|$**				**Flat roofs and De-polarisation ratio ($<\|S_{HV}\|>/<\|S_{HH}\|>$)**			
	Weibull (MLE)	K (MoLM)	Gamma (MoLM)	Fisher (MoLM)	Weibull (MLE)	K (MoLM)	Gamma (MoLM)	Fisher (MoLM)	Weibull (MLE)	K (MoLM)	Gamma (MoLM)	Fisher (MoLM)
K-S test Distance	0.328	0.3400	0.3433	0.343	0.4160	0.458	0.543	0.316	0.534	0.563	0.566	0.564
Correlation	0.910	0.9142	0.9132	0.915	0.9671	0.985	0.963	0.988	0.974	0.994	0.995	0.992

Tab. 5: Mathematical expressions of the Fisher pdf (for radar amplitude), and the corresponding ML criterion. *M*, *L*, μ are the Fisher parameters. *C*(*i,j*): Decided class of the pixel (*i,j*). N_c: number of classes. N_b: number of features (or bands). $u_b(i,j)$: value of the b^{th} feature of the pixel (*i,j*) (pdf = probability density function).

Fisher pdf	$p[\mu, L, M](u) = \frac{2\Gamma(M+L)}{\Gamma(M)\Gamma(L)} \sqrt{\frac{L}{M\mu}} \frac{\left(\sqrt{\frac{L}{M\mu}}u\right)^{2L-1}}{\left(1+\left(\sqrt{\frac{L}{M\mu}}u\right)^2\right)^{L+M}}$
Fisher based ML criterion	$C(i,j) = \underset{k\in[1,N_C]}{\arg\max}\left[\sum_{b=1}^{N_b}\log(p[\mu_{k,b}, L_{k,b}, M_{k,b}](u_b(i,j))) - \log N_c\right]$

correlation coefficient. For example the correlation coefficients related to the Red Pauli descriptor of *bright pixels*, are 0.9323, -0.0817, not exist, 0.0688 and 0.8891, respectively, for the following models: Fisher (MoLM), Weibull (MLE), Gamma (MLE), Gamma (MoLM), K (MoLM). The correlation coefficients using the Fisher model are higher, more than 91 %, than using the other models, and the K-S distances are smaller (significance level < 5 %). This proves the pertinence of the Fisher pdf that can be retained for all bands and all classes using the MoLM. As a result, the Fisher pdfs for each of the eight classes and for different descriptors are shown in Fig. 3.

Besides, such analysis confirms the existence of clear different statistical behaviours such as between *bright pixels* and *highways* using a de-polarization ratio, or between *highways* and *shadows* using the co-polarization ratio.

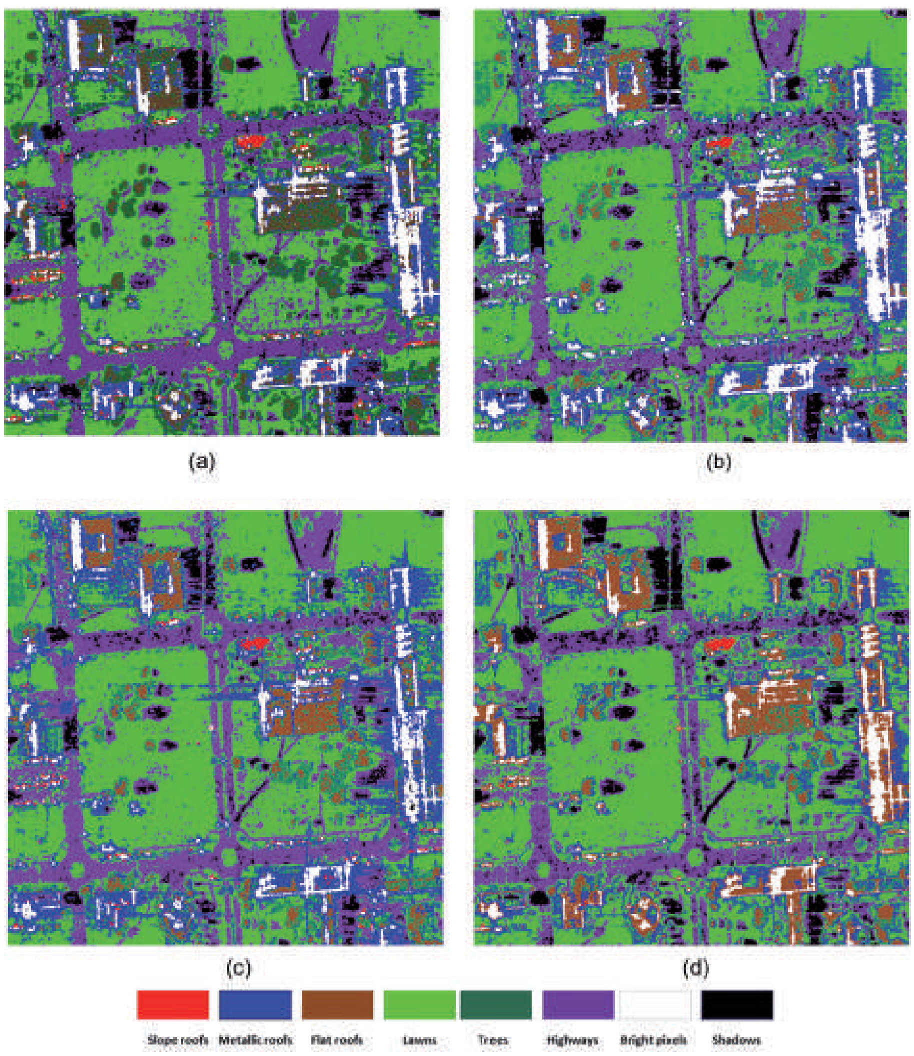


Fig. 4: Classification result: (a) SVM * (b) G-ML * (c) F-ML * (d) W-ML. * is for classification using 20 bands (11 polarimetric descriptors and 9 textural ones, SVM = support vector machine, G-ML = circular Gaussian-based, F-ML = Fisher-based, W-ML = Complex Wishart, ML = Maximum likelihood classification).

4 Classification Methods and Results

Here, we propose a supervised ML classification method based on Fisher pdf as prior knowledge (Tab. 5). It assumes the statistical independency between features. This new F-ML (maximum likelihood based on Fisher pdf) algorithm is implemented in MATLAB code.

We compare this approach with three supervised ones: SVM (using ENVI) (Wu et al. 2004), G-ML (circular Gaussian-based ML classification) (MATLAB code) (Richards 1999) and W-ML (Complex Wishart ML classification) (MATLAB code) (Lee & Pottier 2009). Note that for W-ML, the ML criterion is applied to the polarimetric coherency matrix and not on the feature vectors. SVM, G-ML and F-ML are performed using feature vectors with either 11 polarimetric descriptors (Tab. 2) or 20 bands when adding the 9 texture descriptors (Tab. 3). Besides, due to a limitation of the ENVI SVM algorithm, we had to reduce the size of the training samples (not more than about 1000 pixels for each of them).

Tab. 6: Performance of tested classification methods: producer's accuracy for each class, kappa coefficient and overall accuracy. * stands for a classification using 20 bands with 11 polarimetric and 9 textural descriptors (W-/G-/F-ML and SVM see Fig. 4).

	Shadows	Highways	Lawns	Trees	Bright pixels	Metallic roofs	Flat roofs	Sloped roofs	Kappa coefficient	Overall accuracy
W-ML	87.1%	59.3%	94.4%	33.2%	85.3%	46.8%	66.0%	81.1 %	0.67	76.6%
SVM	67.4%	73.4%	67.4%	36.1%	74.9%	58.2%	49.4%	81.4%	0.50	64.4%
G-ML	69.3%	65.7%	71.4%	46.3%	86.8%	50.4%	47.8%	64.8%	0.52	65.7%
F-ML	78.8%	66.9%	65.2%	41.7%	80.4%	44.4%	65.6%	80.7%	0.51	64.2%
SVM *	63.6%	80.8%	96.7%	63.8%	83.0%	29.3%	70.4%	69.7%	0.77	84.5%
G-ML*	88.0%	68.5%	93.1%	58.6%	90.5%	30.3%	67.3%	59.1%	0.72	80.6%
F-ML*	70.0%	71.0%	94.1%	53.4%	83.4%	22.5%	76.9%	79.9%	0.72	80.7%

Tab. 7: Confusion Matrix (in percent) for F-ML* (see Fig. 4).

Ground-truth ---------- Class	Shadows	Highways	Lawns	Trees	Bright pixels	Metallic roofs	Flat roofs	Sloped roofs	Total
Shadows	70.0	25.0	0	0	1.6	0	0	0	9.9
Highways	21.7	71.0	3.1	0	0	0	0	0	19.4
Lawns	0	1.4	94.1	5.9	0	0	0	0	47.4
Trees	0	0	0.9	53.4	1.3	47.5	21.2	9.8	8.1
Bright pixels	0	0	0	0	83.4	0	0	0	1.8
Metallic roofs	8.4	2.5	1.9	3.4	11.6	22.5	1.9	7.6	3.1
Flat roofs	0	0	0	37.3	2.1	30.0	76.9	2.7	9.8
Sloped roofs	0	0	0	0	0	0	0	79.9	0.6
Total	100	100	100	100	100	100	100	100	100

Some classification results from the different approaches are shown in Fig. 4. Tab. 6 gives the performance of the different tests with and without the textural information. The performance is presented in the different columns of this table: "rate of good classification for each class" (in columns 2 to 9), "kappa coefficient" that is the Cohen's Kappa computed in ENVI (in column 10) and "overall accuracy" defined by the rate of good classification (producer's accuracy) (in column 11).

In this table, we observe that *shadows* and *bright pixels* are better classified with the G-ML using 20 bands. Indeed, the Fisher pdf is the best modelling for each feature according to the analysis in section 3 but our F-ML criterion does not take into account the band statistical dependency, which could explain this result. Despite this limitation, the F-ML using 20 bands leads to the best producer's accuracy for the class *flat roofs*. SVM with 20 bands is preferable for *highways*, *lawns* and *trees*. The classification of sloped roof pixels is better with SVM or F-ML without using the textural descriptors. Note that the image shows only one sloped roof and the training and control samples for this class are smaller than for the other classes. The best result for *metallic roofs* is obtained with SVM without using the textural descriptors.

We can state that using the texture properties such as uniformity, contrast, and homogeneity, the overall accuracy result for all three classification methods is greatly improved; for SVM from 64.4 % to 84.5 %, G-ML from 65.7 % to 80.6 %, and F-ML from 64.2 % to 80.7 % (Tab. 6).

Indeed, the producer's accuracies have increased for most classes.

This is not the case for the classes *metallic roofs*, *shadows* and *sloped roofs* (Tab. 6). If we observe the complete confusion matrices (Tab. 7 for an example), the confusion between *shadows* and *highways*, and between *metallic roofs* and *flat roofs* and *trees*, are more important in the results when using the textural descriptors (that is, 20 bands).

On the other hand, certain rates of misclassification have been reduced. This is the case for *trees*, for which two main classification errors are caused by a confusion with *flat roofs* and *lawns*. Using the F-ML, the confusion with *lawns* is reduced from 20.6 % to 5.9 % (Tab. 7). However, the confusion with *flat roofs* is increased from 31.4 % to 37.3 %. The same tendency is observed using the G-ML algorithm. Similarly, *flat roofs* are confused with the class *trees*. The rates of confusion of *flat roof* pixels that are erroneously classified as *trees* using textural descriptors are equal to 21.2 % with F-ML, 31.8 % with G-ML and 30.0 % with SVM. Then, the best result is obtained with the F-ML approach for *flat roofs*. In the same way, the *lawn* is less frequently confused with *highways* and *trees* if the textural bands are used.

To conclude this analysis, the textural features are relevant to the classification process for most classes. Although we do not obtain the best classification accuracy with the F-ML algorithm, it allows for a better discrimination of *lawns* and *flat roofs*. However, the results obtained for *metallic roofs* with this criterion are quite poor.

5 Conclusions and Perspectives

This paper has presented a validation of the Fisher pdf to model twenty polarimetric and textural descriptors and proposed a new supervised ML classification based on the Fisher pdf. The result is compared with several supervised classification results (SVM, ML based on a Gaussian pdf, Complex Wishart ML algorithm).

The overall classification accuracy using twenty bands obtained from the F-ML classification is 81 % and Cohen's Kappa is equal to 0.72. Although the Fisher ML criterion does not show the best overall accuracy compared to other classifiers, the classification results obtained by F-ML have a less noisy visual appearance, and the algorithm appears to be more efficient than SVM, W-ML or G-ML to discriminate flat roofs from trees. Introducing textural features allows us to greatly increase the overall accuracy of the classification process. Nevertheless, misclassifications still occur between flat roofs and trees and for metallic roofs.

This result could be considered poor compared to those that one can obtain using an optical data. However, a first work has shown that the classified radar image provides information that is complementary to a classified spectral image (Long et al. 2010). For example, the radar data allows to separate road surfaces from flat roofs which was not the case with the spectral data.

As a suggestion for future work, a comparison with the KummerU-distribution based classification (Bombrun & Beaulieu 2008) would be interesting. We will study a new Fisher ML criterion that takes into account the statistical dependencies between features. Besides, we think that contextual information is necessary to properly discriminate different kinds of roofs and to overcome the confusion between some kind of roofs and trees. Contextual information could be other radar images (with a different wavelength), spectral data, or geometric information (shape or elevation). These topics may be investigated in future studies.

Acknowledgements

We thank ONERA/DEMR for providing the RAMSES data.

References

Ainsworth, T.L., Lee, J.-S. & Chang, L.W., 2007: Classification comparisons between dual-pol, compact polarimetric and quad-pol SAR imagery. – IEEE IGARSS: 164–167.

Beckman, P. & Spizzichino, A., 1987: The scattering of electromagnetic waves from rough surfaces. – 512 p., Artech House Radar Library, Norwood, Massachusetts, USA.

Bombrun, L. & Beaulieu, J.M., 2008: Fisher Distribution for Texture Modeling of Polarimetric SAR Data. – IEEE Geoscience and Remote Sensing Letters **5** (3): 512–516.

Cloude, S.R. & Pottier, E., 1996: A Review of Target Decomposition Theorems in Radar Polarimetry. – IEEE TGRS **34** (2): 498–518.

Cloude, S.R. & Pottier, E., 1997: An entropy based classification scheme for land applications of polarimetric SAR. – IEEE TGRS **35** (1): 68–78.

Dreuillet, P., Cantalloube, H., Colin, E., Dubois-Fernandez, P., Dupuis, X., Fromage, P., Garestier, F., Heuze, D., Oriot, H., Peron, J.L., Peyret, J., Bonin, G., Du Plessis, O.R., Nouvel, J.F. & Vaizan, B., 2006: The ONERA RAMSES SAR-latest significant results and future developments. – IEEE Conf. Radar: 2046–2057.

Feller, W., 1948: On the Kolmogorov-Smirnov Limit Theorems for Empirical Distributions. – Annals Mathematical Statistics **19** (2): 177–189.

Frery, A.C., Correia, A.H. & Da Freitas, C.D., 2007: Classifying Multifrequency Fully Polarimetric Imagery With Multiple Sources of Statistical Evidence and Contextual Information. – IEEE TGRS **45** (10): 3098–3109.

Fukuda, S. & Hirosawa, H., 2001: Support vector machine classification of land cover: application to polarimetric SAR data. – IEEE IGARSS: 187–189.

Gao, G., 2010: Statistical Modeling of SAR Images: A Survey. – Sensors **10** (1): 775–795.

Haralick, R.M., 1979: Statistical and structural approaches to texture. – Proceedings of the IEEE **67** (5): 786–804.

Krylov, V., Moser, G., Serpico, S.B. & Zerubia, J., 2009: Modeling the statistics of high resolution SAR images. – INRIA-00342681.

Lardeux, C., Frison, P.-L., Tison, C., Deleflie, D., Souyris, J.-C., Rudant, J.-P. & Stoll, B., 2007: Comparison of Compact Polarimetric with Full Polarimetric Radar Data for Land Use Discrimination Based on SVM. – POLinSAR, Frascati, Italy.

Lee, J.S., Grunes, M.R., Ainsworth, T.L., Du, L.J., Schuler, D.L. & Cloude, S.R., 1999: Unsupervised Classification using Polarimetric Decomposition and the Complex Wishart classifier. – IEEE TGRS **37** (5): 2249–2259.

Lee, J.S., Grunes, M.R. & Pottier, E., 2001: Quantitative Comparison of Classification Capability – Fully Polarimetric versus Multipolarisation SAR. – IEEE TGRS **39** (1): 2343–2351.

Lee, J.S., Grunes, M.R., Pottier, E. & Ferro-Famil, L., 2004: Unsupervised terrain classification preserving scattering characteristics. – IEEE TGRS **42** (4): 722–731.

Lee, J.S. & Pottier, E., 2009: Polarimetric Radar Imaging: From Basics to Applications. – Optical Science and Engineering, B.J. Thompson, New York.

Long, N., Simonetto, E. & Bocher, E., 2010: A combined approach to detect urban features from multi-spectral and radar data. – IEEE IGARSS: 1469–1472.

Moser, G., Krylov, V., Serpico, S.B. & Zerubia, J., 2010: High-resolution SAR-image classification by Markov random fields and finite mixtures. – SPIE Computational Imaging, doi:10.1117/12.838594.

Nicolas, J.-M., 2002: Introduction aux statistiques de deuxième espèce: Application des logs-moments et des logs-cumulants à l'analyse des lois d'images radar. – Traitement du Signal **19** (3): 139–167.

Nicolas, J.-M., 2006: Application de la transformée de Mellin: Etude des lois statistiques de l'imagerie cohérente. – ENST, Paris, France.

Pottier, E. & Lee, J.S., 1999: Application of the H/A/α Polarimetric Decomposition Classification of fully polarimetric SAR data based on the Wishart Distribution. – CEOS SAR Workshop, ES-SP 450: 335–340.

Richards, J.A., 1999: Remote Sensing Digital Image Analysis. – 240, Springer-Verlag, Berlin.

Simonetto, E., Oriot, H. & Garello, R., 2002: Extraction of industrial buildings from stereoscopic airborne radar images. – SPIE Remote Sensing **4543:** 121–129.

Simonetto, E. & Malak, C., 2009: Urban area structuring mapping using an airborne polarimetric SAR image. – SPIE Remote Sensing **7477:** 74771X.

Tison, C., Nicolas, J.-M., Tupin, F. & Maître, H., 2004: A new statistical model for Markovian classification of urban areas in high resolution SAR images. – IEEE TGRS **42** (10): 2046–2057.

Wu, T., Lin, C.-J. & Weng, R.C., 2004: Probability estimates for multi-class classification by pairwise coupling. – Journal of Machine Learning Research **5:** 975–1005.

Zhang, L., Zou, B., Zhang, J. & Zhang, Y., 2010: Classification of Polarimetric SAR Image Based on Support Vector Machine Using Multiple-Component Scattering Model and Texture Features. – EURASIP Journal on Advance in Signal Processing **2010:** 21–29.

Address of the Authors:

Maryam Soheili Majd, Elisabeth Simonetto and Laurent Polidori, L2G, ESGT, 1 Bd Pythagore, 72000 Le Mans, France, Tel.: +33-2-4343-3100, Fax: +33-2-4343-3102, e-mail: {maryam.soheili_majd}{elisabeth.simonetto}{laurent.polidori}@esgt.cnam.fr

Manuskript eingereicht: Oktober 2011
Angenommen: Mai 2012

PFG 2012 / 4, 0409–0420
Stuttgart, August 2012 **Article**

Matching and Evaluating Free-form Linear Features for Georeferencing Space-borne SAR Imagery

Dimitra Vassilaki, Athens, Greece

Keywords: georeferencing, TerraSAR-X, optical images, multitemporal, linear features

Summary: In this paper, 3D free-form linear features are used as ground control information for the computation of the transformation from 3D object space to the 2D SAR image space. This is now feasible thanks to the high resolution imaging capabilities of contemporary satellite SAR sensors which allow the identification of detailed structures. The computation of the transformation parameters is based on a newly introduced general iterative closest point (ICP) based method for single free-form linear feature pair matching, extended to simultaneously match multiple pairs of them. The proposed method is tested with the georeferencing of a whole TerraSAR-X image. Linear features are shown to be a reliable form of ground control information.

Zusammenfassung: *Matching und Evaluation von Freiform-Linien als Passinformation zur Georeferenzierung von satellitengestützten SAR-Bildern*. In diesem Beitrag werden 3D Freiform-Linien als Passinformation für die Berechnung der Parameter zur Transformation vom 3D-Objektraum in den 2D SAR-Bildraum verwendet. Dies ist dank hoch auflösender moderner SAR Satellitenradarsensoren möglich geworden, die eine Identifizierung detaillierter Strukturen erlauben. Die Berechnung der Transformationsparameter basiert auf einer neu eingeführten Methode, welche eine Verallgemeinerung des Iterative Closest Point-Algorithmus (ICP) zur Zuordnung von einzelnen Paaren von Freiform-Linien darstellt und eine gleichzeitige Identifikation und Zuordnung mehrerer solcher Paare erlaubt. Die vorgeschlagene Methode wird an Hand der Georeferenzierung eines TerraSAR-X-Bildes getestet. Die Linien erwiesen sich dabei als eine zuverlässige Form der Passinformation.

1 Introduction

The fine resolution imaging capability of the modern, state-of-the-art, spaceborne SAR sensors makes feasible the identification of certain details on the Earth's surface such as roads and buildings. This was hard or even impossible with the previous generation of SAR images (Eineder et al. 2009). Modern satellite SAR sensors such as TerraSAR-X collect images that are delivered with an absolute geolocation accuracy of 1–2 m (Roth et al. 2004, Bresnahan 2009). However, the identification of salient characteristic points such as road intersections and building corners is still an ambiguous process, due to the inherently speckled and fuzzy nature of SAR images, the severe distortions inherited from the imaging geometry and the absence of true colour. This is not necessarily a drawback for the SAR images; contemporary SAR sensors may not need control information for orientation processes due to accurate orbit data (Vassilaki et al. 2011a). It is more a problem of the co-processing (combined/synergistic use) of slant range (not geometrically corrected) SAR data with heterogeneous (multimodal and multitemporal) remote sensing and geospatial data, in the sense of identifying salient points to correlate the other data with SAR. The concept of co-processing SAR data with other data types is of growing significance (Sörgel et al. 2008, Wegner et al. 2009, Suri & Reinartz 2010, Vassilaki et al. 2011b, Wegner et al. 2011, Reinartz et al. 2011). There is a worldwide abundance of readily available geospatial informa-

DOI: 10.1127/1432-8364/2012/0127 1432-8364/12/0127 $ 3.00

tion as other type of data. Moreover, optical images are of well-established use and of more descriptive nature as they are closer to the human eye perception. On the other hand, SAR is an all-weather, day and night sensor, offering information about the properties of the targets, their 3D geometry and their evolution over time, which complements information found in other data types.

In order to work around the ambiguous identification of salient points on the fuzzy and speckled SAR images, linear features such as road edges and building outlines are investigated as an alternative way to provide ground control information (GCI). The concept of using non salient point GCI for the geometric co-processing of SAR images with other, heterogeneous, remote sensing and geospatial data is relatively new, as the identification of details was hard with previous generations of SAR images; Dare & Dowman (2001) and Karjalainen (2007) used water body boundaries, which are effectively 2D, for registration problems and Zalmanson et al. (2004) attempted an introductory research on a more general 3D approach using parametric curves, with no applications or results. In this paper the georeferencing of TerraSAR-X images is done with a newly introduced ICP-based free-form linear features (FFLFs) matching method (Vassilaki et al. 2008). FFLFs are 2D or 3D linear features of arbitrary geometry, such as road edges and building outlines, defined by nodes of no regularity which are joined by a function of arbitrary type (Vassilaki et al. 2012). In this paper the method is extended to match networks of FFLFs of different dimensionality (3D-2D). The method is applied to the georeferencing of a whole TerraSAR-X image; a network of 3D FFLFs is used as GCI or ground control linear features (GCLFs) in order to compute the parameters of the transformation from 3D object space to 2D SAR image space.

2 Problem Formulation

A SAR image is a 2D projection (x,y) of the 3D object space (X,Y,Z). The georeferencing process recovers the imaging geometry of the sensor at the time of acquisition, by the computation of the parameters of the transformation. Transformations are generally classified into empirical models and into physical sensor models (Toutin 2004). Examples of empirical models are polynomial functions (PFs) and rational polynomial functions (RPFs), while physical sensor models are sensor dependent rigorous mathematical functions. In this paper the georeferencing is computed by using PFs and RPFs as transformation models, and the results are compared to those achieved using the physical model of Vassilaki et al. (2011a). Although not widely adopted for SAR images (Toutin 2004, Dowmann & Dolloff 2000), the use of empirical sensor models is still geometrically meaningful as it was recently explained by Zhang et al. (2011). Traditionally the computation of the transformation parameters, either empirical or physical, requires salient points with known 3D object-space and 2D image-space coordinates (ground control points, GCPs).

In this paper the georeferencing parameters are computed using FFLFs with known 3D object-space and 2D image-space coordinates (GCLFs). In contrast to salient points, the 3D information and the 2D information are in the form of a pair of two distinct heterogeneous FFLFs defined by their nodes (measured points). These nodes may vary widely as they are produced using different processes. Even though a corresponding pair of FFLFs represents the same physical feature, generally no two nodes of the FFLFs correspond to the same physical point (Vassilaki et al. 2012). Homologous point pairs have to be computed (matched). The matching is complicated by the fact that the two FFLFs are of different dimensionality. Furthermore, a single pair of FFLFs may not represent the entire dataset, as it may be confined to a small region of the data. In this case, multiple pairs or a network of corresponding FFLFs are required. The network of FFLFs increases the robustness but it introduces problems such as the FFLFs correspondence and simultaneous matching.

3 Overview of the Method

A novel method based on the iterative closest point (ICP) algorithm (Besl & McKay 1992, Zhang 1994) is used for accurate and robust global matching of heterogeneous FFLFs. The method was initially introduced in Vassilaki et al. (2008) in order to match 2D heterogeneous FFLFs with a rigid transformation. It was further expanded in order to match FFLFs of different dimensionality (2D-3D) with non-rigid projective transformation and it is fully documented in Vassilaki et al. (2012). For efficiency convenience and user friendliness the method has been incorporated into ThanCAD (Stamos 2007), an open-source CAD.

The matching of two FFLFs of the same dimensionality (2D-2D, 3D-3D) is done by an iterative process of determining closest point pairs between the two FFLFs and then using them to compute the transformation parameters by least squares adjustment (LSA). Closest point pairs are computed by splitting a FFLF to a large set of consecutive interpolated points, each one very close to its previous and its next point. Then, the distances of all these points to a node of the other FFLF are computed; the closest points between the two FFLFs are the points with the smallest distances (Vassilaki et al. 2008). This process may be computationally expensive but it is doable with modern computers, and it is relatively easy to speed it up with a divide-and-conquer approach. Further acceleration can be achieved by parallel computing. In the case of FFLFs of different dimensionality (3D-2D) the 3D nodes of the 3D FFLF are projected to the 2D image space using an initial or previous approximation of the transformation parameters, saving the association of each 3D node and its 2D projection. Closest points are computed as in the 2D-2D case, they are converted to 3D-2D pairs through the saved association, and they are used to compute the projection parameters by LSA. Detailed description of the 3D-2D case can be found in Vassilaki et al. (2009b).

The ICP algorithm needs a good initial approximation to converge, which means that the two FFLFs must be close enough to each other. Automatic pre-alignment is done using the rigid similarity transformation computed exploiting physical properties of the two FFLFs, or using the non-rigid first order polynomial (affine) transformation computed exploiting characteristic statistical properties of the FFLFs (Vassilaki et al. 2012).

In the case of networks of FFLFs of the same dimensionality (2D-2D, 3D-3D), the correspondences of FFLFs must be established. Assuming that the two datasets are initially pre-aligned, a FFLF of the first dataset corresponds to the FFLF of the other dataset which is “closest” to it. However, the definition of “closest” is ambiguous for a FFLF which may span many other FFLFs. Different nodes of the same FFLF may be closest to nodes of different FFLFs. Vassilaki et al. (2009a) introduced the term “distance” as an integral measure of how far or how different two FFLFs are. The pair of FFLFs which have the smallest “distance” are assumed to correspond to each other. Four candidates for the “distance” measure were suggested: the Euclidean distance between characteristic homologous points, namely the first nodes (d1), the last nodes (dN), or the centroids (d), and the absolute difference of the FFLFs lengths (ΔS). For robustness, the biggest of these four values can be used as the “distance” of the FFLFs. In the very unlikely case that the “distance” is ambiguous (almost the same) for two or more pairs of FFLFs, application of the full ICP can be used to determine which FFLFs correspond. ICP is the best and more robust approach, but it is very time consuming and should be avoided if possible. Vassilaki et al. (2009a) assumes global matching, implying that the number of FFLFs of each network is the same (N:N matching). The method, however, has partial matching capabilities which can be further exploited in future research. If the number of FFLFs of the two networks is different (N:M matching, N>M), the method finds the “closest” pair of FFLFs, marks these FFLFs as corresponding pair of FFLFs, and continues the same process for the remaining FFLFs (N-1:M-1 matching). In the end N-M single FFLFs are left, which are effectively ignored. The method proceeds with an ICP step, it brings the FFLFs even closer and it then re-evaluates the correspondence, as the “distance” between any pair of FFLFs is obviously changed. Furthermore, ICP automatically rejects two unrelated FFLFs marked to correspond if they do not overlap (Vassilaki et al. 2012).

All the FFLFs of a network share a common transformation. In order to compute the common transformation, LSA is applied to all the pairs of FFLFs simultaneously. The equations produced by the homologous points of all pairs of FFLFs are assembled into the same LSA matrices. The LSA computes the transformation which best fits all the pairs of FFLFs. The computed transformation brings the FFLFs closer together, and thus the correspondences are re-evaluated, in case that the previous, poorer, transformation led to a few false correspondences.

4 Matching 3D-2D Networks of FFLFs

In this paper, the method is extended to match networks of FFLFs of different dimensionality (3D-2D). Vassilaki et al. (2009a) used manually pre-aligned datasets, which in the case of 3D-2D is rather difficult as the varying Z coordinate of the 3D FFLFs and the geometrical distortions on the SAR image (2D FFLFs) makes their shapes incompatible. Manual move, scale and rotate operations, typically provided by CAD software, are not enough to cancel the elongated nature of the SAR projection. Instead, the 3D FFLFs must be projected to the image space of the 2D FFLFs using a good approximation of the unknown transformation parameters. Since all pairs of FFLFs share the same transformation, a good approximation of the projection of a single pair of FFLFs, automatically computed (Vassilaki et al. 2012), can be used to project the 3D FFLFs network to the 2D image space, bringing the datasets close together. The 2D-2D correspondences of FFLFs are then established as outlined in the previous section (Vassilaki et al. 2009a). The correspondence of the single pair of FFLFs is chosen manually by the user, so that in this sense the datasets are manually pre-aligned. However, apart from this, the procedure is fully automated.

5 Design of Tests, Datasets and Results

The experiments were designed to test both the robustness and the accuracy of the method. In the group "A" tests, the georeferencing of a TerraSAR-X image is done using GCI extracted from a 40 years old map. In the group "B" tests, the georeferencing of the same TerraSAR-X image is done using recently captured optical aerial images. Group A is a "crash test" of the proposed method and it may be of practical use in urgent situations when there is no time to collect up-to-date GCI (a day is enough for a satellite scene under the current method). Group B, where the datasets are compatible in terms of time, is a contribution to the operational georeferencing of TerraSAR-X images and their combined use with optical images. Each group of experiments contains tests corresponding to four distinct projection models:

- first order 3D PFs with 8 unknown parameters (3D affine model)

$$\begin{aligned} x &= a_1X + a_2Y + a_3Z + a_4 \\ y &= b_1X + b_2Y + b_3Z + b_4 \end{aligned} \tag{1}$$

- second order 3D PFs (Toutin 2004) with 16 unknown parameters

$$\begin{aligned} x &= a_1X + a_2Y + a_3Z + a_4 + a_5X^2 + a_6Y^2 \\ &\quad + a_7Z^2 + a_8XY \\ y &= b_1X + b_2Y + b_3Z + b_4 + b_5X^2 + b_6Y^2 \\ &\quad + b_7Z^2 + b_8XY \end{aligned} \tag{2}$$

- 3D Direct Linear Transform (DLT) with 11 unknown parameters

$$\begin{aligned} x &= \frac{a_1X + a_2Y + a_3Z + a_4}{c_1X + c_2Y + c_3Z + 1} \\ y &= \frac{b_1X + b_2Y + b_3Z + b_4}{c_1X + c_2Y + c_3Z + 1} \end{aligned} \tag{3}$$

- 3D RPFs with 14 unknown parameters

$$\begin{aligned} x &= \frac{a_1X + a_2Y + a_3Z + a_4}{c_1X + c_2Y + c_3Z + 1} \\ y &= \frac{b_1X + b_2Y + b_3Z + b_4}{d_1X + d_2Y + d_3Z + 1} \end{aligned} \tag{4}$$

where X,Y,Z are the object coordinates and x,y are the SAR image space coordinates (slant range and azimuth, respectively).

The study area is a sub-urban area in the greater north-eastern region of Athens, Greece. It has steep mountainous terrain with an average elevation of 270 m and it is generally covered by sparse vegetation. It also includes two small urban regions. The data used (Fig. 1) are a) a TerraSAR-X image, b) an old medium scale topographic map and c) a recent medium-scale optical orthoimage with the corresponding digital terrain model (DTM) (Fig. 1). The TerraSAR-X image is a single look slant range (SSC) imaging product captured in 2009 with the experimental 300 MHz high resolution spotlight (HS) imaging mode. The whole image, which covers an area of about 50 km^2 (5 km x 10 km), was used. The polarisation is HH and the scene centre incidence angle is 53°. The projected spacing values for range (ps_{Rg}) and azimuth (ps_{Az}) are 0.45 m and 0.87 m, respectively. The medium-scale old topographic map was in analogue form at a scale of 1:5000 and it was compiled by stereo-restitution from aerial photos, captured in 1970. The map was converted to digital form by scanning and digitizing; the contour lines of the map were converted to a DTM. The relative planar accuracy of the map is estimated at 1.5 m (σ_{Pl}). The vertical contour interval is 4 m and the relative vertical accuracy is estimated at 2 m (σ_Z). The absolute planar accuracy is estimated at 2.5 m (σ_{Pl}) and the vertical one at 4 m (σ_Z). The map is part of a series which cover the whole territory of Greece. They are readily available and they cost much less than a GPS survey. The area under study changed widely during the 40 years between the two data collection phases (2009 and 1970). The area used to be an agricultural area but now exhibits a great variety of land uses, as it serves as a holiday resort near Athens. The map and the SAR image share few common features, as is the case in virtually all multitemporal datasets. Most of these common features are roads which, unfortunately, evolve through time. Many sections of the roads have changed considerably during the 40 years between data acquisitions, as it is shown in the next paragraphs. The maps are topographic maps (not illustration or small scale maps) of scale 1:5000 and thus there are no (significant) generalizations. Furthermore, most of the used roads are more than 6 m wide, which means that their width is over 1 mm in the map, and thus, the roads are clearly visible. The medium-scale optical orthoimage and the corresponding DTM were available with pixel sizes of 0.5 m and 5 m, respectively. The orthoimage was produced from aerial images collected in 2008. The orthoimage and the corresponding DTM are of the same nominal accuracy as the paper map, probably because they were also compiled to meet the specifications for mapping at a scale of 1:5000.

The 3D error of the datasets must be expressed in terms of the 2D SAR image space, since this paper deals with the projection (in

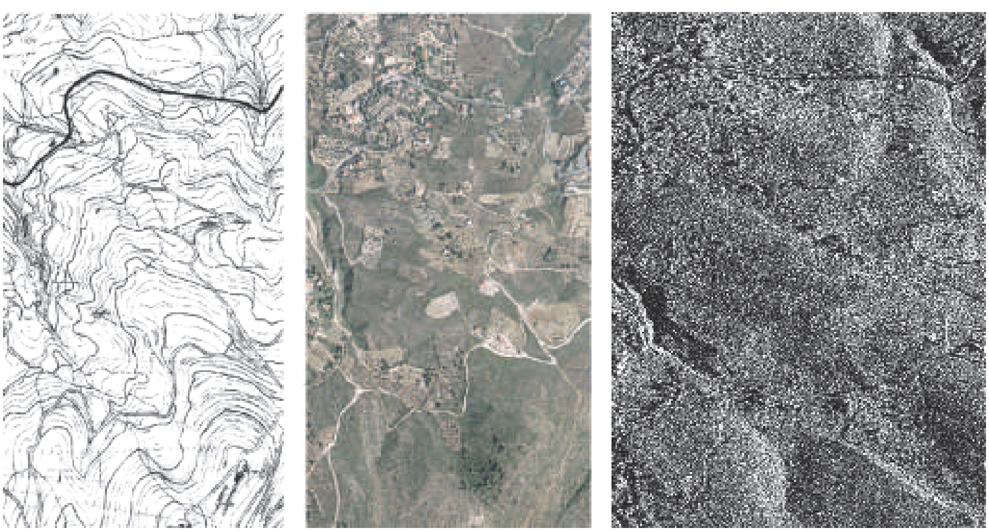


Fig. 1: Datasets: old map (left), orthoimage (middle), SAR image (right).

the mathematical sense) of the 3D object space to the 2D SAR image space. The error propagation of the object space to the SAR image space is given in Appendix 1. The planar relative error (σ_{Pl} = 1.5 m) and the vertical relative error (σ_Z = 2 m) of the 3D object space map are propagated to 1.4 m in the SAR range direction ($\sigma_{Rg,pix}$) and 0.9 m in the SAR azimuth direction ($\sigma_{Az,pix}$), while the planar absolute error (σ_{Pl} = 2.5 m) and the absolute vertical error (σ_Z = 4 m) are propagated to 2.7 m in the SAR range direction ($\sigma_{Rg,pix}$) and 1.7 m in the SAR azimuth direction ($\sigma_{Az,pix}$). These values are valid for the specific SAR image of this paper and are used later for evaluation purposes.

In all tests, the extraction of 2D road edges was done manually by digitizing the lines in the map or in the optical image and in the SAR image. Most of the roads digitized are paved, as they appear better in the SAR images. The road centre lines were obtained from the edges using a skeletonization technique based on the method proposed in Vassilaki et al. (2012). The road centre line is preferred to

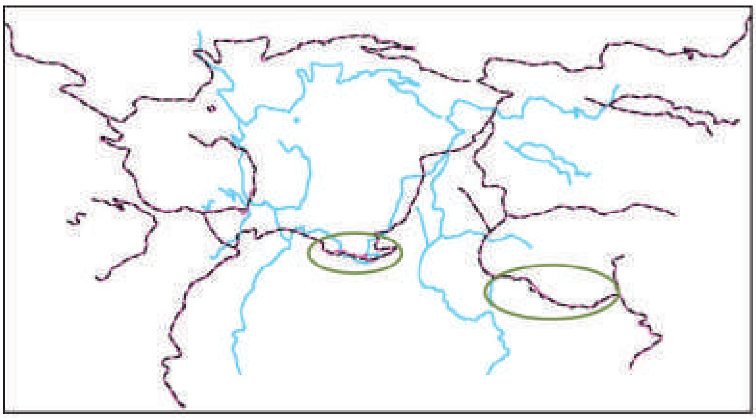


Fig. 2: Matching results of Test A0. 3D GCLFs extracted from the old map (cyan), 2D GCLFs extracted from the SAR image (magenta), matched projection of the 3D GCLFs (black dashed). Fig. 3 shows the road sections inside the green ellipses in a magnified version (GCLF = ground control linear features).

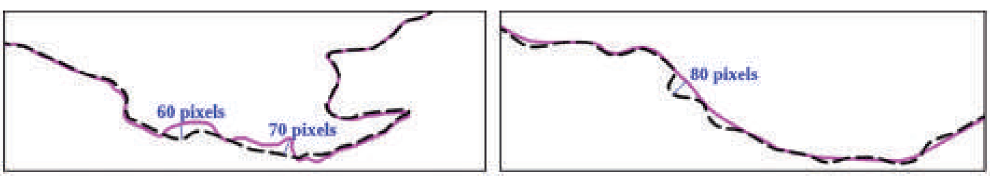


Fig. 3: Test A0: Road sections with wide temporal changes. 2D GCLFs (magenta), matched projection of the 3D GCLFs (black dashed).

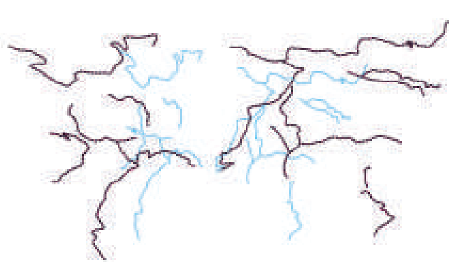


Fig. 4: Matching results of Test A1. GCLFs with wide temporal changes are excluded.

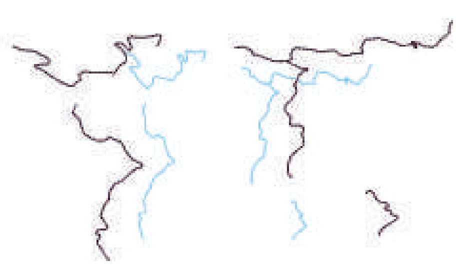


Fig. 5: Matching results of Test A2. Only few GCLFs are used.

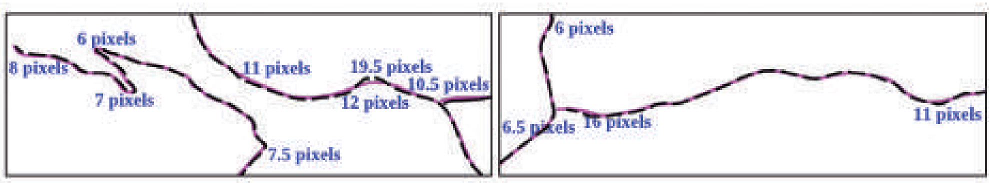


Fig. 6: Test A1: Road sections with temporal changes. 2D GCLFs (magenta), matched projection of the 3D GCLFs (black dashed).

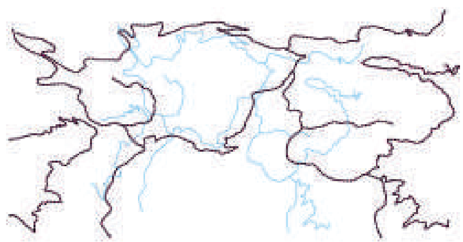


Fig. 7: Matching results of Test B1. The 3D GCLFs are extracted from a recent orthoimage.

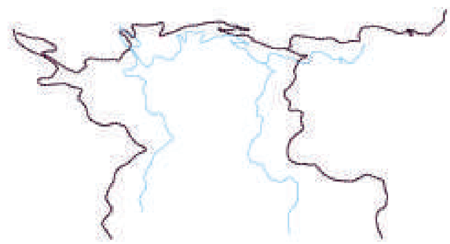


Fig. 8: Matching results of Test B2. Only few GCLFs are used.

road edges as control FFLF, because it is more accurate than the edges and it fully represents the geometry of the road. For all the tests, a preliminary georeferencing was computed using a single pair of corresponding FFLFs that was identified manually. The computed transformation parameters were used to bring the network of FFLFs close together.

In tests A, the length of the road centre lines varies from a few hundred meters to 8.5 km (Fig. 2 and Figs. 4–5). The height profiles of the road centre line are in general rough, due to the scale of the map and the fact that the map does not have elevation information along the surface of the road. The centre line elevation, interpolated in the DTM, was determined by the heights of the surrounding terrain. Three different cases were tested. In the first case (A0) the centre lines of 14 roads were used as ground control linear features (GCLFs). The whole lengths of the roads were used, regardless of the temporal changes that were identified. In the second case (A1) the same 14 centre lines as in (A0) were used, but the sections of the roads which exhibited large temporal changes were eliminated manually as identified by (A0). In the third case (A2) the centre lines of 4 roads, which span the whole scene of the SAR image, were used as GCLFs. The matching results are shown in Fig. 2 and Figs. 4–5. The 3D GCLFs which were extracted from the map appear in cyan colour, the SAR 2D GCLFs appear in magenta colour, and the matched projection of the 3D GCLFs using the proposed method appear with a black dashed line. It must be noted that the coordinates of 3D GCLFs (cyan) and 2D GCLFs (magenta) were too different to fit into the same figure and thus the 3D GCLFS are shown scaled/translated for illustration purposes. In Fig. 3 characteristic sections of the roads with wide temporal changes are presented for case A0. In Fig. 6 characteristic sections with temporal changes are presented for case A1.

In tests B, the length of the road centre lines varies from 1.5 to 12 km (Figs. 7–8). Two different cases were tested. In the first case (B1) the centre lines of 9 roads were used as GCLFs. In the second case (B2) the centre lines of 3 roads, which span the whole scene of the SAR image, were used as GCLFs. The matching results are shown in Figs. 7–8.

6 Validation, Evaluation and Discussion of the Results

The accuracy of the computed georeferencing was checked with independent check points (CPs). In test A, 16 CPs were extracted from the medium scale old map and were also identified in the SAR image. Fig. 9 shows the distribution of the CPs and Tab. 1 shows the RMSE (root-mean-square error) (in m) for the four computed transformation models (1–4). The RMSE is the root-mean-squared differences between the image coordinates of the CPs and the image coordinates of the CPs computed from their object coordinates using the computed transformation model. It is noticeable that the RMSE of the CPs is almost the same regardless of the transformation model used. The differences (1.1 m at most) are not statistically significant as they are less than the accuracy of the maps (1.4 m in range and 0.9 m in azimuth directions) and the uncertainty of the point location on the SAR image is larger than 1 pixel. It is also noticeable that the method is insensitive to temporal changes, given the abundance of GCLFs found in all cases (A0, A1 and A2). The method manages to match robustly and efficiently data which contain sections with gross temporal changes, producing low RMSE. In the first case (A0) the GCLFs contain gross errors in various segments with temporal changes, but they cancel out. In fact, the method can also be used to identify the erroneous segments and exclude them from the matching, but this remains a subject of future research. The erroneous segments were removed manually in the second case (A1). After running tests A0, the distance of some segments of the matched roads were far larger than the rest. The length of these segments was small compared to the length of the roads. If the difference was greater than a threshold (more than 3 times the average difference of the rest) and the length of the segment was greater than 5 times this threshold, the segment was excluded manually. Case A1 tends

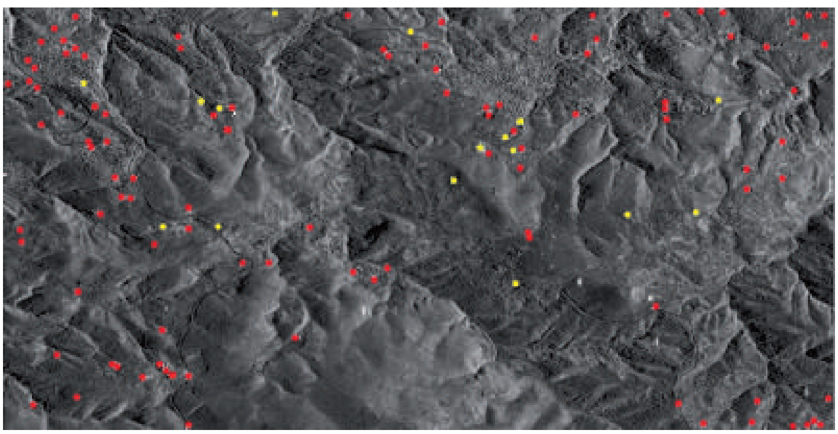


Fig. 9: The CP distribution used for tests A (yellow) and B (red), shown on the SAR image.

Tab. 1: RMSE of the transformation models using CPs (m). The error is dominated by the a-priori error of the GCI and/or CPs (GCI = ground control information, GCLF see Fig. 2).

Test	GCI: GCLFs								GCI: none	
	1st order PFs		2nd order PFs		DLT		RPFs		Physical model	
	dRg	dAz	dRg	dAz	dRg	dAz	dRg	dAz	dRg	dAz
A0	1.9	3.4	2.4	4.1	2.6	3.7	2.3	4.5		
A1	1.8	3.1	2.2	3.2	1.9	3.2	2.0	3.3	2.4	3.5
A2	2.1	3.9	2.2	4.1	1.8	4.1	2.0	3.4		
B1	1.3	2.4	1.1	2.9	1.3	2.5	1.2	2.7	2.3	2.6
B2	1.3	2.6	1.2	3.5	1.5	2.8	1.3	2.9		

to give the best RMSE. In the last case (A2) the number of GCLFs is only 4, but they cover adequately the whole scene, and give almost the same RMSE as (A0).

In tests B, 108 CPs were extracted from the optical image and the corresponding DTM and were also identified on the SAR image. The abundance of salient points which can be identified both on the optical and the SAR image offered the opportunity to collect a great number of CPs in order to improve the reliability. Features such as roads, pools, big buildings and field boundaries are identified clearly on both image types. Metallic objects appear brightly on the SAR image but it was impossible to identify them on the optical image. Built-up areas appear clearly in the optical image, but their interpretation in the SAR image was hard due to the layover and shadow effects. The elevation of features above the terrain surface could not be determined as only ground elevation information was available. Thus, the majority of CPs are corners of pools, field boundaries, and characteristic points of roads. Fig. 9 shows the distribution of the CPs, Fig. 10 shows some of the CPs in detail and Tab. 1 shows the RMSE (in m) for the four computed transformation models (1–4). The conclusions of tests A are verified in tests B: 1) the RMSE of the CPs is almost the same regardless of the transformation model used and 2) the RMSE is not sensitive to the number of GCLFs, if they cover adequately the scene. It should also be noted that the RMSE is very close to the relative accuracy of the orthoimage and the corresponding DTM, computed in Appendix 1 (1.4 m and 0.9 m in the SAR range and azimuth directions, respectively).

The results computed with the present method (GCLFs) tend to be of better quality than those computed by salient point based approaches. In Vassilaki & Ioannidis (2010) a terrain dependent approach was applied to the datasets of tests A using the 3D DLT (11 parameters) and the 1st order 3D RPF (14 parameters) transformations. The RMSE of independent CPs was about 3.2 m in range and about 3.0 m in azimuth direction. About the same results were later achieved by Crespi et al. (2010), who used a terrain independent approach to georeference a SpotLight COSMO-SkyMed image, using 3rd order 3D RPFs (78 parameters, 20 of which proved to be statistically significant). In Nonaka et al. (2008) a digital map at a scale of 1:2500 was used in order to evaluate the accuracy of the orthorectified EEC SpotLight TerraSAR-X products. The accuracy revealed to be better than 5 m in a flat area while it degraded to more than 10 m in mountainous areas.

Comparing the RMSE computed with the present method (GCLFs with empirical models) to those computed with the SAR physical model (without GCI), one might erroneously conclude that the present method is better. However, in both cases the error of the GCI

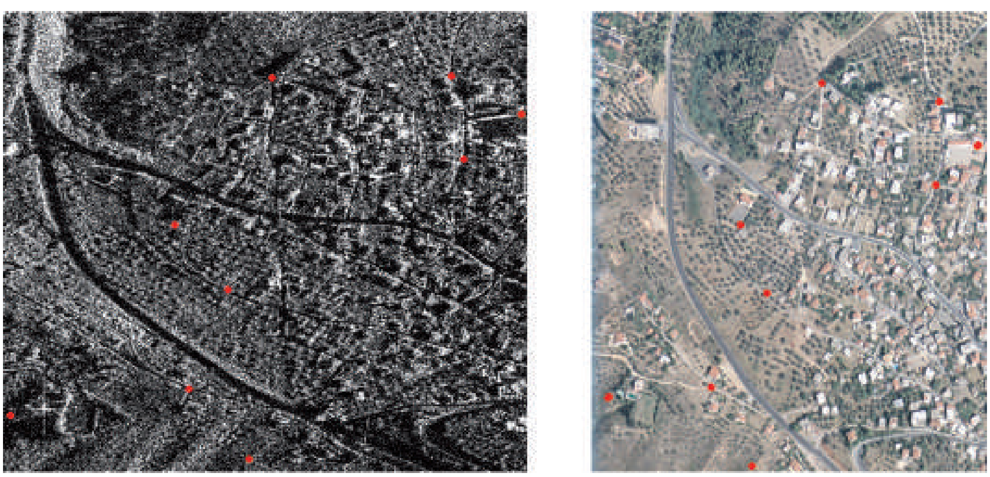


Fig. 10: Some of the CPs used for test B, shown on the SAR (left) and optical (right) image.

and/or the CPs (map, orthoimage, uncertainty of the point location in the SAR image) dominates the (much smaller) error of the sensor. Furthermore, the empirical models reflect the relative accuracy of the GCI, while the physical model reflects the absolute accuracy of the GCI which is worse than the relative one. In Vassilaki et al. (2011a), where the physical model was used with the datasets of tests A, the RMSE of the independent CPs is 2.4 m in range and 3.5 m in the azimuth direction. The absolute accuracy of the map, computed in Appendix 1, was 2.7 m and 1.7 m in the SAR range and azimuth directions, respectively. The RMSE in range is close to the one to be expected, but in azimuth it is somewhat larger. The overall 2D errors (geometric mean of range and azimuth), which are 4.2 m and 3.2 m for our results and the expected accuracy, respectively, are also relatively close. The relatively large RMSE values in azimuth require further investigations. In test B the RMSE of the 108 CPs computed using the physical model is 2.3 m in range and 2.6 m in the azimuth direction. This is close to the absolute accuracy of the orthoimage and the corresponding DTM (the same as in test A) in range and also closer in azimuth than in test A. Again, the overall 2D error is close to the one expected (an overall RMSE of 3.4 m vs. an expected accuracy of 3.2 m). Thus, the results computed with the present method are of the same quality as the ones computed by the physical model without GCI.

7 Conclusions

The pixel location accuracy of TerraSAR-X images has been validated in the past by its operator and by independent researchers, with dedicated projects which demand the installation of corner reflectors on strategic places on the Earth's surface and the acquisition of datasets appropriate for this purpose. Although the accuracy was proven to be well below 1 m, the use of corner reflectors either as control or as check points is generally not desirable and/or feasible in everyday operational cases of research and practice, even more so in emergency situations. The objective of this paper was to evaluate the performance of linear features as GCI for the operational georeferencing of TerraSAR-X images. Research conducted so far, in a sub-urban mountainous area, shows that contemporary satellite SAR sensors and the proposed method is promising with respect to robustness and accuracy: 1) linear features with large temporal changes, due to the 40 years time interval between the data acquisitions, are matched robustly and lead to a reliable recovery of the imaging geometry of the sensor (RMSE on independent CPs: 1.9 m in range and 3.2 m in azimuth), 2) linear features extracted from an optical image and the corresponding DTM are a reliable form of GCI (RMSE on independent CPs: 1.3 m in range and 2.5 m in azimuth).

The RMSE differences between the models (PFs, DLT, RPFs) are inconclusive as all models approach the accuracy of the map. The use of more accurate check points is needed for further evaluation. Automated extraction of the linear features is also expected to serve the goal of further refinement. However the latter will be not a trivial task because the road surfaces may be paved or not, and the material of the pavement may vary across a road's surface due to additional lanes. E.g. the deceleration lane may be paved, semi-paved or not paved at all. These facts differentiate the appearance of a road surfaces in the optical and the SAR images and make it hard to identify the differences, even for the human eye.

Two issues revealed in this paper need further research: 1) the good performance of simple transformations such as the first order PFs when linear features are used as GCI, and 2) the impact of the absolute and relative error of the ground information (CPs, GCPs and GCLFs) on the performance of GCI-based methods as compared to the physical sensor model without GCI.

Appendix 1: Error Propagation from Object Space to SAR Image Space

A point P(X,Y,Z) measured on a map contains a planar error σ_{Pl} and a vertical error σ_Z inherited from the accuracy of the map. There is no reason to assume that the σ_{Pl} error is different in any particular direction so that $\sigma_X = \sigma_Y = \sigma_{PL} / \sqrt{2}$. The directions x, y are taken as the range and azimuth directions of the SAR sensor, respectively. The error σ_Y (m) is also the expected SAR image error in the azimuth direction of the SAR image. The σ_Z, σ_X errors lead to SAR image error in the range direction, which is more complicated due to the slant range geometry of the SAR sensor. The slant distance (S_D) from the sensor (X_0,Y_0,Z_0) to the point (X,Y,Z) is given by (5) where $k = Z/Z_0 << 1$ and φ is the mean incidence angle of all points in the SAR image. The error of the slant distance is then computed by (6). As the pixel spacing is different in the range (ps_{Rg}) and azimuth (ps_{Az}) direction of the SAR image, the error in pixels is computed by (7).

Acknowledgements

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) – Research Funding Program: "Heracleitus II. Investing in knowledge society through the European Social Fund". The author is also grateful to the School of Rural and Surveying Engineering of NTUA for providing the data.

$$S_D = f(Z,X) = \sqrt{(Z_0 - Z)^2 + (X_0 - X)^2} \approx \sqrt{(Z_0 - Z)^2 + Z_0^2 \tan^2\varphi} = Z_0\sqrt{(1-k)^2 + \tan^2\varphi} \qquad (5)$$

$$\sigma_{S_D}^2 = (\frac{\partial f}{\partial Z})^2\sigma_Z^2 + (\frac{\partial f}{\partial X})^2\sigma_X^2 = \frac{(Z_0 - Z)^2}{(Z_0 - Z)^2 + (X_0 - X)^2}\sigma_Z^2 + \frac{(X_0 - X)^2}{(Z_0 - Z)^2 + (X_0 - X)^2}\sigma_X^2 \approx$$

$$\approx \frac{Z_0^2(1-k)^2}{Z_0^2[(1-k)^2 + \tan^2\varphi]}\sigma_Z^2 + \frac{Z_0^2\tan^2\varphi}{Z_0^2[(1-k)^2 + \tan^2\varphi]}\sigma_X^2 \approx \frac{1}{1+\tan^2\varphi}\sigma_Z^2 + \frac{\tan^2\varphi}{1+\tan^2\varphi}\sigma_X^2$$

$$= \cos^2\varphi\sigma_Z^2 + \sin^2\varphi\sigma_X^2 \Rightarrow \sigma_{S_D} \approx \sqrt{\cos^2\varphi\,\sigma_Z^2 + \sin^2\varphi\,\sigma_X^2} \qquad (6)$$

$$\sigma_{Rg,pix} \approx \frac{1}{ps_{Rg}}\sqrt{\cos^2\varphi\,\sigma_Z^2 + \sin^2\varphi\,\sigma_X^2} \qquad \sigma_{Az,pix} \approx \frac{1}{ps_{Az}}\sqrt{\frac{\sigma_{PL}^2}{2}} \qquad (7)$$

References

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

Bresnahan, P., 2009: Absolute Geolocation Accuracy Evaluation of TerraSAR-X Spotlight and Stripmap Imagery – Study Results. – Civil Commercial Imagery Evaluation Workshop, 31 March – 2 April, 2009, USGS, Fairfax Virginia, USA.

Crespi, M., Capaldo, P., Fratarcangeli, F., Nascetti, A. & Pieralice, F., 2010: DSM generation from very high optical and radar sensors: Problems and potentialities along the road from the 3D geometric modeling to the surface model. – IEEE International Geoscience and Remote Sensing Symposium: 3596–3599.

Dare, P. & Dowman, I., 2001: An improved model for automatic feature-based registration of SAR and SPOT images. – ISPRS Journal of Photogrammetry and Remote Sensing **56** (1): 3–28.

Dowman, I. & Dolloff, J., 2000: An evaluation of rational function for photogrammetric restitution. – International Archives of Photogrammetry and Remote Sensing **33** (B3): 254–266.

Eineder, M., Adam, N., Bamler, R., Yaque-Martinez, N. & Breit, H., 2009: Spaceborne Spotlight SAR Interferometry with TerraSAR-X. – IEEE Transactions on Geoscience and Remote Sensing **47** (5): 1524–1535.

Karjalainen, M., 2007: Geocoding of synthetic aperture radar images using digital vector maps. – IEEE Geoscience and Remote Sensing Letters **4** (4): 616–620.

Nonaka, T., Ishizuka, Y., Yamane, N., Shibayama, T., Takagishi, S. & Sasagawa, T., 2008: Evaluation of the Geometric Accuracy of TerraSAR-X. – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **37** (B7): 135–140.

Reinartz, P., Müller, R., Schwind, P., Suri, S. & Bamler, R. 2011: Orthorectification of VHR optical satellite data exploiting the geometric accuracy of TerraSAR-X data. – ISPRS Journal of Photogrammetry and Remote Sensing **66** (1): 124–132.

Roth, A., Huber, M. & Kosmann, D. 2004: Geocoding of TerraSAR-X Data. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **35** (3): 840–845.

Sörgel, U., Cadario, E., Thiele, A. & Thoennessen, U., 2008: Feature Extraction and Visualization of Bridges over Water from high-resolution InSAR Data and one Orthophoto. – IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **1** (2): 147–153.

Stamos, A.A., 2007: ThanCad: a 2dimensional CAD for engineers. – Europython2007 Conference, Vilnius, Lithuania, http://thancad.sourceforge.net/thancad10.pdf, 11 p. (29.2.2012).

Suri, S. & Reinartz, P., 2010: Mutual-information-based registration of TerraSAR-X and Ikonos imagery in Urban areas. – IEEE Transactions on Geoscience and Remote Sensing **48** (2): 939–949.

Toutin, T., 2004: Review Paper: Geometric processing of remote sensing images: models, algorithms and methods. – International Journal of Remote Sensing **25** (10): 1893–1924.

Vassilaki, D.I, Ioannidis, C. & Stamos, A.A., 2008: Registration of 2D Free-Form Curves Extracted from High Resolution Imagery using Iterative Closest Point Algorithm. – EARSeL's Workshop on Remote Sensing – New Challenges of High Resolution **2008:** 141–152, Bochum, Germany.

Vassilaki, D.I., Ioannidis, C. & Stamos, A.A., 2009a: Multitemporal Data Registration through Global Matching of Networks of Free-form Curves. – 2009 FIG Working, Eilat, Israel, cd.

Vassilaki, D.I., Ioannidis, C. & Stamos, A.A., 2009b: Registration of Unrectified Optical and SAR imagery over Mountainous Areas through Automatic Free-form Features Global Matching. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **38** (1-4-7/W5): cd.

Vassilaki, D.I. & Ioannidis, C., 2010: Georeference of High Resolution TerraSAR-X Images with RPFs. – 30th EARSeL Symposium Proceedings: 573–580.

Vassilaki, D.I., Ioannidis, C. & Stamos, A.A., 2011a: Georeference of TerraSAR-X images using Science Orbit Data. – 31st EARSeL Symposium Proceedings: 472–480.

Vassilaki, D.I., Ioannidis, C. & Stamos, A.A., 2011b: Fusion of optical and SAR images in slant range SAR geometry. – 31st EARSeL Symposium Proceedings: 552–560.

Vassilaki, D.I., Ioannidis, C. & Stamos, A.A., 2012: Automatic ICP-based global matching of free-form linear features. – Photogrammetric Record, doi: 10.1111/j.1477-9730.2012.x.

Wegner, J.D., Auer, S. & Sörgel, U., 2009: Accuracy assessment of building height estimation from high resolution optical image combined with a simulated SAR image. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **38** (1-4-7/WS): cd.

Wegner, J.D., Hänsch, R., Thiele, A. & Sörgel, U., 2011: Building Detection From One Orthophoto and High-Resolution InSAR Data Using Conditional Random Fields. – IEEE Journal of selected topics in applied Earth Observations and Remote Sensing **4** (1): 83–91.

Zalmanson, G.H., Doytsher, Y. & Schenk, T., 2004: Towards a new paradigm for registration of SAR images using linear features. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **35** (B3): 41–45.

Zhang, L., He, X., Balz, T., Wei, X. & Liao, M., 2011: Rational function modeling for spaceborne SAR datasets. – ISPRS Journal of Photogrammetry and Remote Sensing **66** (1): 133–145.

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

Address of the Author:

M.-Eng., PhD Candidate, Dimitra Vassilaki, National Technical University of Athens, School of Rural and Surveying Engineering, Athens, Greece, Tel.: +30-21-0772-2653, Fax: +30-21-0772-2677, e-mail: dimitra.vassilaki@gmail.com

Manuskript eingereicht: Oktober 2011
Angenommen: Mai 2012

PFG 2012 / 4, 0421 – 0432
Stuttgart, August 2012

Article

Orientation and Dense Reconstruction from Unordered Wide Baseline Image Sets

Jan Bartelsen, Helmut Mayer, Neubiberg, Heiko Hirschmüller, Oberpfaffenhofen, Andreas Kuhn & Mario Michelini, Neubiberg

Keywords: automatic image orientation, dense 3D reconstruction, unordered image sets, wide baseline, semiglobal matching

Summary: In this paper we present an approach for detailed and precise automatic dense 3D reconstruction using possibly unordered image sets from consumer cameras. Opposed to other approaches we focus on wide baseline image sets. We have combined and improved several methods for robust matching and parameter estimation, particularly, least squares matching, RANSAC, scale-space maxima and bundle adjustment. Point correspondences and the five-point algorithm lead to relative orientation without a need for approximate values. Due to our robust matching method we can orient images under much more unfavourable conditions, for instance concerning illumination changes or scale differences, than it would be possible based on commonly used operators such as SIFT. For dense reconstruction, we use our orientation as input for semiglobal matching (SGM) resulting in dense depth images. The latter can be fused into a 2.5D model for eliminating the redundancy of highly overlapping depth images. However, some applications and acquisition scenarios have a need for full 3D modelling, for which preliminary results are presented. Using small unmanned aerial systems (micro UAS), it is possible to acquire images which have a similar perspective as terrestrial images and can thus be combined with them. Such a combination is useful for almost complete 3D reconstruction of urban scenes. We have applied our approach to blocks of several hundred aerial and terrestrial images, generating detailed 2.5D and 3D models of urban areas.

Zusammenfassung: *Orientierung und dichte Rekonstruktion aus ungeordneten Bildverbänden mit großer Basis*. In diesem Beitrag wird ein Ansatz für die detaillierte und genaue automatische dichte 3D Rekonstruktion auf Grundlage von möglicherweise ungeordneten Bildverbänden, welche mit Consumer Kameras aufgenommen wurden, vorgestellt. Im Gegensatz zu anderen Ansätze zielt der vorgestellte Ansatz auf Bilddatensätze mit großer Basis ab. Dafür wurden verschiedene Methoden, insbesondere Kleinste Quadrate Zuordnung, RANSAC, Maßstabsraum Maxima und Bündelausgleichung, für robuste Zuordnung und Parameterschätzung kombiniert und verbessert. Punktkorrespondenzen und der Fünf-Punkt Algorithmus führen zur relativen Orientierung ohne Bedarf für Näherungswerte. Die verwendete robuste Zuordnungsmethodik ermöglicht es, Bilder unter sehr viel ungünstigeren Bedingungen, z.B. bezüglich Beleuchtungsbedingungen und Maßstabsunterschieden, zuzuordnen, als häufig verwendete Operatoren, wie z.B. SIFT. Für die dichte Rekonstruktion wird die berechnete Orientierung als Eingabe für Semiglobal Matching (SGM) verwendet, mit dessen Hilfe dichte Tiefenbilder bestimmt werden. Diese können, um die Redundanz in den oft hochgradig überlappenden Tiefenbildern zu eliminieren, in 2,5D Modellen fusioniert werden. Einige Anwendungen und Aufnahmekonfigurationen benötigen aber eine volle 3D Modellierung, wofür erste Ergebnisse vorgestellt werden. Mit kleinen Drohnen / Unmanned Aerial Systems (Micro UAS) wird es möglich, Bilder zu erfassen, die eine ähnliche Perspektive auf die Szene haben wie terrestrische Bilddaten und daher mit diesen kombiniert werden können. Eine solche Kombination ist für eine fast vollständige 3D Rekonstruktion von städtischen Szenen sehr hilfreich. Der entwickelte Ansatz wurde auf Blöcke mit hunderten von Bildern aus der Luft und vom Boden angewandt und damit detaillierte 2,5D und 3D Modelle von Siedlungsbereichen erzeugt.

DOI: 10.1127/1432-8364/2012/0128 1432-8364/12/0128 $ 3.00

1 Introduction

Pollefeys et al. (2000) have demonstrated that sets of images from consumer cameras in combination with dense 3D reconstruction form a good basis for photo realistic visualization. Pollefeys et al. (2002) presented one of the first approaches for relative orientation for a larger number of images in a general configuration, i.e., without known approximate values such as for aerial images. By employing uncalibrated images, for which the camera constant / principal distance etc. is not known, the approach is very flexible, yet, on the other hand, reliant on sufficient 3D structure in the scene for the determination of calibration parameters. More recently, Pollefeys et al. (2008) have reconstructed 3D structures from sequences with more than one hundred thousand calibrated images.

For the work above, the overlap of the images is assumed to be known, either implicitly in form of the order in a sequence, or explicitly, e.g. in the form of an aerial flight plan. Schaffalitzky & Zisserman (2002) proposed one of the first methods which can automatically determine the overlap of images in unordered image sets. While Havlena et al. (2010) have proposed an approach which works efficiently for thousands of images based on graph optimization. Agarwal et al. (2009) and Frahm et al. (2010) have recently presented approaches which can deal with hundred thousands or even millions of unordered images from Community Photo Collections from the Internet to model urban areas. A major difference between the two is that the former runs on a cloud, the latter on just one multi-GPU (graphics processing unit) PC system. While both approaches are impressive, one has to note that they are based on certain characteristics of the data and several assumptions to make them tractable:

- At tourist attractions many images are taken from nearly the same spot. Thus, many similar images can be found even when extremely downsampling the images.
- The restricted goal is to reconstruct the obvious 3D structure. This leads to impressive 3D reconstructions of highlights, such as the Colosseum in Rome. Yet, there might be images, possibly with wider baselines, which are not linked, but could be used to extend the geometrical coverage or connect the tourist attractions. They are not considered, as it would mean a detailed comparison of many more images.

Opposed to the above approaches, we aim at a detailed and more complete modelling by making use also of wide baseline image sets. We assume that we know the calibration of the cameras, e.g. from the Exif (Exchangeable image file format) tags of the images in combination with a database about different cameras.

In section 2 we present the methods for point detection, matching and robust parameter estimation that we have improved and combined for orientation of possibly unordered wide baseline image sets. For dense reconstruction, the results of our orientation procedure are used as input for semiglobal matching – SGM (Hirschmüller 2008), which we are about to extend to 3D (section 3). We found that due to our precise relative orientation, very good depth estimates were possible also for wider baselines.

Section 4 gives results. We have processed blocks consisting of hundreds of images acquired from small unmanned aircraft systems (micro UAS). For a combination of UAS images with terrestrial images we have generated a preliminary dense 3D reconstruction of a building comprising the roof as well as the facades. In addition, we present preliminary results for dense 3D surface reconstruction from terrestrial images only. Finally, section 5 gives conclusions and discusses future work.

2 Orientation of Unordered Image Sets

This section is split into two parts: In the first part, we describe our orientation procedure. Contrary to state-of-the-art approaches, it relies on given information on overlap between images, e.g. in the form of a sequence. Yet, it is suitable for wide baseline image sets and it provides very accurate results by consequent propagation of covariance information through all steps.

In the second part, we present a preliminary approach for overlap detection for unordered image sets. Due to restrictions of the fast matching approach (Wu 2007) it is based on, it can only deal with short baselines for the time being.

2.1 *Orientation of Wide Baseline Image Sets*

With the scale invariant feature transform (SIFT) Lowe (2004) has presented a powerful solution for the estimation of point correspondences mainly for short baselines. Yet, reliable matching of points for wide baseline images is much harder and, thus, there is a need for improved matching methods. Our approach to wide baselines is based on scale invariant point matching, least squares matching and robust bundle adjustment (Fig. 1).

The developed approach for point matching produces reliable results even in case of major scale differences as well as viewpoint and illumination changes. It is based on normalized cross correlation (NCC), which is relatively invariant against the latter, but only weak compared to the former two. NCC is weak against scale changes, because in this case image patches with the same size in pixels around corresponding points contain different scene parts. To deal with scale differences, we make use of the work of Lindeberg (1994) and determine points in the form of scale space maxima for Differences of Gaussians (DoG). With the information on scale, we downsample image patches with higher resolution, so that they match to the same scene part. Potential correspondences are refined by least squares matching using an affine geometric model (Grün 1985). This results in subpixel relative point positions. Matched points in two and more images and their covariance information are employed for relative orientation of pairs, triplets and image sets.

With the five-point algorithm (Nistér 2004), one can directly compute the relative orientation from calibrated image pairs. That is, no approximate values for the orientation are needed regardless of the geometric configuration of the two images during acquisition. We have embedded a version of the five-point algorithm into RANdom SAmple Consensus – RANSAC (Fischler & Bolles 1981). This allows us to deal with the large proportion of outliers of possibly more than 80 % typical when matching wide baseline images. By using the geometric robust information criterion (GRIC) (Torr & Zisserman 1997), we compare models not only based on the number of

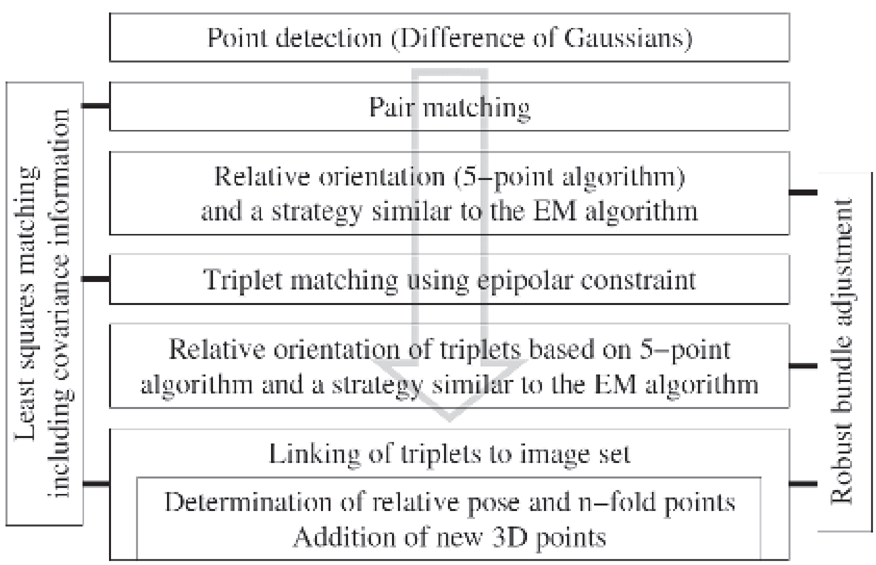


Fig. 1: Image orientation based on scale invariant point matching, least squares matching and robust bundle adjustment.

inliers as in standard RANSAC, but we also take the distance from the ideal solution, i.e., the epipolar line for image pairs, as well as the estimated covariance into account.

As we empirically found that incorrect models can be evaluated very similarly as correct models even when using GRIC, we refine the solution based on a strategy similar to the expectation maximization (EM) algorithm (Bartelsen & Mayer 2010). Partial solutions which initially stem from RANSAC are extended by alternating between robust bundle adjustment with all current inliers (maximization) and the determination of possibly new inliers for the adjusted solution (expectation). This strategy is employed for pairs as well as triplets, with robust bundle adjustment using the covariance information from image matching and reweighting in the form of an M-estimator (Huber 1981) at its core.

For the following reasons, we use triplets as basic geometric building block for image sets:

- While pairs of points can only be checked in one dimension by means of their distance from their respective epipolar lines, triplets permit an unambiguous geometric checking. This does not only result in much more reliable points, but also leads to improved, more reliable information for the cameras.
- Triplets can be directly linked into larger sets by determining their relative pose (translation, rotation and scale) from two common images.

When determining matching points for triplets, we make use of the information derived for image pairs by restricting the matching to a corridor around the epipolar lines. The relative orientation for triplets is computed by taking one image of a triplet as reference, twofold application of the five-point algorithm for the reference image and the other two images, and finally robust determination of the relative scale between the two pairs.

The construction of the relatively oriented image set starts with selecting one triplet. Triplets are linked to the image set based on common image pairs. One common image allows the propagation of translation and rotation. The distance to the second image provides scale. Matched points are propagated from the image set to the linked triplet by means of least squares matching leading to n-fold points, i.e., points that can be seen in $n > 3$ images. New 3D points from the linked triplet are added. Finally, robust bundle adjustment is employed to improve the accuracy, but also the reliability by eliminating wrong matches that could not be detected in the triplets due to the limited redundancy.

Based on the highly reliable relative orientation thus derived, we have shown in (Bartelsen & Mayer 2010) how to calculate the absolute orientation from unreliable and imprecise GPS data of low cost sensors, e.g. in a GPS camera, also in areas with strong occlusions, e.g. cities, similarly to Strecha et al. (2010). Taking other information for absolute orientation, such as ground control points, into account is even easier, as one does not have to deal with a possibly larger number of gross errors as for GPS cameras.

2.2 *Overlap Determination for Unordered Image Sets*

To deal with unordered image sets, we employ automatic overlap detection consisting of the following steps:

- determination of similarity between images,
- construction of a two-view graph and
- construction of a three-view graph.

A fast GPU implementation (Wu 2007) of SIFT is used for detecting points and correspondences by pair-wise image matching. The two-view matching graph consists of images as nodes, whereas its edges connect similar images. The weight of an edge, i.e., the image similarity, is assumed to be given by the number of correspondences between connected images. Images with the number of correspondences below a threshold will be considered as dissimilar and no edge is inserted in the matching graph. To reduce the complexity, available GPS information in the Exif tags of the images is used to derive the distance between images and thus to sort out unlikely pairs.

Once the similarities between the images have been derived, we determine a connected image set by constructing the maximum span-

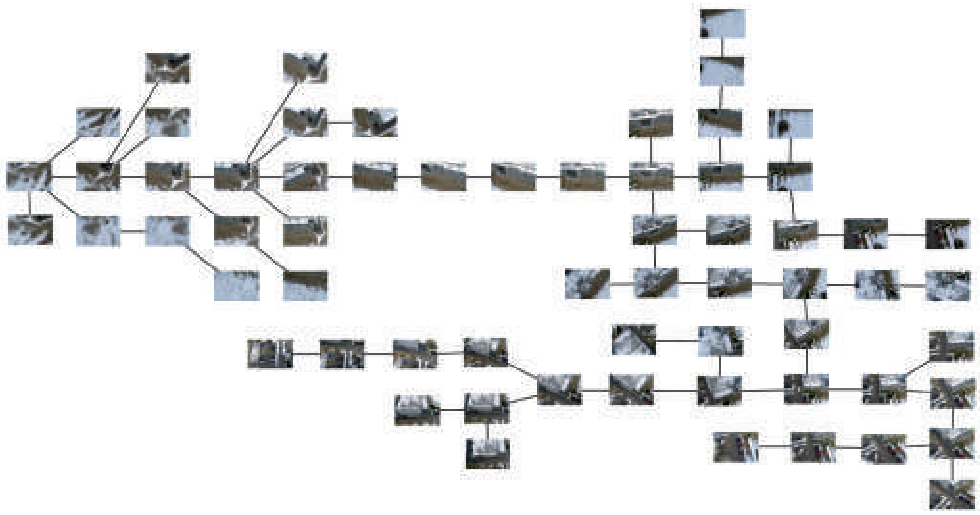


Fig. 2: Minimum spanning tree (MST) for image pairs.

ning tree (MST) of the two-view matching graph using the modified algorithm of PRIM (1957). Fig. 2 is an example, for which results are presented further below in Fig. 6.

Finally, triplets are derived from the connected image set by iterating through the MST using the depth-first traversal algorithm. We discard triplets with the images having a number of correspondences or a normalized overlap area below a threshold (Fig. 3). For the determination of the overlap area between the images of triplets, we calculate the convex hull of correspondences between all three images using the algorithm of SKLANSKY (1982).

The state concerning unordered image sets is still preliminary and, thus, it is only used for the example presented in Fig. 6. Many pairs which can be oriented by our robust matching method described in section 2.1 are not found due to the limited capability of the employed fast matching method (WU 2007). Therefore, whereas we can orient wide baseline combinations of images from the air and from ground, the detection of overlapping images has to be conducted manually at the moment in this case. The integration of our capabilities for wide baseline matching into the overlap determination for unordered image sets is our most important goal for the future.

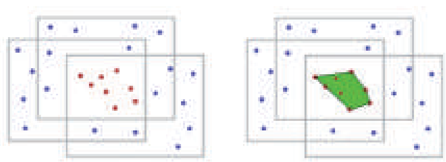


Fig. 3: Number of common points (red – left) and area (green – right) for overlap determination between three images.

3 Dense Reconstruction

For dense reconstruction semiglobal matching – SGM (HIRSCHMÜLLER 2008) is employed. It is based on

- mutual information (MI) or the census filter for cost computation and
- the substitution of a 2D smoothness term by a combination of 1D constraints (semiglobal).

MI presents the conditional probability distribution for the intensities in the matching image given an intensity in the reference image, without resorting to a parametric model. Thus, MI can compensate a large class of global radiometric differences. Because the conditional probability is computed for the whole image, problems can arise for local radiometric changes, e.g. if materials with very different reflection characteristics exist in the scene or lighting conditions change.

In Hirschmüller & Scharstein (2009) the census filter was found to be the most robust variant for matching cost computation. It defines a bit string with each bit corresponding to a pixel in the local neighbourhood of a given pixel. A bit is set if the intensity is lower than that of the given pixel. Census thus encodes the spatial neighbourhood structure. A 7×9 neighbourhood can be encoded in a 64 bit integer. Matching is conducted via computing the Hamming distance between corresponding bit strings.

The smoothness term of SGM punishes changes of neighbouring disparities (operator $T[]$ is 1 if its argument is true and 0 otherwise):

$$E(D) = \sum_{\mathbf{p}} \Big(C(\mathbf{p}, D_{\mathbf{p}}) + \sum_{\mathbf{q} \in N_{\mathbf{p}}} P_1 \mathrm{T}[|D_{\mathbf{p}} - D_{\mathbf{q}}| = 1] + \sum_{\mathbf{q} \in N_{\mathbf{p}}} P_2 \mathrm{T}[|D_{\mathbf{p}} - D_{\mathbf{q}}| > 1] \Big). \quad (1)$$

- The first term consists of matching costs for all disparities of disparity image D.
- The second term adds a constant penalty P_1 for all pixels $\mathbf{q}$ from the neighbourhood $N_{\mathbf{p}}$ of $\mathbf{p}$ for which the disparity changes only slightly (1 pixel).
- The third term adds a larger constant penalty P_2 for larger disparity changes. Because it is independent of the size of the disparity change, it preserves discontinuities.
- As discontinuities in disparity are often visible as intensity changes, P_2 is calculated depending on the intensity gradient in the reference image (with $P_2 \geq P_1$).

Global minimization in 2D is NP complete for many discontinuity preserving energies $E(D)$. Opposed to this, in 1D, minimization can be done in polynomial time via dynamical programming. The latter is usually applied within image lines. Because the solutions for neighbouring lines are computed independently, this typically leads to streaking. For the semiglobal solution, 1D matching costs are computed in different, (practically 8) directions, which are aggregated without weighting. In the reference image, straight lines are employed, which are deformed in the matching image.

By computing the disparity images D for exchanged reference and matching image one can infer occlusions or matching errors by means of a consistency check. If more than one pair with the same reference image is matched, the consistency check is conducted for all pairs only once (Hirschmüller 2008).

With SGM, very dense disparity maps having one disparity per image pixel can be computed. Using the orientation parameters, all points can be projected into 3D space, leading to dense 3D point clouds. While the original work of Hirschmüller (2008) has shown how to derive 2.5D surface models, work on the derivation of 3D surface models by means of triangulation of the 3D points, dealing also with outliers, has been started only recently.

Because modelling large-scale scenes fully 3D can produce billions of points, efficient processing with regard to the computational and memory costs is a must. We consider octrees to be very suitable for this purpose. Hence, we use a triangulation based on balanced octrees for meshing (Bodenmüller 2009). Besides removing redundancy, octrees are particularly useful for visibility-checks in multiple-view geometry.

Before mesh generation, normal vectors are determined from the neighbours of a point, and points with uncertain normal vectors are eliminated. The triangle mesh is built incrementally. Iterating through all remaining points, the temporary mesh is projected on the tangent plane in a neighbourhood of a

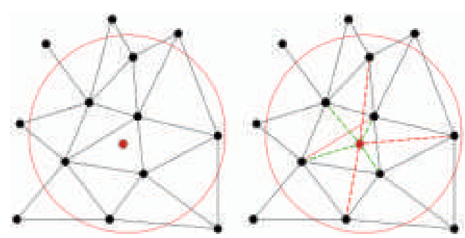


Fig. 4: Local update of triangulation by adding a new vertex to a temporary mesh. Left: New vertex *v* (red) and projected neighbourhood. Right: New candidate edges (dashed lines). Green lines were accepted and red lines were removed because of intersection with shorter edges.

new point (Fig. 4). The new point is connected with all vertices within the neighbourhood. If a new edge intersects an old edge in the plane, the longer one is removed.

4 Results and Discussion

We have applied our approach for orientation and dense 2.5D and 3D reconstruction to several image sets. As our focus is on wide baseline scenarios, we have manually determined the overlap for all but the results in Fig. 6. For the latter, we demonstrate the potential of our

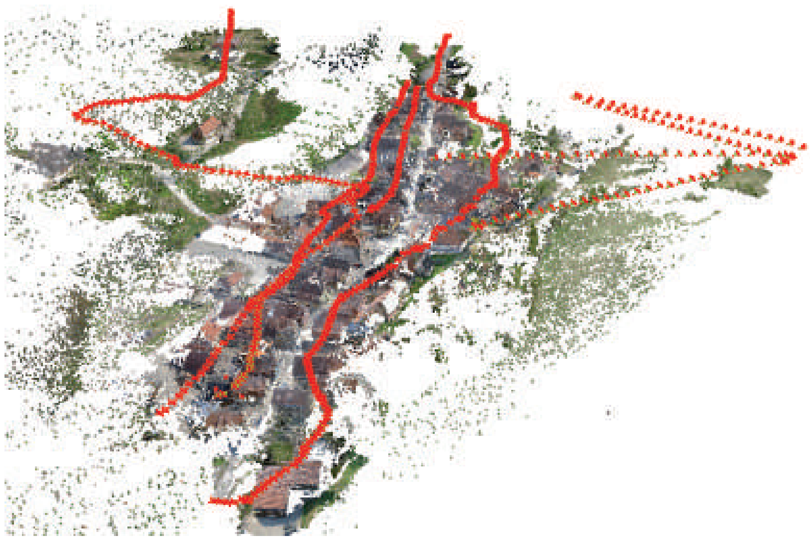


Fig. 5: 3D points and cameras (red pyramids) for a model generated from more than 600 images from a micro unmanned aircraft system (micro UAS).

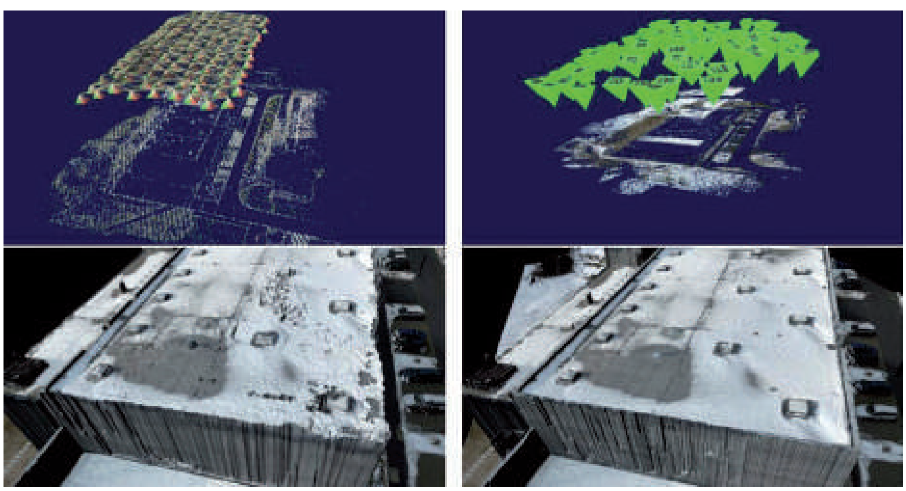


Fig. 6: Top row: 3D points and cameras (pyramids) for a set of 166 images of a large building in Wessling. Bottom row: Resulting 2.5D models. Left column: Result for Bundler (Snavely 2010). Right column: Result derived by our approach – the more precise relative orientation allows for a much more detailed scene reconstruction.

approach for unordered image sets introduced in section 2.2 also in comparison with Bundler (Snavely 2010). In all cases, the census filter (section 3) was used for cost computation for SGM.

Fig. 5 presents 3D points and camera positions (red pyramids) for more than 600 images acquired by a micro unmanned aircraft system (micro UAS) from about 50 m above the ground. Orientation was possible in spite of the lack of approximate values for exterior orientation of this highly non-regular flight configuration. Particularly, for overlapping areas of the flight strips, images with wide baselines could be matched, leading to a more stable geometry.

The result in Fig. 6 is based on 166 images, acquired by a micro UAS. Although the flight was controlled automatically, the obtained image set is not very well structured. Because of too small overlap, many triplets could not be matched. For this image set, we have compared our approach with Bundler (Snavely 2010). In particular, we found that the relative orientation produced by Bundler is not very precise and the 3D point cloud contains many obviously false points. Thus, SGM could only be applied in a meaningful way after downsampling the images to half the original resolution. In contrast, the relative orientation obtained by our approach is much more precise (Tab. 1) and could be used as basis for SGM on

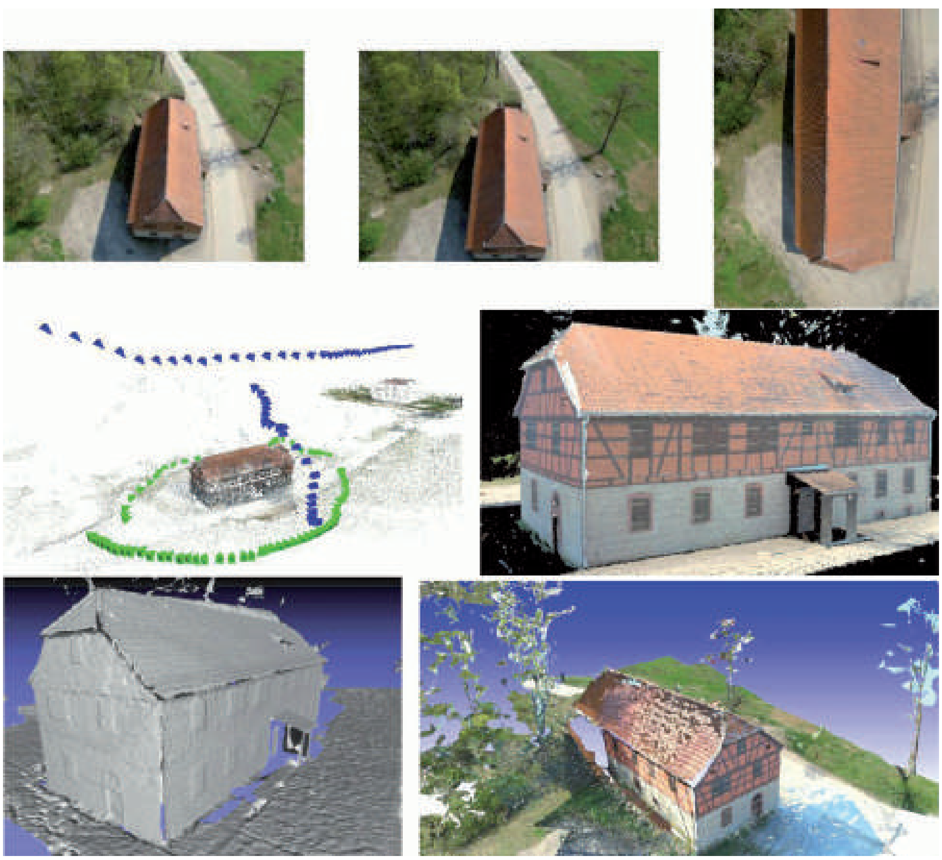


Fig. 7: Top row: Two images from the aerial and one from the 'ascending' sequence which could be matched and oriented. Middle row left: 3D points and cameras derived from a combination of images from the air (blue pyramids) and from the ground (green pyramids) of a building near Hammelburg. Middle row right: 3D point cloud generated by SGM from the flight and the terrestrial sequence. Bottom row: Preliminary result for dense 3D surface reconstruction – shaded with wire frame as well as textured.

the original resolution, leading to a more detailed and realistic 2.5D model.

Fig. 7 presents preliminary results for our new approach for 3D surface reconstruction based on a combination of images from UAS and from ground. The set of 205 images contains three different sequences (Fig. 7 centre row, left):

- The flight sequence was taken from about 20 m above ground.
- The terrestrial image-sequence was acquired around one building.
- The 'ascending' sequence connects images from the air and from the ground. The images change in small steps from the aerial to the ground perspective. This image configuration is only feasible for micro UAS, which can be flown very close to facades and roofs.

The combination of the flight and the 'ascending' sequence is quite difficult, because of major scale differences in combination with small overlap and perspective distortion. Fig. 7 top row shows an example for an image triplet which could be matched. The dense 3D point cloud (Fig. 7, centre right) was generated from images from the terrestrial and the flight sequences (with the relative orientation of both sequences determined using the 'ascending' sequence). It illustrates that roof and walls exactly fit to each other (please note the roof overhang) and thus, that our relative orientation is very precise. Finally, the bottom row of Fig. 7 shows our preliminary 3D surface reconstruction as a shaded view and also as a textured model. The colour variations of the roof texture are caused by the different lighting conditions during acquisition of the ground and the flight sequences.

Fig. 8 shows another preliminary result of our work for full dense 3D surface reconstruction. We have applied our approach to the image sets 'fountain-R25' and 'castle-R20'

Fig. 8: Top: Textured preliminary result for 3D surface reconstruction of the facade of Ettlingen castle. Bottom left: Dense 3D point cloud. Bottom right: Shaded 3D surface with wire frames. The results base on the image-sets 'fountain-R25' and 'castle-R20'of (Strecha et al. 2008).

of Strecha et al. (2008). Due to our robust matching approach, a combination of both sets was possible.

Finally, Tab. 1 presents the average back projection errors σ_0 as well as computation times for all examples. For all experiments one PC with an Intel Xeon processor with four cores with 2.3 GHz and a 2GB Nvidia Geforce GTX 285 graphics card was used. All back projection errors for our approach are around or below 0.3 pixels, with the majority in the range of 0.15 pixels. Although Bundler (Snavely 2010) can produce quite accurate results too, it had severe problems with our challenging wide baseline dataset, reflected by the large reprojection error (0.44 pixels for half resolution) and leading to the inferior 2.5D model presented in Fig. 6.

The computation times for our approach for all but the first experiment are reasonable. Concerning the first experiment we note that it was computed with the sequential version of our orientation algorithm, where robust bundle adjustment is computed for every linked image triplet. This is not necessary and we are, thus, about to replace this by a hierarchical solution, where image sets are combined, which is much more efficient for large sets. For the dataset of Fig. 6, images downsampled to half resolution (i.e., 2.5 Mpixel instead of 10 Mpixel) were used for Bundler for reducing the number of SIFT features. Still, Bundler required longer to finish, because it simply tries to match each image with every other image. In contrast, we use the fast matching approach (Wu 2007) for overlap determination and we only try to orient the rather limited number of image triplets for which sufficient overlap could be determined.

5 Conclusions and Future Work

In this paper, we have presented an approach for automatic orientation and dense 3D reconstruction from wide baseline image sets. As key characteristics, it aims at a high precision in every step from least squares matching to robust bundle adjustment. Currently, our approach for full 3D reconstruction does not maintain all the details that are available in the high resolution depth images of semi-global matching. We plan to take into account the uncertainty of the 3D points from different pairs as well as smoothness priors. How this can be effectively and efficiently done is subject of our current research. Our ability for very precise relative orientation is of fundamental importance for accurate 3D modelling.

Although our point matching approach is pretty robust against scale and illumination changes, it is still not robust enough concerning viewpoint changes. Currently, arguably the best known concept for matching which is robust concerning viewpoint changes is the approach of Morel & Yu (2009) simulating off-image-plane rotations. This concept has not yet been integrated into an approach for 3D reconstruction, also due to its prohibitive computational complexity. The improvement of our approach by a similar, yet more efficient procedure is also part of our future work.

Acknowledgements

We gratefully acknowledge the support of Dipl.-Ing. Florian Seibel, Dr. rer. nat. Patrick Reidelstürz and Prof. Dr.-Ing. Peter Stütz during the acquisition of the aerial image sets and we thank the reviewers for the helpful comments.

Tab. 1: Average reprojection error σ_0, number (#) images and computing times for orientation of the image sets in the given figures.

	σ_0 (pixels)	# images	Computing time (hours)
Fig. 5	0.12	603	170
Fig. 6: Our approach	0.31	166	5
Fig. 6: Bundler (half resolution)	0.44	166	8
Fig. 7	0.14	205	8
Fig. 8	0.11	47	2

References

AGARWAL, S., SNAVELY, N., SIMON, I., SEITZ, S. & SZELISKI, R., 2009: Building Rome in a Day. – Twelfth International Conference on Computer Vision: 72–79.

BARTELSEN, J. & MAYER, H., 2010: Orientation of Image Sequences Acquired from UAVs and with GPS Cameras. – Surveying and Land Information Science **70** (3): 151–159.

BODENMÜLLER, T., 2009: Streaming Surface Reconstruction from Real Time 3D Measurements. – Dissertation, Technische Universität München.

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

FRAHM, J.-M., GALLUP, D., JOHNSON, T., RAGURAM, R., WU, C., JEN, Y.-H., DUNN, E., CLIPP, B., LAZEBNIK, S. & POLLEFEYS, M., 2010: Building Rome on a Cloudless Day. – Eleventh European Conference on Computer Vision **IV:** 368–381.

GRÜN, A., 1985: Adaptive Least Squares Correlation: A Powerful Image Matching Technique. – South African Journal of Photogrammetry, Remote Sensing and Cartography **14** (3): 175–187.

HAVLENA, M., TORII, A. & PAJDLA, T., 2010: Efficient Structure from Motion by Graph Optimization. – Eleventh European Conference on Computer Vision **II:** 100–113.

HIRSCHMÜLLER, H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. – IEEE Transactions on Pattern Analysis and Machine Intelligence **30** (2): 328–341.

HIRSCHMÜLLER, H. & SCHARSTEIN, D., 2009: Evaluation of Stereo Matching Costs on Images with Radiometric Differences. – IEEE Transactions on Pattern Analysis and Machine Intelligence **31** (9): 1582–1599.

HUBER, P., 1981: Robust Statistics. – John Wiley & Sons, Inc. New York, USA.

LINDEBERG, T., 1994: Scale-Space Theory in Computer Vision. – Kluwer Academic Publishers, Boston, USA.

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

MOREL, J. & YU, G., 2009: ASIFT: A New Framework for Fully Affine Invariant Image Comparison. – SIAM Journal on Imaging Sciences **2** (2): 438–469.

NISTÉR, D., 2004: An Efficient Solution to the Five-Point Relative Pose Problem. – IEEE Transactions on Pattern Analysis and Machine Intelligence **26** (6): 756–770.

POLLEFEYS, M., VERGAUWEN, M. & VAN GOOL, L., 2000: Automatic 3D Modeling from Image Sequences. – International Archives of Photogrammetry and Remote Sensing **33** (B5/2): 619–626.

POLLEFEYS, M., VERBIEST, F. & VAN GOOL, L., 2002: Surviving Dominant Planes in Uncalibrated Structure and Motion Recovery. – Seventh European Conference on Computer Vision, Volume **II:** 837–851.

POLLEFEYS, M., NISTÉR, D., FRAHM, J.-M., AKBARZADEH, A., MORDOHAI, P., CLIPP, B., ENGELS, C., GALLUP, D., KIM, S.-J., MERRELL, P., SALMI, C., SINHA, S., TALTON, B., WANG, L., YANG, Q., STEWÉNIUS, H., YANG, R., WELCH, G. & TOWLES, H., 2008: Detailed Real-Time Urban 3D Reconstruction from Video. – International Journal of Computer Vision **78** (2–3): 143–167.

PRIM, R., 1957: Shortest Connection Networks and some Generalizations. – Bell Systems Technical Journal **36:** 1389–1401.

SCHAFFALITZKY, F. & ZISSERMAN, A., 2002: Multiview Matching for Unordered Images Sets, or "How Do I Organize My Holiday Snaps?". – Seventh European Conference on Computer Vision **I:** 414–431.

SKLANSKY, J., 1982: Finding the Convex Hull of a Simple Polygon. – Pattern Recognition Letters **1:** 79–83.

SNAVELY, N., 2010: Bundler: Structure from Motion for Unordered Image Collections. – phototour.cs.washington.edu/bundler (9.5.2012).

STRECHA, C., VON HANSEN, W., VAN GOOL, L., FUA, P. & THOENNESSEN, U., 2008: On Benchmarking Camera Calibration and Multi-View Stereo for High Resolution Imagery. – Computer Vision and Pattern Recognition: 1–8.

STRECHA, C., PYLVÄNÄINEN, T. & FUA, P., 2010: Dynamic and Scalable Large Scale Image Reconstruction. – Computer Vision and Pattern Recognition: 358–365.

TORR, P. & ZISSERMAN, A., 1997: Robust Parametrization and Computation of the Trifocal Tensor. – Image and Vision Computing **15:** 591–605.

WU, C., 2007: SiftGPU: A GPU Implementation of Scale Invariant Feature Transform (SIFT). – cs.unc.edu/~ccwu/siftgpu (9.5.2012).

Addresses of the Authors:

Dipl.-Inf. JAN BARTELSEN, Prof. Dr.-Ing. HELMUT MAYER, Dipl.-Inf. ANDREAS KUHN, M.Sc. MARIO MICHELINI, Bundeswehr University Munich, Institute of Applied Computer Science, D-85577 Neubiberg, Tel.: +49-89-6004-3429, Fax: +49-89-6004-

3906, e-mail: {jan.bartelsen}{helmut.mayer}{andreas.kuhn}{mario.michelini}@unibw.de

Dr. Heiko Hirschmüller, Dipl.-Inf. Andreas Kuhn, Department of Perception and Cognition, Institute of Robotics and Mechatronics, German Aerospace Center (DLR), e-mail: heiko.hirschmueller@dlr.de

This article is a modified version of a paper accepted for the ISPRS Congress 2012 in Melbourne, Australia.

Manuskript eingereicht: März 2012
Angenommen: Mai 2012

PFG 2012 / 4, 0433 – 0442
Stuttgart, August 2012

Article

Overview of Experimental Setups in Spectroscopic Laboratory Measurements – the SpecTour Project

András Jung, Christian Götze & Cornelia Glässer, Halle (Saale)

Keywords: spectral laboratory measurements, portable spectrometers, experimental setup, SpecTour

Summary: The spectral signatures differ in many ways; even for the same material the captured data is significantly influenced by data collection method, equipment and other parameters influencing the general experimental setup. Spectral measurements are primary information sources in hyperspectral remote sensing. Portable device-based spectroscopy has significantly accelerated high resolution data production and the risk of uncontrolled error propagation. The number of spectral library producers and users is increasing. The growing challenges of spectral libraries initiated the project SpecTour. SpecTour networks portable spectrometer users in order to provide an overview of their measurement routines, equipment, general attitudes and, most of all, the spectra. The project requires active attendance by its participants and is performed in a round-robin approach. The philosophy of the project is "Just measure like you always do!". The illumination parameters, workflow or other technical parameters are not predetermined, but the participants have to fill out a setup protocol. More than 36 spectrometers from Europe, Israel and USA have been involved in the experiment (last update: March 2012). All of the instruments are commercially available. The primary results show that the differences between the normalized spectra for the same reference panels are over 20 % and the quality of the spectra vary considerably. Evidence from the present and realistic measurement situation can be used to quantitatively highlight the lack of standards. This paper gives a general overview of the measurements and data acquisition techniques. The results of the setup protocols are evaluated and discussed here. The project is still running (SpecTour 2012) and is open to anyone who is an active field spectrometer user and would like to share spectra and practice.

Zusammenfassung: *Bei spektrometrischen Messungen werden oftmals für gleiche Objekte unterschiedliche Ergebnisse der Spektralsignaturen ermittelt.* Dabei können die Probenahme, die Laborausstattung und andere Parameter den Aufbau sowie den Ablauf des Experimentes beeinflussen. In der hyperspektralen Fernerkundung sind spektrale Messungen wichtige primäre Informationsquellen. Aufgrund der steigenden Anzahl von Nutzern und der dabei entstehenden spektralen Bibliotheken wird auch das Risiko von nicht aufgespürten Fehlern erhöht. Vor diesem Hintergrund ist das Hauptziel des Projektes „SpecTour" das Networking von Gruppen mit Spektrometern, um einen Überblick über ihre Messroutinen, ihre Ausstattung, die Gewohnheiten und die entstandenen Spektren zu schaffen. Das Projekt, angelegt als Ringversuch, benötigt eine aktive Beteiligung aller Teilnehmer. Die Philosophie des Projekts ist dabei „just measure like you always do!", d.h. die Beleuchtungsparameter, der Messaufbau sowie -ablauf werden nicht festgelegt. Alle Parameter werden aber in einem Messprotokoll festgehalten. Mehr als 36 im Handel verfügbare Spektrometer aus Europa, Israel und den USA sind bisher in dem Experiment zum Einsatz gekommen (Stand: März 2012). Die ersten Ergebnisse zeigen, dass zwischen den zu messenden Referenzmaterialien eine Variabilität von über 20 % auftritt. Ebenso sind Qualitätsunterschiede festzustellen. Die aktuelle Messsituation innerhalb der Community zeigt das Fehlen von Normen und Standards deutlich. Dieser Artikel präsentiert einen Überblick über die Auswertung der Messprotokolle aller bisherigen Teilnehmer. Das Projekt ist erreichbar unter SpecTour (2012). Jedem aktiven Nutzer von Feldspektrometern bieten wir die Möglichkeit, an diesem Experiment teilzunehmen.

DOI: 10.1127/1432-8364/2012/0129 1432-8364/12/0129 $ 2.50

1 Introduction

Hyperspectral laboratory and field measurements were mostly made by expert groups until recent years. Profound changes can be recognized to date in hyperspectral remote sensing, which are characterized by the increasing number of spectrometers on the market and a growing scientific community. This community shows variability in technical and scientific background regarding the users' experience, equipment and applications. Due to this development nowadays we have a mixed user community ranging from beginners to experienced groups. Additionally, there are new and less experienced companies on the market for hyperspectral remote sensing.

Spectral records are stored and managed in databases such as the spectral database of the USGS or by working groups in own databases called spectral libraries. The quality is influenced and affected by many parameters. Exemplary studies investigated individual parameters and their influence on the spectral signature. The variability of the field of view (FOV) of spectral signals was analyzed by Caras et al. (2011). For reflectance spectra reference panels are usually used. A comparison of different reflectance standards were investigated by Sanches et al. (2009). It is of high importance that standards have a stable reflectance throughout the optical region. In the study of Castro-Esau et al. (2006) significant differences were found between different spectrometers depending on measurement and illumination geometry. Field spectrometers with full-range detectors show sensitivity changes following temperature changes (Markham et al. 1995). Detectors are often characterized by 'steps' between ranges which most frequently appear at the 'overlapping' regions (Milton et al. 2009). These can be reduced by a longer warm-up time. The GER3700 spectroradiometer (Brown et al. 2001) shows the changes of 'steps' in a three hour warm-up experiment. An analysis of the realistic technical environment in remote sensing laboratories does not exist yet. A "five-spectrometer-experiment" carried out under the same conditions in geometry and illumination showed that post correction methods could improve the comparability of the reflectance spectra (Jung et al. 2010, 2011).

Library spectra should be used as a reference for identification with no doubt to their quality and reliability. A spectral library is as good as its spectra and data structure. The development of spectral databases like SPECCHIO (Bojinski et al. 2003) shows also the need of standardized spectral libraries.

Our approach is focusing on the spectrometer user and the individual experimental setup. The object of our investigation was to analyze and compare spectrometric data using the same reference materials by different spectrometer users, because less attention has recently been paid to monitor and evaluate the individuality of the spectral measurements.

The analysis focuses on easy-to-capture technical parameters such as lamps, experiment setup and age-of-device. For the sake of simplicity, the laboratory seemed to be the most stable and suitable setting for the experiment. Furthermore, the concept of the experiment and initial conclusions with practical experiences are presented as well.

At this point it is important to note that development of a metadata structure, or best practice on how to archive spectral data, was not our intention, and a technical comparison or inter-calibration of the spectrometers was also not the focus of this work. In the literature there are comprehensive research studies focusing on calibration issues or database development with proven techniques and successful implementations (Schaepman & Dangel 2000, Hueni et al. 2009, Milton 1987, Milton et al. 2009).

SpecTour is a project of the German Society of Photogrammetry, Remote Sensing and Geoinformation (DGPF) and the Martin-Luther-University Halle-Wittenberg that investigates and analyses the variability of the measurement environments in laboratories in order to learn more about the real situation of spectral measurements in daily work. The round-robin test is not a tool to completely compensate low quality spectra or inappropriate measurement conditions (Price 1994, 1998). The goal of this paper is to communicate the present situation, a better understanding of the experimental setups and influences on spectral signals. Finally, we would like to minimize the knowledge gaps of spectrometer users.

2 Materials and Methods

This experimental is a round-robin test that involves different scientific groups to measure the same objects (four reference panels and one rock sample) in their own laboratory und environment. The participants vary from sophisticated experts to recently formed groups. The philosophy is "Just measure like you always do!". It was of high importance during the round-robin test to reflect and document the reality of the spectrometric laboratory measurements without any preliminary technical instructions but with the same reference panels. The test started in 2009 and it is still ongoing.

Up to now, the project has attracted 25 participants from six countries. The number of members is increasing since the project is ongoing and the community is growing. For more information about the participants, please visit SpecTour (2012).

The spectrometers deployed here are from industry leading manufacturers such as Analytical Spectral Devices (USA), FOSS NIRSystems (USA), LI-COR Biosciences (USA), Ocean Optics (USA) und Tec5 (Germany). The instruments vary in resolution, age, spectral range, technology and performance (Tab. 1). All of them were commercially available and active spectral library producers.

Tab. 1: Short technical overview of the involved spectrometers.

Spectrometer	Spectral range (nm)		Light source	
	350–1000	1000–2500	external	internal or contact
ASD FieldSpec Pro JR	**X**	**X**	**X**	
ASD FieldSpec Pro FR	**X**	**X**	**X**	**X** (with contact probe)
ASD FieldSpec 3	**X**	**X**	**X**	**X** (with contact probe)
ASD Handheld	**X**		**X**	
Tec5 HandySpec	**X**		**X**	
Ocean Optics HR2000+	**X**			**X**
FOSS XDS Rapid Content Analyzer	**X**			**X**
Li-Cor Li1800	**X**		**X**	

Tab. 2: Basic technical documentation of the experiment.

Basic setup information	Distance between reference and measurements unit (a) Distance between reference and the source of illumination (b) Angle between *a* and *b* (α)
Properties of the spectrometer	Name (serial number, type, version) Year of purchase Start of use Last calibration (when and where)
Properties of the lamp	Name (serial number, producer, type) Power (W) and voltage (V) values Operation since Total hours of operation
Administrative information	Name of the operator Date and place Affiliation and head of unit

In Tab. 2 we see which parameters were asked for documentation. These data were of high importance in understanding the experimental setups. Only those data were additionally recorded in the protocol that could not directly be derived from the spectra. Simplicity was important because of limited time and readiness. Consideration and documentation of all technical aspects of the measurement would have been beyond the capacity (of both the project and the participants) and would have been out of the original scopes as well.

The calibrated reference panels provide the most valuable details on the individuality of the measurements. The white reference measurement was taken with a calibrated 90 % reflectance panel, and this became the master reference for all further targets and samples. Each spectral measurement was repeated four times for the same target and saved separately. Before reflectance values were generated, the raw data or digital counts were also saved. The normalized reflectance factors were calculated and used for analysis.

Milton et al. (2009) noted that reflectance data would remain a convenient method to represent the energy interactions occurring at the surface. This approach was followed in this work and used for data interpretation. It is important to note that the real quantity acquired by the used spectrometers was the reflectance factor. The reflectance factor is the ratio of reflected and incident flux, irradiated under the same conditions (Nicodemus et al. 1977).

Each involved spectrometer collected 20 reflectance spectra (4 white, 4 grey, 4 black, 4 uncalibrated reference material and 4 chlorite) plus 4 reflectance spectra for the individual background materials. When chlorite was

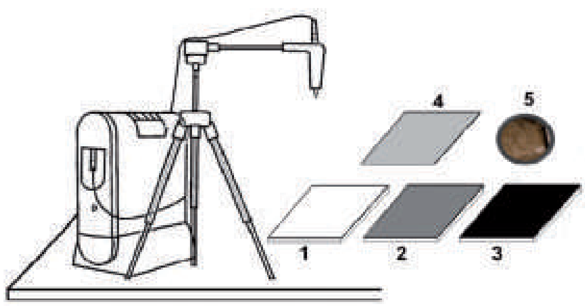


Fig. 1: Schematic illustration of the panels and targets with a field spectrometer in the laboratory, 1–3: calibrated panels, 4: uncalibrated reference panel, 5: chlorite.

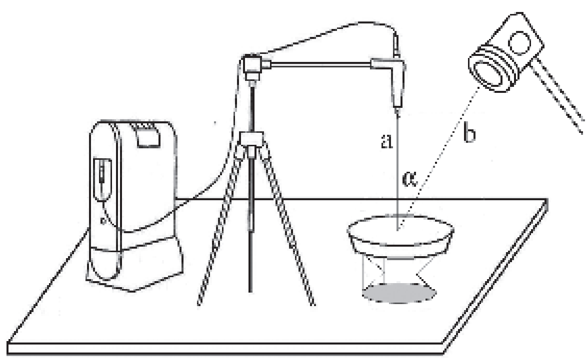


Fig. 2: Geometric parameters of the experiment protocol: a = distance between reference and measurements unit, b = distance between reference and the source of illumination, α = angle between a and b.

measured the size of the sample was often less than the FOV of the spectrometer and a spectrum-neutral background was needed, which was a challenge in many cases (Fig. 5).

The schematic illustration of the panels and targets with a spectrometer in a laboratory set-up can be seen in Fig. 1. Fig. 2 shows the basic geometric parameters that were documented in the protocol (Tab. 2).

3 Reference Materials

Three calibrated reference targets (Fig. 1, 1–3) and two material samples (Fig. 1, 4–5) were measured by each participant. The two samples consisted of an uncalibrated reference target and a chlorite mineral (Chlorite schist, Great St. Bernard, Switzerland, $(Mg,Al,Fe,Mn)_3[(OH)_2Si_4O_{10}]*3Mg(OH)_2$). The calibrated targets had 5 %, 20 % and 90 % absolute reflectance.

The absolute reflectance values for the reference targets were defined in the following way: The hemispherical spectral reflectance was measured for the calibrated reference targets and the calibration was performed with a standard from U.S. National Institute of Standards and Technology (NIST), serial number 2044a-01-15. The absolute reflectance was determined by using a Perkin-Elmer Lambda 19 UV-VIS-NIR spectrometer (serial number: 1260) equipped with a 150 mm PTFE (Polytetrafluoroethylene) sphere, certified by the National Metrology Institute of Germany (PTB, Braunschweig, Calibration PTB 4.52-0208). The random error for the reference targets was 0.006 for 250–400 nm, 0.005 for 400–1100 nm, 0.006 for 1100–2200 nm and 0.010 for 2200–2500 nm. The size of the calibrated reference panels was 200 x 200 mm². The uncalibrated reference target shows typical reflectance properties in the visible spectral range and very characteristic features both in the NIR and SWIR (near and short wave infrared). Its size was also 200 x 200 mm². The chlorite sample (Fig. 4) with an average size of 50 x 50 mm² was an ideal mineral for a spectrometric round-robin test because it is robust and has easy-to-find absorption peaks and stand-alone features. The second reason to work with this sample was its small size. This pre-condition (50 x 50 mm²) was a challenge to the participants (because of FOV) and forced them to use individual background solutions, which strongly affected the final outcome and enriched the experiment with valuable information (Fig. 5).

The reference curves are presented (Fig. 3) to have basic spectral information on the reference panels and targets. It provides an illustration of the characteristics of the six curves (white, grey, black, uncalibrated, chlorite and background). The participants had to deliver these kinds of spectra after having finished the individual experiment with their own spectrometer. In Fig. 6 can be seen what kind of variety of the reflectance spectra were captured by individual measurements. As an example, three curves were shown in Fig. 6 to present obvious anomalies.

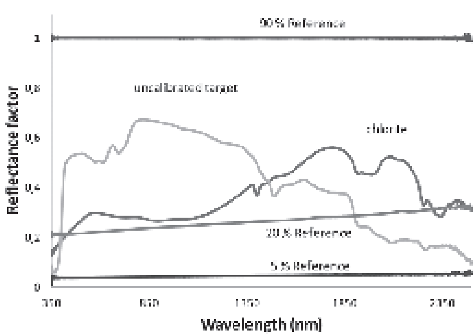


Fig. 3: The reference spectra and chlorite.

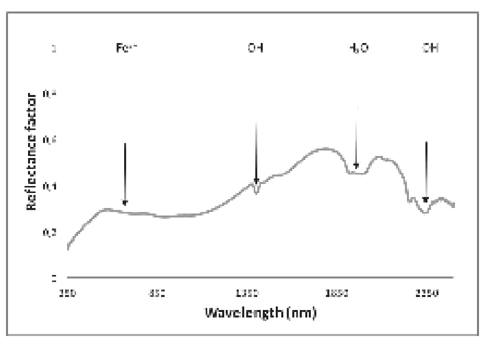


Fig. 4: Chlorite spectrum with absorption peaks.

4 Results

In this paper more attention is paid to general aspects and easy-to-recognize anomalies in the spectra and attitudes in the community. The results will be discussed through examples taken from the SpecTour database. The first one summarises the content of the protocols and highlights the tendencies of the user community. The second example will point out how important it is to be aware of the background materials when the FOV of the spectrometers exceeds the size of the sample. The third example shows typical errors for white reference measurements (Figs. 7 and 8). The SpecTour coordinators checked the spectra and gave feedback in case of anomalies. It helped localize and communicate typical problems and showed the advantages of a networking community.

4.1 *Evaluation of the Protocols*

The protocols of the experiment were of high significance for an understanding of the spectra and for the documentation of the non-spectral properties of a spectrometric measurement. These results came mostly from the measurements' protocol (Tab. 2 and Fig. 2) or from the header part of a measurement file. In Tab. 3 the geometrical properties of the individual measurement setup were summarised. Half of the participants worked with a target-measurement unit distance (P_a) of 21–30 cm, with a target-lamp distance (P_b) of 31–60 cm and with a lamp-to-measurement angle (P_α) of 36–45°. These dimensions document a typical laboratory environment and are achievable without extra investments. Another reason is the analogy to the field measurements which demand the source of illumination being significantly further away compared to the height of the measurement to guarantee diffuse irradiance conditions. In the laboratory real conditions are narrowed to compromises and feasible solutions.

Tab. 3: Geometric properties of the measurements.

Distance between target and measurement optic (cm)	% n=24
Contact (<1)	16
1–5	8
11–15	13
16–20	13
21–30	**50**

Distance between target and light source (cm)	% n=24
Contact (<1)	17
1–30	13
31–60	**50**
61–90	8
91–140	12

Angle between P_a and P_b in degree (°)	% n=25
Contact (<1)	16
15–25	12
26–36	8
36–45	**56**
46–55	8

n: number of elements (instrument or other parameters)
Contact: stands for measurement with distance up to 1 cm

Tab. 4: Auxiliary parameters of the spectra.

Integration time (ms)	% n=29
8	3
17	7
34	7
68	15
136	**62**
272	3
1000	3

Count of samplings before saving a measurement	% n=28
10	4
25	**50**
30	4
40	7
50	25
100	10

Considering the integration time it can be stated that more than 60 % of the participants operate with 136 ms. In many spectrometers the integration time is chosen automatically and manual adjustment is not needed. This avoids saturation and improves comparability. Another observed issue is "the count of samplings before saving". It points out how many measurements (samplings) were made and averaged before one spectrum was saved and shows that the saved spectra have different statistical robustness and half of the participants made 25 samplings before saving one spectrum (Tab. 4). According to the measurement protocol, most of the spectrometers were purchased between 2007 and 2009. This reflects a growing market and scientific interest (Tab. 5).

The number of calibrations is not very meaningful because most of the instruments were bought in the last 5–10 years (Tab. 5). Usually, the manufacturers recommend recalibrating the instrument every 2–3 years. However, as many instruments are quite new, 62 % of them have only been recalibrated once or twice. The frequency of recalibration also depends highly on budget and necessity in public research facilities.

In laboratories, artificial sources of illumination are commonly used. These illumination sources should be stable and robust over a long period of time. In this experiment, queries were made in reference to three parameters about lamps. The parameter "number of lamps" shows that more than 80 % of the participants use only one lamp in the laboratory setup (Tab. 6). One reason for this is that also in nature the sun is the "standalone source" of illumination. In some cases, the number of lamps is increased to minimise directional light effects and multiple shadowing for layered targets like vegetation. More than 60 % of the lamps had a power of 50 W. This power seems to be an appropriate solution that can service the observed 31–60 cm distance (P_b) and does not negatively affect vital or wet organic materials. A logical outcome is the year of the lamp's purchase that correlates with the

Tab. 5: Auxiliary parameters for the spectrometers.

Year of instrument purchase and last calibration	% n=30	% n=27
1991 – 1995	10	---
1996 – 2000	17	---
2001 – 2003	17	7
2004 – 2006	13	15
2007 – 2009	**37**	**59**
2010 – 2011	7	18

Number of calibrations	% n=16
1	**31**
2	**31**
3	13
5	6
6	13
7	6

n: number of elements (instrument or other parameters)

Tab. 6: Properties of illumination sources.

Number of lamps	% n=31
1	**84**
2	7
4	2
6	7

Power of light source (W)	% n=28
6.5	18
50	**63**
150	4
235	4
1000	7
2000	4

Year of purchase for light sources	% n=23
1997 – 2003	22
2004 – 2006	26
2007 – 2009	**43**
2010 – 2011	9

n: number of elements (instrument, lamp or other countable parameters) year of the spectrometer's purchase since most systems were obtained simultaneously and as a unit.

4.2 *Variability of Background Signals*

Multiple effects substantially influenced the chlorite spectra (Fig. 6). At the single spectrum level the integration time and the warm-up time of the instrument often were likely not to be properly addressed, which caused typical jumps (around 1000 nm and 1750 nm) in many spectra (Fig. 6, middle curve). The background and its interactions with the chlorite spectrum must be carefully discussed. The background was an unknown material behind the sample in almost all cases. The spectrometer user was asked to find the appropriate substance on his own in order to properly measure the chlorite sample. The ideal background in spectroscopy is a low reflectance material (<5 %) covering the range of 400–2500 nm (Fig. 5). Deployment of materials not meeting this requirement should be avoided. The example shows that there are no standards for background material (standard deviation of chlorite is 23 %). Unfortunately, not all spectrometer users are aware of this effect and do not pay enough attention to this (Fig. 5). It is evident that an inappropriate background can fundamentally change the reflectance values of a sample. Generally spoken, the spectrometer users need more information about what to consider when the FOV of the spectrometer exceeds the size of the sample.

4.3 *Variability of White Reference Signals*

Most of the white reference measurements seem to be appropriate at the present level of scaling (Fig. 7). Only one spectrum shows un-

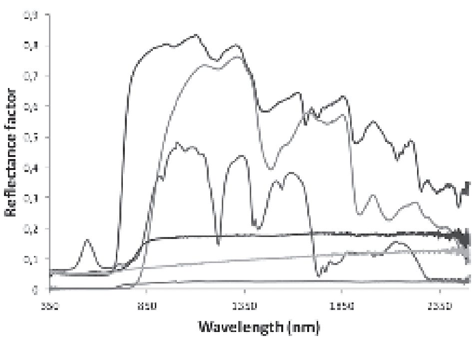


Fig. 5: Characteristics of 6 different background signals.

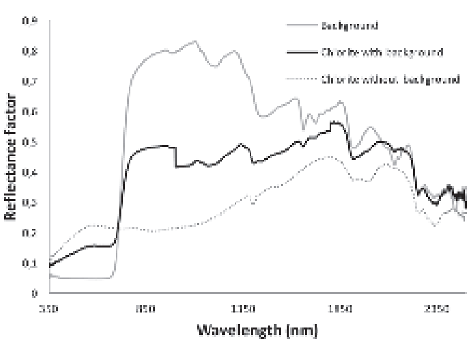


Fig. 6: Effect of background signals on chlorite.

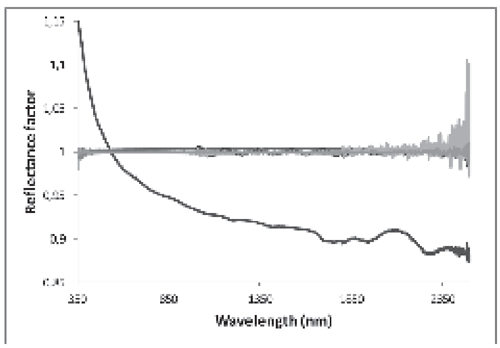


Fig. 7: White reference measurements.

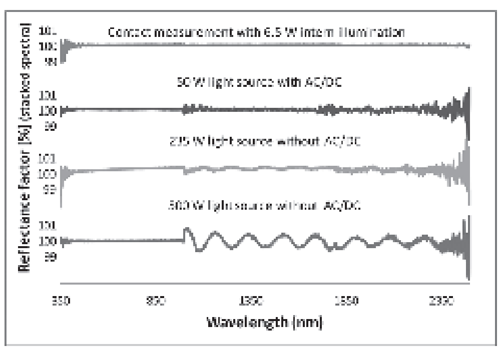


Fig. 8: Influence of light source effects (with or without AC/DC) on the reflectance factor of the white reference panel.

acceptable discrepancies and indicates a fundamental measurement failure.

In Fig. 8 four different reflectance curves are depicted (measurement on white panel) that are typically used by spectrometer users. The 6.5 W and 50 W sources were very common (70 % of the participants), but in the remaining 30 % of the measurements some uncommon phenomena could be observed. Five or six waves can be seen in the first SWIR detector of the spectrometer. The oscillations (sine wave) can be caused by an AC (alternating current) light source (ASD Manual 2007). It is a remarkable phenomenon to observe how electric adapters can influence the spectroscopic data. Laboratory lamps without AC/DC (direct current) adapters have a significant effect on the spectrum that was often overseen or neglected.

After taking the white reference measurement it is recommended to have a detailed look at the measured values to select and avoid malfunctioning signals.

The controlled illumination sources play a significant role in the in-door measurement practise, but they have a lower priority for spectrometer users who take the measurements mostly in the field and are less experienced in the laboratory.

5 Conclusions

This paper gives a brief overview on the diversity of laboratory environment and is more focused on the evaluation of the non-spectroscopic or auxiliary properties of the experiment. Some basic considerations could already be concluded from the present level of the analysis. It became evident that more technical communication is needed among spectroscopic laboratories. For newcomers training courses are necessary and standards or detailed reference books are still missing.

Taking part in the network increases the statistical robustness, helps the mapping procedure and avoids the typical anomalies of measurements.

Spectra must be evaluated regarding quality before entering the final spectral library. This quality is influenced and affected by many parameters, as we could see in this experiment as well. More attention has to be paid to background materials, interactions between white reference measurements and illumination sources. The high variance of background signals and their effects in mineral measurements provided surprising results. The AC/DC converter phenomenon was an interesting, although atypical, observation that can easily be eliminated through the use of a proper adapter. These results encourage us to network more intensively and effectively in the future.

6 Outlook

Our aim is to provide to the community recommendations regarding best practices, guidelines and participation in standardisation studies. Spectra are used as reference and reliability is the highest priority. The lack of measurement environment standards and the significant variability of the results encourage the detailed spectral analysis and interpretation of the parameters. Laboratory work does not cover the whole activity of a user. Field work is just as important as laboratory work and even more complicated. Open access to the results of the project will be available for teaching and training purposes. In the next level of this experiment the SpecTour project will be extended to field measurements.

Acknowledgement

We are very thankful to all colleagues who took part in this project. The calibrated targets and the NIST traceable calibration were provided by SphereOptics GmbH, Germany. The uncalibrated grey panel was provided by courtesy of the Remote Sensing and GIS Laboratory at the Tel Aviv University. The mineral sample (chlorite) was provided by the Institute of Geosciences at the MLU. We would like to express our gratitude to Michael Denk and Florian Beyer from Halle University for their great support. Without their help, support and engagement this project would not have been possible to maintain. The project is still open, and new participants are welcome at any time.

References

ASD Manual, 2007: Quality Spec Manual. – ASD Document 600548 Rev. C.

Bojinski, S., Schaepman, M., Schläpfer, D. & Itten, K., 2003: SPECCHIO: A spectrum database for remote sensing applications. – Computers & Geosciences **2:** 27–38.

Brown, S.W., Johnson, B.C., Yoon, H.W., Butler, J.J., Barnes, R.A. & Biggar, S.F., 2001: Radiometric characterisation of field radiometers in support of the 1997 Lunar Lake, Nevada, experiment to determine surface reflectance and top-of-atmosphere radiance. – Remote Sensing of Environment **77:** 367–376.

Caras, T., Karnieli, A. & Hedley, J., 2011: Exploration field-of-view non uniformities produced by hand-held spectrometer. – Journal of Spectral Imaging **2** (1): 1–11.

Castro-Esau, K.L., Sánchez-Azofeifa, G.A. & Rivard, B., 2006: Comparison of spectral indices obtained using multiple spectroradiometers. – Remote Sensing of Environment **103:** 276–288.

Hueni, A., Nieke, J., Schopfer, J., Kneubühler, M. & Itten, K., 2009: The spectral database SPECCHIO for improved long term usability and data sharing. – Computers & Geosciences **35** (3): 557–565.

Jung, A., Götze, C. & Glässer, C., 2010: White-reference Based Post-correction Method for Multi-source Spectral Libraries. – PFG **2010** (5): 361–368.

Jung, A., Götze, C. & Glässer, C., 2011: A Practical Post-Correction Method for Spectral Libraries. – RS & GIS **2:** 4–12.

Markham, B.L., Williams, D.L., Schafer, J.R., Wood, F. & Kim, M.S., 1995: Radiometric characterization of diode-array field spectroradiometers. – Remote Sensing of Environment **51:** 317–330.

Milton, E.J., 1987: Principles of field spectroscopy. – International Journal of Remote Sensing **8:** 1807–1827.

Milton, E.J., Schaepman, M.E., Anderson, K., Kneubühler, M. & Fox, N., 2009: Progress in field spectroscopy. – Remote Sensing of Environment **113:** 92–109.

Nicodemus, F.E., Richmond, J.C., Hsia, J.J., Ginsberg, I.W. & Limperis, T., 1977: Geometrical Considerations and Nomenclature for Reflectance. – National Bureau of Standards (U.S.).

Price, J.C., 1994: How unique are spectral signatures? – Remote Sensing of Environment **49:** 181–186.

Price, J.C., 1998: An approach for analysis of reflectance spectra. – Remote Sensing of Environment **64** (3): 316–330.

Sanches, I.D., Tuohy, M.P., Hedley, M.J. & Bretherton, M.R., 2009: Large, durable and low-cost reflectance standard for field remote sensing applications. – International Journal of Remote Sensing **30** (9): 2309–2319.

Schaepman, M.E. & Dangel, S., 2000: Solid laboratory calibration of a nonimaging spectroradiometer. – Applied Optics **39:** 3754–3764.

SpecTour, 2012: spectour.org (12.5.2012).

Addresses of the Authors:

Dr.-Ing. András Jung, Institut für Lasertechnologien in der Medizin und Messtechnik, Universität Ulm, D-89081 Ulm, Tel.: +49-731-1429997, e-mail: info@andrasjung.de

Dr. rer. nat. Christian Götze, Prof. Dr. Cornelia Glässer, Institut für Geowissenschaften und Geographie, Martin-Luther-Universität Halle-Wittenberg, Von-Seckendorff-Platz 4, D-06120 Halle (Saale), Tel.: +49-345-5526020/21, Fax: +49-345-5527168, e-mail: {christian.goetze}{cornelia.glaesser}@geo.uni-halle.de

Manuskript eingereicht: März 2012
Angenommen: Mai 2012

PFG 2012 / 4, 0443 – 0454
Stuttgart, August 2012 **Article**

Semi-automatic Integration of Panoramic Hyperspectral Imagery with Photorealistic Lidar Models

Aleksandra Sima, Simon J. Buckley, Tobias H. Kurz, Bergen, Norway & Danilo Schneider, Dresden

Keywords: registration, hyperspectral, TLS, lidar, photorealistic models

Summary: This paper presents a method for increasing the automation of the registration of panoramic hyperspectral images with co-registered conventional digital imagery and a point cloud acquired with a terrestrial laser scanner (lidar), for geological purposes. The SIFT (scale invariant feature transform) interest operator is used to find homologous points between the two imagery types which, because they are recorded at different spectral ranges (visible and short wave infrared) using different geometric projections, differ significantly in appearance. After reducing false matches using RANSAC (random sample consensus), a geometric model for panoramic cameras is applied to retrieve the orientation parameters of the hyperspectral scenes. Once registered, hyperspectral classifications can be combined with the lidar geometry in photorealistic models, allowing material information to be linked to object geometry. Improved automation of the data registration reduces processing time and makes the domain more accessible for non-specialists.

Zusammenfassung: *Semi-automatische Integration von hyperspektralen Panoramabildern mit fotorealistischen Lidar-Modellen*. Mit diesem Beitrag wird eine Methode zur semi-automatischen Registrierung von hyperspektralen Panoramabildern mit konventionellen digitalen Fotos sowie mit den Punktwolken eines terrestrischen Laserscannners (Lidar) beschrieben. Die Methode wird für geologische Fragestellungen angewendet. Es wird gezeigt, dass der SIFT (Scale Invariant Feature Transform) Operator zur automatischen Auffindung homologer Punkte in den verschiedenen Bilddaten trotz der verschiedenen Spektralbereiche (sichtbarer sowie infraroter Bereich) und unterschiedlicher Bildgeometrien (zentralperspektivische Bilder und Panoramabilder mit zylindrischer Projektion) verwendet werden kann. Falsch positive Punkte werden mit der RANSAC (Random Sample Consensus) Methode ausgeschlossen. Unter Anwendung eines Panoramakameramodells können anschließend die äußere und innere Orientierung der hyperspektralen Bilder ermittelt werden. Mit Kenntnis der inneren und äußeren Orientierung können dann einerseits die vermaschten Lidar-Modelle fotorealistisch texturiert und andererseits mit hyperspektralen Klassifikationen geometrisch korrekt überlagert werden, womit eine Verlinkung von geometrischen mit geochemischen Informationen ermöglicht wird. Eine verbesserte Automatisierung der Datenregistrierung verkürzt die Bearbeitungszeit und erleichtert die Datenintegration für fachfremde Anwender.

1 Introduction

The technique of hyperspectral imaging has been used for many years, mostly applied from airborne or satellite-borne platforms. Recent technological advances have given rise to terrestrial hyperspectral imaging sensors, and these are being applied in an increasingly diverse range of laboratory and field settings, including surgery, food safety, surveillance and geology. These hyperspectral sensors record the light reflected from an object in hundreds

 www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0130 1432-8364/12/0130 $ 3.00

of narrow (3–15 nm) spectral bands, allowing creation of near-continuous spectral curves that can be used for the analysis of absorption properties (van der meer 2004), detection of subtle geochemical differences (Clark et al. 1990) or quantitative analysis of pixel composition (Clark & Roush 1984). Hyperspectral imaging is an established method, allowing different materials to be mapped and quantified, even where differences are indistinguishable with the naked eye.

Photorealistic 3D models created from terrestrial laser scanners (TLS, also known as lidar) and digital semi-metric imagery have been widely used in geological applications (Bellian et al. 2005, Buckley et al. 2008), and geomatics techniques are becoming prevalent in the field of outcrop geology. Geological outcrops are exposed cliff sections or quarries that, when captured in 3D, can help to improve the understanding of geometrical relationships between geological features. High-accuracy and high-resolution 3D models of geological outcrops are used to extract statistical data, such as sizes of features, thicknesses of strata, or surface orientation, as well as for interpretation and educational purposes (Bellian et al. 2005, Buckley et al. 2008, 2010). However, mineralogical and lithological mapping is normally limited to interpretation of the three standard spectral bands (red, green, blue; RGB) acquired in the visible spectrum by the conventional digital camera. Many minerals exist that cannot be distinguished using visible light, but that exhibit specific absorption and reflectance properties in the infrared region of the electromagnetic spectrum (e.g. limestone and dolomite; Clark et al. 1990). Therefore, complementing the photorealistic 3D outcrop models with auxiliary spectral data, in the form of hyperspectral imagery, can provide domain experts with additional geochemical information, adding great potential to studies of mineralogy and lithology (Kurz et al. 2012). This goal relies on the accurate registration of the component data types.

In this paper the first results are presented of a semi-automatic method for registration of terrestrial panoramic hyperspectral imagery with data acquired by a laser scanner with an integrated standard digital camera. The resulting orientation parameters of the hyperspectral images establish their position and orientation relative to the point cloud and enable enhancement of the photorealistic outcrop models with classification products. The required processing steps are implemented as a series of Matlab (MathWorks) scripts. As such, automated image registration methods are highly desired to reduce processing time, improve registration accuracy, and make the method more accessible for non-specialists. The data used in this project were acquired for the purposes of geological studies, though the workflow is application-independent, and can be applied analogously in other disciplines.

2 Background

Although lidar systems primarily record geometric data, in the form of the point cloud, many systems record ancillary information pertaining to the strength of the returned laser signal, known as intensity or amplitude. Lidar intensity data can be used to register auxiliary image data in the scanner coordinate system. Furkuo & King (2004) registered conventional photos using tie points between the photos and synthetic images created from the point cloud intensity values. Böhm & Becker (2007) and Gonzalez-Aguilera et al. (2009) registered multiple scan positions on the basis of control points matched between digital photos and the intensity images recorded by a TLS. The success of using lidar intensity for registration purposes is reliant on the signal digitization depth (dynamic range) being of sufficient quality to identify strong features for determining correspondence.

Several successful attempts have been made to integrate terrestrial hyperspectral data with lidar and image data. Kurz et al. (2011) automatically measured reflective targets and manually selected additional natural tie points between the point cloud and hyperspectral image in order to ensure a uniform point distribution. Exterior image orientations and intrinsic panoramic hyperspectral camera parameters were determined in a bundle block adjustment. The authors reported difficulties with the manual, and time-consuming, measurement of tie points between the point cloud and the spectral image, and they instead manually located homologous points between a textured photorealistic model and the hyperspectral images. Nieto et al. (2010) developed a framework aiming to obtain a detailed 3D geological map of the environment in an automated manner, using TLS, visible spectrum photography and very near infrared spectral imagery. Piecewise linear transformation was used to register the hyperspectral data, based on natural features found by the scale invariant feature transform (SIFT; Lowe 1999) interest operator between a 2D image formed from the coloured (RGB) point cloud and the true colour image composition created from the hyperspectral data.

In contrast to the work presented in Nieto et al. (2010), the hyperspectral imagery used in the current study does not cover the visible light spectrum, and therefore automated measures must handle the significantly different grayscale values between the conventional photographs and the short wave infrared (SWIR) imagery (Fig. 1).

The SIFT operator (Lowe 1999) is one of the most frequently used interest operators, and is employed in computer vision for many different purposes, such as 3D matching, 3D scene reconstruction, panorama stitching and motion tracking. The advantage of SIFT over other interest operators traditionally used in photogrammetry, such as the Förstner operator (Förstner 1986) or Harris operator (Harris & Stephens 1988), is the capability to deliver reliable results under difficult geometric and radiometric conditions (Jazayeri & Fraser 2010, Ostiak 2006). These properties were important in this study, where homologous (corresponding) points needed to be found between images acquired in varying parts of the electromagnetic spectrum (Fig. 1). The behaviour and efficiency of the SIFT interest operator has been mostly documented when applied to images of man-made features with many edges, corners and relatively regular patterns (Picard et al. 2009). In contrast, geological outcrops have a completely different scene composition and character, with natural features, few hard edges and irregular texture (Fig. 1). Additionally, the two image types to be matched in this study are acquired using different projections (central perspective and cylindrical panoramic) and with different resolutions. Such environments can be challenging for interest operators, especially if the distribution of identified points, used as tie points in later stages of processing, is of high importance.

3 Data Characteristics

The datasets used in this study were acquired in two different locations, in Spain and Norway. Data collection in the Pozalagua quarry in Cantabria, northern Spain, had the geological aim to map the distribution of sedimentary products such as limestone, different types of dolomite and calcite (Kurz et al. 2012). The main objective of the data collection in the second location, the Gusdal quarry near Åheim in central-West Norway, was to map the degree of serpentinization in ultramafic rocks. Because the rock types in the two

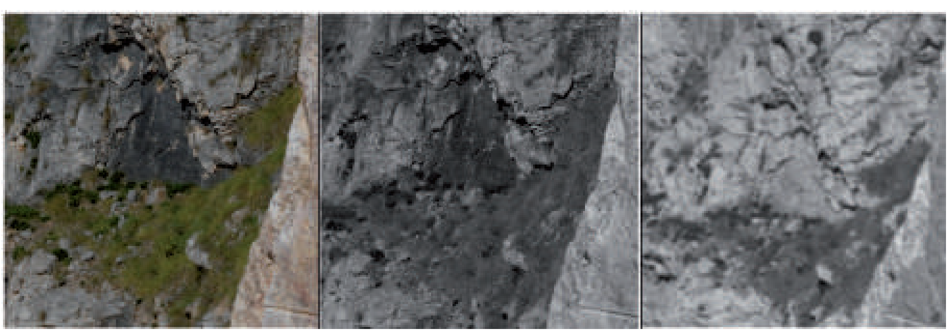


Fig. 1: Detail of a Nikon D200 photo; left: RGB; middle: greyscale; right: the corresponding part of a HySpex image, 1.629 μm wavelength.

quarries were different, their appearance in the images varies, creating different environments for tie point extraction.

Data were acquired with two separate systems, a portable hyperspectral HySpex SWIR-320m camera (Norsk Elektro Optikk AS), and a Riegl LMS-Z420i (Riegl 2012) terrestrial laser scanner with a Nikon D200 camera mounted on top. Both systems were located close to each other (very short baseline, c. 1.5 m) at the time of data acquisition in order to reduce differences in the viewing angle of the sensors. The exposure position of the HySpex sensor was included in an overview lidar scan so that the relative position of the two instruments was recorded.

3.1 *Hyperspectral Imagery*

The HySpex SWIR-320m is a portable terrestrial hyperspectral line scanner, with a 14° field of view across track. The sensor covers the spectral range from 1.3 µm to 2.5 µm over 241 bands, with a sampling interval of 5 nm and has 320 pixels in the across track direction. As a pushbroom linear-array sensor, the system uses a rotation stage to construct the image in the along-track direction. This results in cylindrical imaging geometry, which must be modelled to allow a precise coupling with the lidar and conventional image data. The sensor can be successfully represented by a geometric model for panoramic cameras (Schneider & Maas 2006), as described in (Kurz et al. 2011).

From the Pozalagua quarry two hyperspectral scans were used, of size 320 pixels x 1200 pixels (image B1) and 320 pixels x 1500 pixels (B2), and with a ground pixel size varying between 0.06 m and 0.09 m at ranges between 60 m and 130 m. From the dataset acquired at Gusdal quarry, one hyperspectral scan was used (A1), of size 320 pixels x 1000 pixels. With a scanning range of c. 20 m, this resulted in a ground sample distance of c. 0.015 m. Imagery was collected to maximize coverage of the geology in the limited field time, rather than for the photogrammetric configuration (Kurz et al. 2011).

3.2 *Digital Photos*

A 10.2 megapixel Nikon D200 single-lens reflex (SLR) camera was mounted on top of the terrestrial laser scanner. Two different lenses were used while acquiring data for this study: a Nikkor 85 mm lens (Pozalagua quarry), and a Nikkor 50 mm lens (Gusdal quarry). The interior orientation of the lenses was calibrated and the focal lengths fixed for the duration of the data collection. The position of the camera was calibrated relative to the scanner centre, so that the photographs could be easily registered in the lidar coordinate system and further used for texturing of the lidar surface model (Buckley et al. 2010).

The difference in optical construction and much higher sensor resolution of the Nikon camera resulted in a significantly better ground sampling distance when compared to the hyperspectral images, being between 0.004 m and 0.009 m for the Spain dataset and around 0.003 m in the Gusdal quarry. Horizontal and vertical fields of view of the Nikon D200 camera are 10.6° and 15.8° for the 85 mm lens and 18.0° and 26.6° for the 50 mm lens. The photos were collected with a 10 % horizontal overlap, so that several photos are often necessary to cover the area of one HySpex image or lidar scan (see example in Fig. 4).

3.3 *Lidar Point Cloud*

The Riegl LMS-Z420i is a long range, high performance terrestrial scanner capable of delivering points with accuracy (1σ) of 0.01 m and average repeatability of 0.004 m at the range of 50 m (Riegl 2012). The scan distance at the Pozalagua quarry varied between 60 m and 130 m, and the resulting point density was 0.05 m to 0.10 m. At the Gusdal quarry the scanner was located at a distance of 20 m from the outcrop wall, resulting in a point cloud density of around 0.03 m. The collected point clouds and digital imagery were processed as in Buckley et al. (2010) to obtain a photorealistic model for each site.

4 Data Integration Workflow

In order to integrate hyperspectral imagery into the lidar coordinate system, 3D control points need to be determined between the hyperspectral panoramic images and the data collected by the TLS system. Because the dynamic range of the point cloud intensity values recorded by the LMS-Z420i scanner was very low, it was not possible to create a synthetic intensity image with quality high enough for interest point extraction. Instead, the registered Nikon photos were used as the reference dataset, and hyperspectral panoramas were matched with these. The SIFT interest operator was used to find natural tie points between the hyperspectral imagery and the photos acquired by the Nikon camera (orange block in Fig. 2). Because of significant disparities between these two datasets, the interest operator algorithm was configured to yield as many points as possible, though a high number of false matches resulted. Therefore, robust homography model fitting using random sample consensus method (RANSAC) (Fischler & Bolles 1981) was applied in order to ensure the correct linking of conjugate points. 3D object coordinates of the matched points were obtained by subpixel interpolation within the point cloud backprojected on the image plane (blue block in Fig. 2). Finally, exterior orientation parameters of the hyperspectral image were established by spatial resection using the identified point pairs and the geometric model for panoramic cameras described in Schneider & Maas (2006) (green block in Fig. 2).

4.1 *Preprocessing of the Hyperspectral Imagery*

Firstly, raw spectral data acquired by the HySpex SWIR-320m sensor were converted into at-sensor radiances, according to the spectral calibration report provided by the sensor manufacturer. Each hyperspectral image consists of 241 very narrow spectral bands. Some of these bands contain only random noise, as certain wavelengths are completely absorbed by the atmosphere, and others are highly correlated with neighbouring bands. Many different methods exist to reduce the dimensionality of the data, and most of these are based on principal component analysis (PCA), relying on several preprocessing steps (de-striping, bad pixel interpolation, correction of brightness gradients). All these corrections are crucial for the final spectral mapping procedures but they also change locally the digital number values of the image. The influence of this change depends strongly also on the correction algorithm applied. In order to make the hyperspectral image registration independent of the spectral processing workflow, it was decided to make use of the almost unchanged at-sensor radiance images.

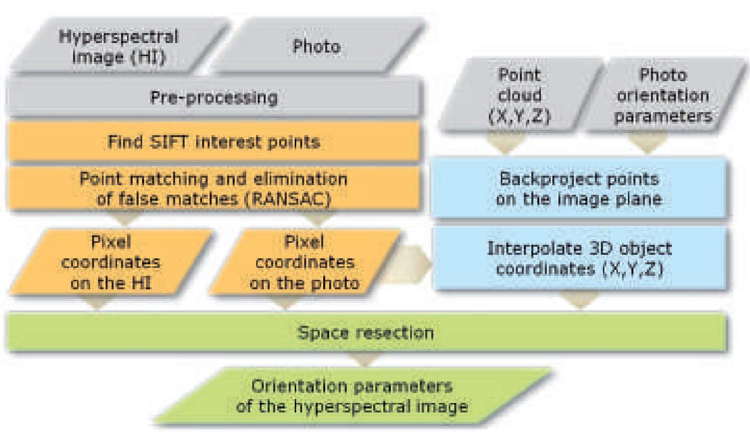


Fig. 2: General processing workflow.

4.2 *Preprocessing of the Photos*

Because the SIFT interest operator is applied to a single image band, the conventional photos (RGB) needed to be transformed into a single-band greyscale image. This conversion was performed according to the national television systems committee (NTSC) standard conversion (Durda 2012) used for calculating the effective luminance of a pixel.

Even though the interest point operator used in this study is robust to small projection differences (Lingua et al. 2009, Lourenco et al. 2012), in order to guarantee optimized conditions for the automatic tie point selection procedure, photos were reprojected from the central perspective projection to the (hyperspectral) cylindrical panoramic projection taking into account interior orientation parameters and lens distortion. Furthermore, in order to optimize images for the matching purposes and shorten the processing time, the Nikon photos were downsampled by a scale factor approximately equal to the ratio of the differences in pixel dimensions between the two data types. Special attention must be paid when selecting the downsampling algorithm, as the entire resampling process must be traceable so that the photo interior orientation parameters from the downsampled image can be restored. The result of this stage is that both hyperspectral and standard photos are in the same geometric projection, and have near-identical ground sample distances.

4.3 *SIFT Interest Points and Image Matching*

Hyperspectral images consist of hundreds of spectral bands, but, as previously mentioned, the SIFT detector operates only on a single image band. An attempt to merge all the points correctly matched between all the hyperspectral image bands and the Nikon photos revealed a small band mis-registration in the HySpex SWIR-320m sensor, most likely caused by a keystone effect (Neville et al. 2004). It was therefore decided to identify one spectral band that was optimal for matching with the overlapping Nikon photos. The criterion used in this selection procedure was the maximum total number of points matched between the spectral band and all the overlapping Nikon images. In the cases where several spectral bands had an equal total number of control points found, priority was given to the band with the highest minimum number of points matched for a single photo.

Using the described selection procedure, bands 61 (1.629 µm) and 78 (1.712 µm) were selected respectively as the best wavelengths for tie point extraction purposes for the Pozalagua B1 and B2 hyperspectral images (Fig. 3). In the dataset from the Gusdal quarry the 2.326 µm wavelength (band 204) was chosen. The troughs in Fig. 3 where very few points were matched correspond to spectral bands strongly affected by atmospheric absorption.

The VLFeat open source library (Vedaldi & Fulkerson 2008), containing implementations of the SIFT detector and descriptor, together with a matching procedure, was used for the purposes of this study. In order to yield

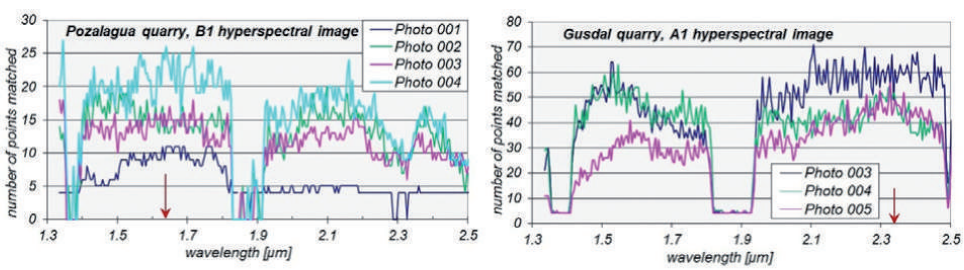


Fig. 3: Number of points matched between bands of the hyperspectral image and the overlapping photos. The wavelengths selected for matching are marked by red arrows.

as many interest points as possible, the SIFT detector was configured with values of 0 for the peak threshold, so that even points with the lowest contrast were detected, and 10 for the edge response threshold to filter weak edge responses. To reject matches that were too ambiguous, the VLFeat library uses a nearest neighbour algorithm with a distance ratio threshold, as suggested in Lowe (1999).

Due to many incorrect matches identified during a visual quality check, a homography model was fitted to the matched points using RANSAC. A homography model was used (instead of a fundamental matrix), due to a very short baseline between matched images, and the fact that after pre-processing (section 4.2) the Nikon images were very close to a rotated, scaled, or warped version of the hyperspectral imagery. It is worth stressing that this model was only used for false match elimination and not utilized in any of the further orientation steps.

The number of homologous points found at each stage of the processing workflow is summarized in Tab. 1. Only overlapping image areas were taken into account in detec-

Tab. 1: Number of homologous points at different processing stages (SIFT = scale invariant feature transform).

HySpex image	SIFT points	Photo (pre-processed)	detected SIFT points		Matched points				RANSAC inliers			
					per photo	total	% per photo	% total	per photo	total	% per photo	% of total matches
B1	1490	Photo 001	375	1528	62	274	16.5	18	15	110	24.2	40
		Photo 002	337		68		20.2		24		35.3	
		Photo 003	370		71		19.2		31		43.7	
		Photo 004	446		73		16.4		40		54.8	
B2	1662	Photo 005	467	1727	64	253	13.7	15	7	66	10.9	26
		Photo 006	459		57		12.4		14		24.6	
		Photo 007	381		70		18.4		26		37.1	
		Photo 008	420		62		14.8		19		30.6	
A1	1581	Photo 003	441	1231	105	303	23.8	24	61	165	24.2	54
		Photo 004	396		104		26.3		53		35.3	
		Photo 005	394		94		23.9		51		43.7	

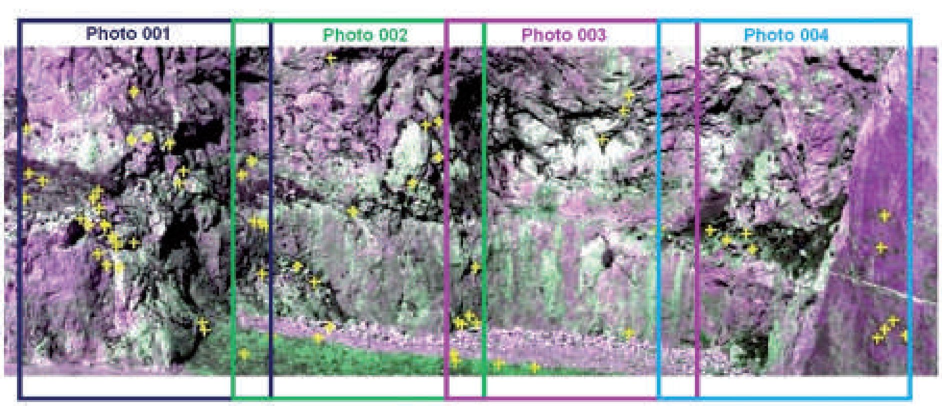


Fig. 4: Correctly matched points between B1 HySpex image (false colour composite shown here) and the overlapping Nikon photos.

tion of the SIFT interest points presented in Tab. 1. Due to the fact that the Nikon photos were acquired with 10 % side overlap (see Fig. 4) the total number of points found on the hyperspectral image and on the conventional photos are not directly comparable. The rate of points preliminarily matched in comparison to the number of SIFT keypoints detected is at the level of 12 %–26 % (column *Matched points* in Tab. 1). Taking into account the dissimilar spectral appearances of the two types of matched images, these values are satisfying and confirm the robustness of the scale invariant feature transform to image radiometric differences. The number of homologous points indicated by RANSAC as inliers was used to assess the matching accuracy (last column in Tab. 1), that is, a percentage of the correct matches in comparison to the total number of points matched. The estimated matching accuracy (26 %−54 %) is much lower than results reported in the case of matching images acquired at identical wavelengths (50 %–85 %, Lowe 2004). Nevertheless, the correct homologous point pairs could be reliably identified with the support of RANSAC and used in the later stages of the processing workflow.

4.4 *Point Cloud Back-Projection*

By themselves, the image coordinates of the homologous points in the hyperspectral imagery and the photos are not sufficient for the registration process. 3D object coordinates (X, Y, Z) are required, which may be derived indirectly from the lidar point cloud. In order to achieve that task, two processing steps were carried out for all conventional photos. Firstly, the point cloud was back-projected onto each photo plane using the collinearity condition, so that the relationship between each photo pixel and the object coordinates (X, Y, Z) was determined. This stage used the known projection centre, and the interior and exterior orientations from the pre-registered Nikon camera data. The effect of the photo downsampling algorithm was taken into account when defining the back-projection, in order to keep the correspondence between pixel and object point coordinates. In the second step, the exact object coordinates (sub-pixel) of each matched point were interpolated in the back-projected point cloud image pixel array.

4.5 *Spatial Resection*

The orientation parameters of each HySpex image were established by a spatial resection (Luhmann et al. 2007) using the geometric model for rotational linear-array panoramic cameras. The image and object coordinates of the control points, established automatically in previous processing steps, and the interior camera calibration parameters were supplied as input to the software Bundle (Schneider & Maas 2007).

5 Results and Discussion

For the Pozalagua dataset, 110 control points were successfully matched using the SIFT interest operator between the 4 overlapping conventional photos and the HySpex B1 image (Fig. 4). 66 points were matched for the B2 image. The matching accuracy rates were estimated as 40 % for image B1 and 26 % for image B2. For the Gusdal A1 image, 165 homologous points were correctly matched with the three overlapping photos. In this dataset the matching accuracy was 54 %.

The values expressing matching accuracies (Tab. 1.) are not directly comparable between datasets acquired in the two different locations. Not only were the image scales different (different imaging range, camera lens), but also the imaged geology was of a different mineral composition. The number of correctly found homologous points depends on the content/texture of the matched images, and thus on the both the 3D roughness and the 2D heterogeneity of the imaged surface. However, in the case of matching images acquired at different spectral ranges, the resulting number of points matched seems to be also material dependent. The difference in visual appearance between the materials in image A1 (Gusdal) and the conventional photos is much smaller than in the case of images B1 and B2 (Pozalagua). This can explain the relatively high (54 %) matching accuracy result for the A1 image. On the other hand, images B1 and

B2 contain very similar rock types, as they show the same outcrop wall, though a relatively large area of image B1 contains a freshly cut rock face that has a very distinctive texture pattern, and hence a higher matching accuracy (40 %). Other factors, such as the amount of surfacing weathering and the infrared lighting conditions at the time of exposure, and therefore the signal strength, may also influence the number of extracted interest points.

The identified homologous points were successfully used in a spatial resection (RMS of planimetric coordinates of ground points below 0.8 pix) to retrieve orientation parameters of the HySpex SWIR-320m scenes in the lidar project coordinate system. This in turn enabled the possibility to texture the lidar mesh model (Buckley et al. 2010) with the HySpex processing products. An example of an outcrop model textured with both conventional photographs and results of hyperspectral classification (hyperspectral map) is shown in Fig. 5.

The hyperspectral image orientation parameters established using the processing workflow presented above have been used in

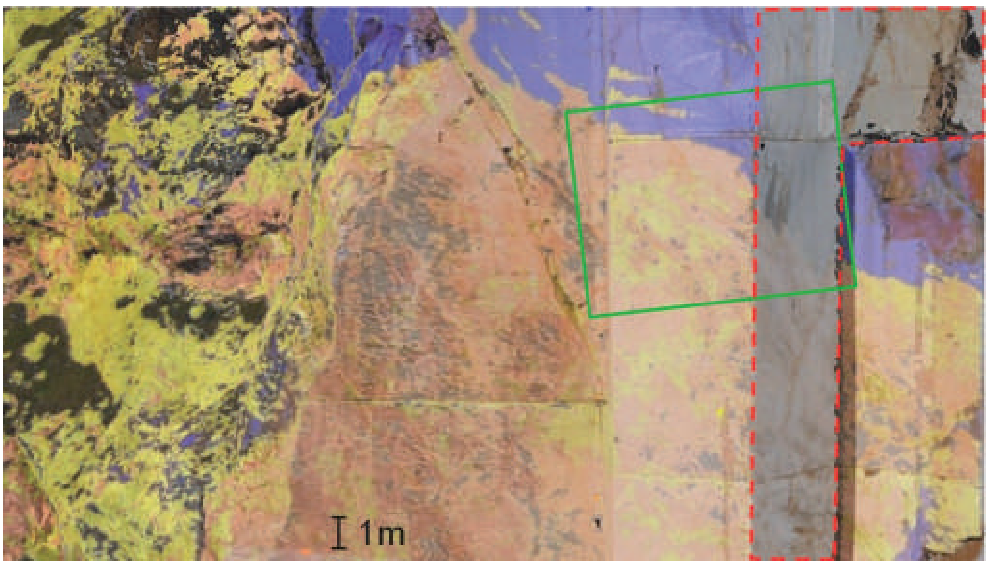


Fig. 5: 3D model textured with conventional photos (area outlined in red) and the results of hyperspectral image classification, where blue colours represent limestone and pink dolomite. Green rectangle marks the model sample shown in Fig. 6, right.

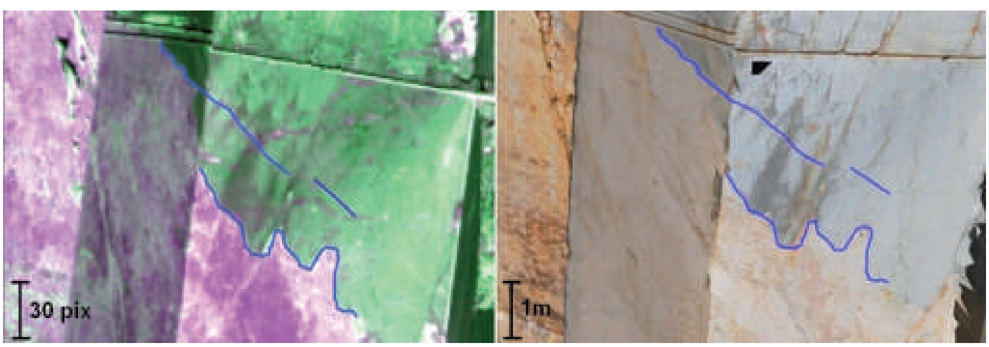


Fig. 6: 2D vectors digitized manually on a false colour RGB composite of the B1 HySpex image (left), and the corresponding 3D vectors displayed on the photorealistic model (right). Larger model overview shown in Fig. 5.

a visual quality check. Several clearly distinguishable features were digitized manually in the hyperspectral B1 image (Fig. 6, left). The collinearity condition with additional parameters (Schneider & Maas 2006) was applied to the 2D pixel coordinates of each point of the digitized vector line on the hyperspectral image, using the newly established image orientation parameters. For each point, a ray was intersected with the lidar mesh to obtain a 3D point in the project coordinate space. The resulting 3D vectors were displayed together with the photorealistic outcrop model (Fig. 6, right). The correct position of the 3D vectors on the photorealistic model shows that HySpex image orientation parameters obtained using the proposed method can be used in the data integration process.

6 Conclusions and Future Work

In this paper a semi-automated method for integrating panoramic hyperspectral imagery into photorealistic outcrop models was presented. The automation level was increased by automatic point matching between the hyperspectral images and conventional photos. The SIFT interest operator proved to be robust to very different "radiometric conditions", and capable of finding homologous points on images acquired in significantly different spectral ranges, i.e. in visible light (Nikon photos) and the short wave infrared spectrum (HySpex SWIR-320m imagery). The resulting matching accuracy rates are low when compared to the results obtained between data collected in similar spectral ranges; however, the elimination of false matches using RANSAC yields enough correct points to register the hyperspectral panoramic images to the lidar project coordinate system.

Although the control point distribution was not optimized for photogrammetric processing, the resulting HySpex image orientation enables the possibility to complement lidar outcrop models with spectral products, and is therefore a valuable technique for the domain end-users, geologists.

More work is required in order to evaluate the geometric accuracy of the determined orientation parameters. The quality of the developed data registration procedure can be further improved by using a bundle block adjustment to simultaneously orientate multiple images, instead of the single image spatial resection applied here. Application of a bundle block adjustment would be beneficial to check the stability of the interior orientation parameters of the HySpex SWIR-320m camera. In this case, matching should be restricted to areas of high data overlap, to reduce processing time. The control point distribution should be controlled and optimized for the purposes of photogrammetric applications. Influence of contrast enhancement on the final number of correct control points found will be investigated in order to force successful matches in low texture image areas.

Acknowledgements

This research is supported by the Research Council of Norway and FORCE (project number 193059). Statoil ASA and Sibelco Nordic are acknowledged for supporting the data collection. Norsk Elektro Optikk AS and Riegl GmbH are thanked for providing hardware and software support.

References

Bellian, J.A., Kerans, C. & Jennette, D.C., 2005: Digital Outcrop Models: Applications of Terrestrial Scanning Lidar Technology in Stratigraphic Modeling. – Journal of Sedimentary Research **75** (2): 166–176.

Böhm, J. & Becker, S., 2007: Automatic Marker-Free Registration of Terrestrial Laser Scans Using Reflectance Features. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XXXVI** (5/W17): 338–344.

Buckley, S.J., Enge, H., Carlsson, C. & Howell, J.A., 2010: Terrestrial Laser Scanning for Use in Virtual Outcrop Geology. – The Photogrammetric Record **25** (131): 225–239.

Buckley, S.J., Howell, J.A., Enge, H.D. & Kurz, T.H., 2008: Terrestrial Laser Scanning in Geology: Data Acquisition, Processing and Accuracy Considerations. – Journal of the Geological Society **165** (3): 625–638.

Clark, R.N. & Roush, T.L., 1984: Reflectance Spectroscopy: Quantitative Analysis Techniques for Remote Sensing Applications. – Journal of Geophysical Research **89** (B7): 6329–6340.

Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G.A. & Vergo, N., 1990: High Spectral Resolution Reflectance Spectroscopy of Minerals. – Journal of Geophysical Research **95** (B8): 12653–12680.

Durda, F., 2012: NTSC Primer. http://nemesis.lonestar.org/reference/internet/web/color/ntsc_primer.html (February 2012).

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

Forkuo, E.K. & King, B., 2004: Automatic Fusion of Photogrammetric Imagery and Laser Scanner Point Clouds. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XXXV** (B4): 921–926.

Förstner, W., 1986: A Feature Based Correspondence Algorithm for Image Matching. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XXVI** (3/3): 150–166.

González-Aguilera, D., Rodríguez-Gonzálvez, P. & Gómez-Lahoz, J., 2009: An Automatic Procedure for Co-Registration of Terrestrial Laser Scanners and Digital Cameras. – ISPRS Journal of Photogrammetry and Remote Sensing **64:** 308–316.

Harris, C. & Stephens, M., 1988: A Combined Corner and Edge Detector. – **4th** Alvey Vision Conference: 147–151, Alvey, UK.

Jazayeri, I. & Fraser, C.S., 2010: Interest Operators for Feature-Based Matching in Close Range Photogrammetry. – The Photogrammetric Record **25** (129): 24–41.

Kurz, T.H., Buckley, S.J., Howell, J.A. & Schneider, D., 2011: Integration of Panoramic Hyperspectral Imaging with Terrestrial Lidar Data. – The Photogrammetric Record **26** (134): 212–228.

Kurz, T.H., Dewit, J., Buckley, S.J., Thurmond, J.B., Hunt, D.W. & Swennen, R., 2012: Hyperspectral Image Analysis of Different Carbonate Lithologies (Limestone, Karst and Hydrothermal Dolomites): the Pozalagua Quarry Case Study (Cantabria, North-west Spain). – Sedimentology **59** (2): 623–645.

Lingua, A., Marenchino, D. & Nex, F., 2009: Performance Analysis of the SIFT Operator for Automatic Feature Extraction and Matching in Photogrammetric Applications. – 10th CESCG Conference **9** (5): 3745–3766.

Lourenço, M., Barreto, J.P. & Malti, A., 2010: Feature Detection and Matching in Images with Radial Distortion. – ICRA'**2010:** 1028–1034.

Lowe, D., 1999: Object Recognition from Local Scale-Invariant Features. – 7th IEEE International Conference on Computer Vision **2:** 1150–1157.

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

Luhmann, T., Robson, S., Kyle, S. & Harley, I., 2007: Close range photogrammetry: principles, techniques and applications. – 528 pp, Wiley, UK.

Neville, R.A., Lixin, S. & Staenz, K., 2004: Detection of Keystone in Imaging Spectrometer Data. – SPIE, April 2004: 208–217, Orlando, FL, USA.

Nieto, J.I., Monteiro, S.T. & Viejo, D., 2010: 3D Geological Modelling Using Laser and Hyperspectral Data. – IGARSS **2010** Symposium: 4568–4571.

Ostiak, P., 2006: Implementation of HDR Panorama Stitching Algorithm. – **10th** CESCG Conference **2006**: Castá-Papiernicka.

Picard, D., Cord, M. & Valle, E., 2009: Study of SIFT Descriptors for Image Matching Based Localization in Urban Street View Context. – CMRT2009 Conference, Paris.

Riegl LMS-Z420i, Brochure, http://www.riegl.com/uploads/tx_pxpriegldownloads/10_DataSheet_Z420i_03-05-2010.pdf (February 2012).

Schneider, D. & Maas, H.-G., 2006: A Geometric Model for Linear-Array-Based Terrestrial Panoramic Cameras. – The Photogrammetric Record **21** (115): 198–210.

Schneider, D. & Maas, H.-G., 2007: Integrated Bundle Adjustment with Variance Component Estimation – Fusion of Terrestrial Laser Scanner Data, Panoramic and Central Perspective Image Data. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XXXVI:** 373–378.

Van Der Meer, F.D., 2004: Analysis of Spectral Absorption Features in Hyperspectral Imagery. – International Journal of Applied Earth Observation and Geoinformation **5** (1): 55–68.

Vedaldi, A. & Fulkerson, B., 2008: {VLFeat}: An Open and Portable Library of Computer Vision Algorithms, http://www.vlfeat.org (February 2012).

Addresses of the Authors:

Aleksandra A. Sima, Simon J. Buckley & Tobias H. Kurz, Centre for Integrated Petroleum Research (Uni CIPR), Postboks 7810, 5020 Bergen, Norway, Tel.: +47-555-83-670, -458, -646, e-mail: {aleksandra.sima}{simon.buckley}{tobias.kurz}@uni.no

Danilo Schneider, Technische Universität Dresden, Germany, Tel.: +49-351-463-33144, e-mail: danilo.schneider@tu-dresden.de

Manuskript eingereicht: März 2012
Angenommen: Mai 2012

PFG 2012 / 4, 0455–0466
Stuttgart, August 2012

Article

Concepts for Cartography-Oriented Visualization of Virtual 3D City Models

Amir Semmo, Dieter Hildebrandt, Matthias Trapp & Jürgen Döllner, Potsdam

Keywords: 3D city models, cartography-oriented visualization, style description languages, real-time rendering

Summary: Virtual 3D city models serve as an effective medium with manifold applications in geoinformation systems and services. To date, most 3D city models are visualized using photorealistic graphics. But an effective communication of geoinformation significantly depends on how important information is designed and cognitively processed in the given application context. One possibility to visually emphasize important information is based on non-photorealistic rendering, which comprehends artistic depiction styles and is characterized by its expressiveness and communication aspects. However, a direct application of non-photorealistic rendering techniques primarily results in monotonic visualization that lacks cartographic design aspects. In this work, we present concepts for cartography-oriented visualization of virtual 3D city models. These are based on coupling non-photorealistic rendering techniques and semantics-based information for a user, context, and media-dependent representation of thematic information. This work highlights challenges for cartography-oriented visualization of 3D geovirtual environments, presents stylization techniques and discusses their applications and ideas for a standardized visualization. In particular, the presented concepts enable a real-time and dynamic visualization of thematic geoinformation.

Zusammenfassung: *Ansätze zur kartographischen Gestaltung von virtuellen 3D-Stadtmodellen*. Virtuelle 3D-Stadtmodelle repräsentieren ein wirkungsvolles und mannigfaltiges Medium in Geoinformationssystemen und -diensten. Häufig wird eine Vielzahl an 3D-Stadtmodellen photorealistisch dargestellt. Eine effektive Kommunikation von Geoinformation beruht allerdings darauf, wie wichtige Informationen innerhalb eines gegebenen Anwendungskontextes gestaltet und kognitiv verarbeitet werden können. Eine Möglichkeit wichtige Informationen visuell hervorzuheben, beruht auf nichtphotorealistischer Bildsynthese, die künstlerische Darstellungsstile umfasst und durch ihre Ausdrucksstärke und kommunikativen Aspekte gekennzeichnet ist. Deren direkte Anwendung resultiert jedoch in einer primär monotonen Gestaltung, bei der kartographische Gestaltungsaspekte nicht berücksichtigt werden. In diesem Beitrag stellen wir Ansätze zur kartographischen Gestaltung virtueller 3D-Stadtmodelle vor. Diese koppeln Techniken der nichtphotorealistischen Bildsynthese mit semantischen Informationen für eine nutzer-, kontext- und medienadäquate Repräsentation von thematischen Informationen. Wir identifizieren Herausforderungen einer kartographischen Gestaltung in 3D geovirtuellen Umgebungen, präsentieren Bildsynthese-Techniken zur kartographischen Gestaltung und diskutieren deren potentielle Anwendungen und Ansätze für einen standardisierten Visualisierungsprozess. Im Besonderen ermöglichen die vorgestellten Ansätze eine echtzeitfähige und dynamische Gestaltung von thematischen Geoinformationen.

1 Introduction

Virtual 3D city models and, more general, 3D geovirtual environments, play an important role for communication of and interaction with complex, 2D and 3D geoinformation. They represent key elements in a growing number of applications, systems, and services, for example, in mass-market systems (e.g. Google Earth or Ovi Maps 3D), expert tools

DOI: 10.1127/1432-8364/2012/0131 1432-8364/12/0131 $ 2.75

for urban planning and development, disaster management and environmental information systems. Although most virtual 3D city models abstract from reality to serve comprehension, photorealistic renderings commonly suffer from an inefficient communication of geoinformation. As soon as a compact, yet attractive presentation is required, perceptional, cognitive and graphical design issues need to be considered.

Non-photorealistic rendering (NPR) is known to enhance communication aspects (Gooch et al. 2011). It generally builds on two principles: highlighting of important or prioritized information, and abstraction of less important information. Nowadays, various rendering techniques and abstraction concepts exist (Gooch et al. 2011) to enable a non-photorealistic visualization of 2D and 3D *features* (i.e., abstractions of real-world phenomena, ISO 19101:2002). However, applying non-photorealistic rendering techniques to a single 3D city model results in a monotonic visualization. For an efficient communication, it requires having an adequate representation on feature level. One approach is to adopt design principles from cartography (Jobst & Döllner 2008) to efficiently communicate thematic information and geospatial relationships (Fig. 1), for instance by using semantics-based information (Hildebrandt 2011).

In this work, we present concepts for cartography-oriented visualization of virtual 3D city models that seamlessly integrate into a visualization pipeline. These are based on semantic information to enable the communication of thematic information by principles of cartography. For this, we make several contributions. First, we identify challenges and principles for the cartography-oriented visualization of virtual 3D city models. We then propose ideas for a standardized visualization that is optimized for cartographic design. Finally, we outline application examples of our concept to show how cartography-oriented visualization can be used.

The remainder of this work is organized as follows. We review related work on non-photorealistic rendering for virtual 3D city models (section 2) before we summarize our concept of a semantics-based and cartography-oriented visualization including challenges and standardization by means of a visualization pipeline (section 3). We then propose stylization techniques on top of this concept (section 4) and outline their applications and use cases (section 5). Finally, we state ideas for possible future extensions and conclude this work (section 6).

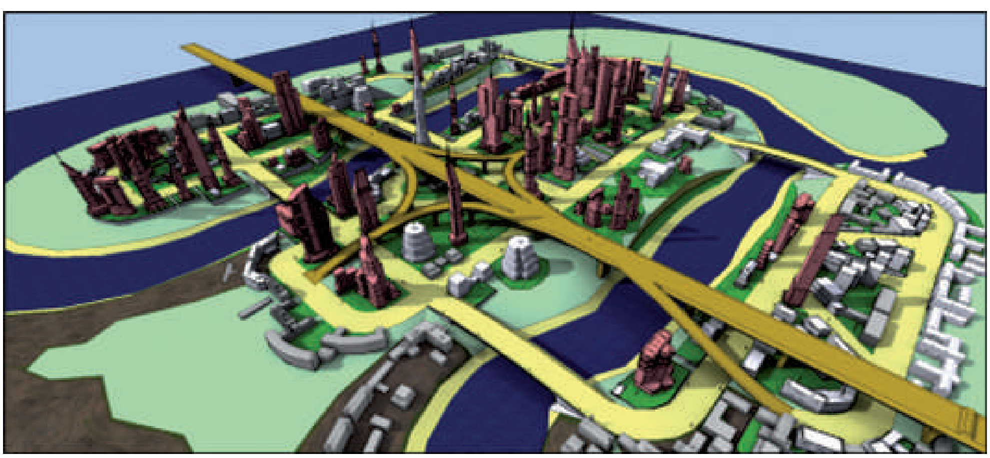


Fig. 1: Exemplary visualization of a virtual 3D city model based on feature semantics. Features are coloured and stylized according to their feature class (i.e., building, street, green space, terrain, and water surface).

2 Related Work

Prominent examples for a non-photorealistic rendering are computer-generated imitations of watercolours, pen-and-ink and pencil drawings (Gooch et al. 2011). Previous work showed that NPR is able to improve visual clarity in 3D geovirtual environments (Döllner & Walther 2003, Döllner et al. 2005, Fig. 2B/C). Our work uses an edge enhancement in image-space (Nienhaus & Döllner 2003) and object-space (Döllner & Walther 2003) to improve the perception of city structures, and techniques for the stylization of 2D textures (Kyprianidis & Döllner 2008) to highlight important information and filter less important information.

A context-aware abstraction has the potential to improve the perception of structures in images, as is shown exemplary for aerial images (Semmo et al. 2010). In previous work, we already introduced and discussed a semantics-based, selective stylization of virtual 3D city models, the concurrent use of photorealistic and non-photorealistic stylization within one single image, and a set of style operators suitable for virtual 3D city models (Hildebrandt 2011). Virtual 3D city models can be enriched with semantic information by using, e.g. ontology-based approaches (Métral et al. 2009). The OGC's encoding standard CityGML (Kolbe 2009) describes an approach for modeling and exchanging semantically enriched virtual 3D city models.

First theoretical work on cartography-oriented visualization is found in the context of web visualization services (Zipf 2005). A previous work described a first approach how to render virtual 3D city models with cartography-oriented design (Semmo et al. 2011). We extend this work by identifying challenges for computer-generated cartography-oriented design, use cases and a discussion for a standardized visualization.

The *Open Geospatial Consortium* (OGC) proposed the languages *Styled Layer Descriptor 3D* (SLD3D) and *Symbology Encoding 3D* (SE3D) (Neubauer & Zipf 2009) as standards for the stylization of 3D geoinformation. These languages represent simple adaptations of the SLD (Müller 2006) and SE (Lupp 2007), which are used for 2D geoinformation, but are still in an early stage of development. In this work, we highlight the importance of style description languages and discuss their applicability for advanced visualization techniques.

3 Principles of Cartography-Oriented Visualization

This section discusses principles of an effective cartography-oriented visualization: design principles from traditional cartography and interactive 3D environments for effective cartography-oriented illustrations, real-time rendering techniques, and concepts for a standardized visualization.

3.1 *Cartography-Oriented Illustrations*

In contrast to 2D maps (Fig. 2A) and paintings by famous cartographers like Merian

Fig. 2: Comparison between (A) digital 2D maps and virtual 3D city models rendered with (B) photorealistic rendering techniques, and (C) non-photorealistic rendering techniques.

(Fig. 3A), virtual 3D city models have special characteristics in terms of their data complexity and interactive presentation. We identified the following principles for effective computer-generated, cartography-oriented illustrations:

Automated and generic abstraction techniques. A stylization of virtual 3D city models is based on abstraction techniques that should work independently from model contents. Fig. 3 exemplifies a hand-drawn map in comparison to a computer-generated stylization.

Interaction with 3D geovirtual environments. A user should be able to adapt the stylization of a 3D scene to the performed task and according to prioritized information. One example is an adaptive configuration of colour schemes for prioritized feature classes.

Real-time rendering. Rendering techniques should perform in real-time to maintain immersion (MacEachren et al. 1999) within 3D geovirtual environments.

Frame-to-frame coherence. While navigating through a 3D geovirtual environment, the viewing direction can be continuously changed. Correspondingly, visualization techniques need to maintain a frame-to-frame coherence to avoid visualization artifacts by changes in viewing parameters, depth values and occlusion of models.

3.2 *Visualization Pipeline*

A cartography-oriented visualization of virtual 3D city models is primarily based on two concepts: geometric abstraction using generalization operators to transform aspects of reality into datasets or maps (e.g. for 3D city models, Glander & Döllner 2008, or street networks, Agrawala & Stolte 2001), and visual abstraction using non-photorealistic rendering techniques.

The generalization model of Grünreich proposes three passes of generalization towards the computation of a *cartographic city model*: object generalization, model generalization and cartographic generalization (Grünreich 1992). Common generalization operators comprise the enhancement, displacement, elimination, typification or amalgamation (Foerster et al. 2007), and apply at each visualization stage of the visualization pipeline (Ware 2004). Fig. 4 exemplifies a visualization pipeline that structures the visualization of virtual 3D city models into three stages.

In this work, cartographic stylization is based on level-of-abstraction (LoA) that defines the spatial and thematic granularity at which model contents is represented (Glander & Döllner 2008). In contrast to level-of-detail (LoD), which focuses on geometric abstraction, LoA directly influences visual abstraction, e.g. performed by shading effects at the rendering stage of the visualization pipeline (Fig. 4). To use visual abstraction in the context of a cartographic stylization, it is essential

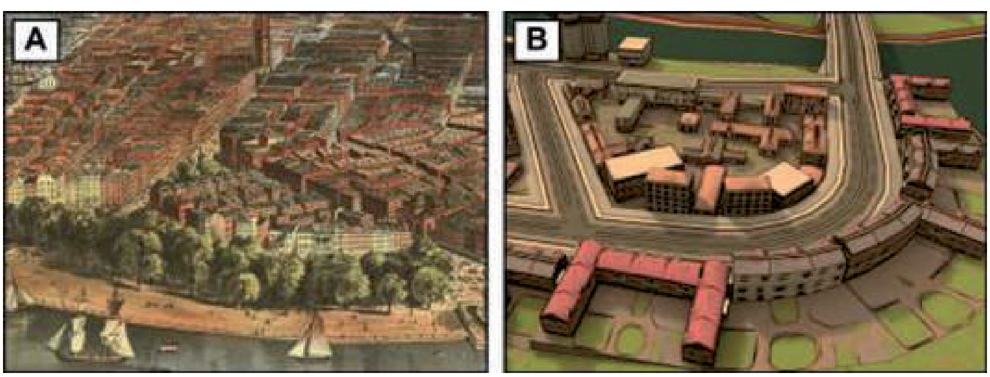


Fig. 3: Comparison between (A) a hand-drawn map (Library of Congress – New York) and (B) an automatically computer-generated, cartography-oriented illustration of a virtual 3D city model.

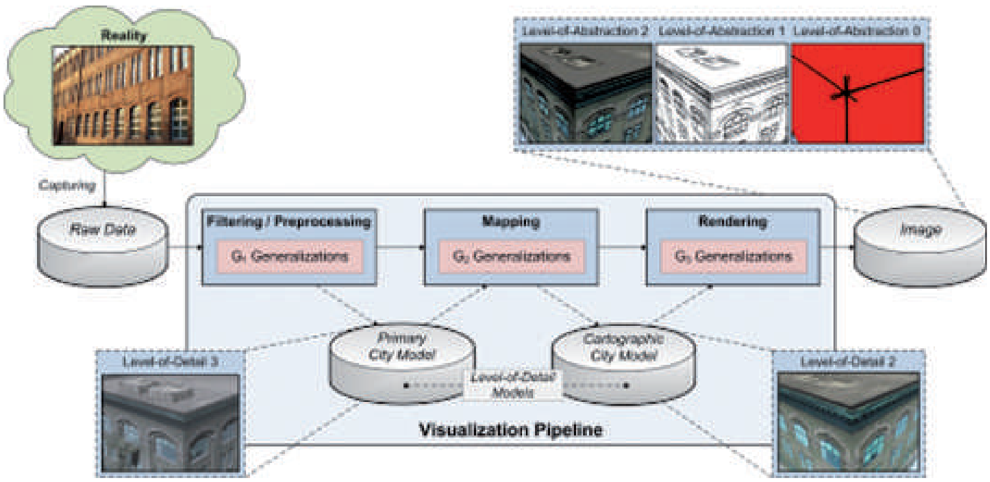


Fig. 4: Visualization pipeline (Ware 2004) that incorporates generalization operators for a cartography-oriented visualization by geometric and visual abstraction.

to map features to feature classes. The classification of objects represents a generalization at the pre-processing stage of the visualization pipeline to enable a semantics-based rendering. Nowadays, data formats exist for an exchange and storage of virtual 3D city models with thematic information, such as the OGC standard CityGML (Kolbe 2009). However, most virtual 3D city models do not incorporate explicit thematic information (e.g. reconstructed from laser scans), and are therefore not suitable for a cartography-oriented visualization. In our concept, the appearance of 3D objects is utilized for a semi-automatic classification. Here, we distinguish between two approaches: (1) features are grouped according to their appearance (e.g. defined by textures or materials) or (2) are manually grouped at run-time.

3.3 *Style Description Languages*

To enable an effective and efficient communication of geoinformation, in general, the used stylization needs to be adapted to the specific user, the user's task, the application, and the model. In principle, stylization can be described by a formal style description language. This language basically defines the mapping of geodata to geometric primitives and visual attributes. For a generic, computer-based processing, it is required that this language is defined with well-defined syntax and semantics. If such a style description language exists, we see several advantages. First, style specifications can be viewed, modified, compared, exchanged and thus reused among instances of the same system without the need to modify the systems processing of the specification. Second, if the used language complies to a standard, specifications are independent from specific systems and implementations, and can be exchanged and thus reused between different, standard conforming systems. Third, in general, standards represent the formal specification of approved concepts of a domain by domain experts. They are important to disseminate best practices and technological progress (Hildebrandt & Döllner 2010).

For style description languages designed for 3D geovirtual environments, we identified various requirements, in addition to those already presented in section 3.1. A style description language should

- use or extend existing standards,
- be declarative to increase the level-of-abstraction,
- consider proven concepts of the visualization pipeline (Ware 2004),
- consider proven concepts for modeling geodata (e.g. feature, layer, and filter) and maintain a semantics-based and selective stylization (Hildebrandt 2011),

- be able to parametrize stylizations by well-defined operators defined on the abstraction level of the domain (and not, e.g. the abstraction level of technology) to compose high-level stylizations,
- be able to provide user-defined style operators,
- support the usage of 3D-specific and advanced style operators, and
- specifically support style operators working in image-space at a post-processing stage that are designed to improve the rendering performance and expressiveness, and ease flexibility and parametrization (Akenine-Möller & Haines 2008).

In the domain of virtual 3D city models and 3D geovirtual environments, the only relevant standards-based style description languages are SLD3D and SE3D presently proposed as drafts by the OGC. These drafts already support many of the identified requirements. However, several important requirements are not supported. Due to the lack of user-defined style operators, users are constrained to use only the fixed set of operators defined in the standards and their combinations. 3D-specific and advanced style operators are not supported. For instance, it is not possible to specify global illumination, advanced material properties or environmental effects. Further, since the drafts have no support for specific style operators working in image-space, the significant potential of this type of operators cannot be exploited. Examples of operators working efficiently in image space are global illumination (e.g. by *deferred shading*, *shadow mapping*, or *ambient occlusion*) (Akenine-Möller & Haines 2008) and the abstraction of the image content by removing unimportant features while at the same time highlighting important features (Hildebrandt 2011). Standardized style description languages exist in other domains. However, these typically do not meet the identified requirements for the domain of interest in this work (e.g. specifying styling in OpenGL from the domain of computer graphics is not declarative, lacks direct support of domain concepts, and does not provide composable operators), or their application to this domain was not yet demonstrated.

To summarize, standardized style description languages have a high potential. However, current drafts for such languages in the domain of interest in this work cannot express advanced stylizations that go beyond the straightforward mapping of features to simple geometries and their colourization. In section 4 we propose stylization techniques used for cartography-oriented visualization, and in section 5 exemplify that these stylizations cannot be expressed with the current drafts proposed by the OGC.

4 Techniques for Cartographic Stylization and Implementation

This section outlines visualization techniques that serve a cartography-oriented visualization (Fig. 5) by considering the challenges and principles of section 3.1.

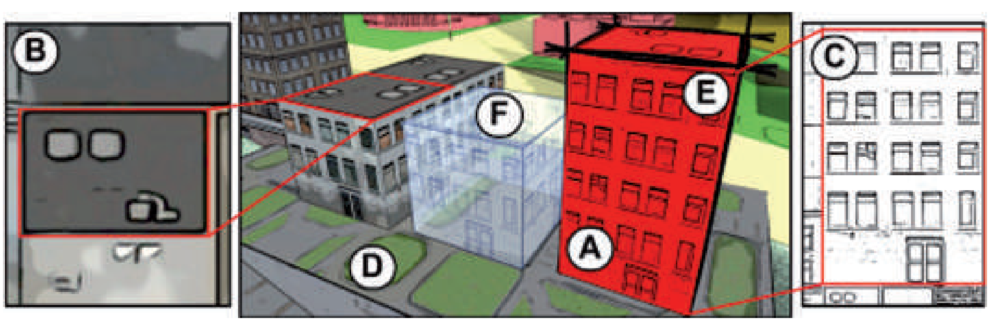


Fig. 5: Overview of stylization techniques: (A) colourization, (B/C) stylization of textures, (D/E) edge enhancement in image-space and object-space, and (F) transparency effects.

4.1 *Colourization*

The colourization of geometric primitives adopts design principles known from cartography for visual abstraction and efficient communication of thematic geoinformation. In combination with an edge enhancement, colour schemes can improve the perception of city structures. BREWER proposes conventions for using colours in cartography (BREWER 1994). For example, qualitative colour schemes can be used to communicate feature classes efficiently and ease comprehension of a city's structure. By contrast, sequential colour schemes can be used to communicate thematic information for analysis purposes, for instance to illustrate the distribution of green spaces in a virtual city model (Fig. 6D).

4.2 *Stylization of Object Textures*

A *bilateral filter* and a *difference-of-gaussians-filter* (DoG) (KYPRIANIDIS & DÖLLNER 2008) can be used for an automated abstraction of textures. Because these techniques operate in colour space and are content-independent, they can be applied to aerial images, terrestrial photography or synthesized textures. The stylization of textures is performed resolution-dependently to enable a view-dependent level-of-abstraction (SEMMO et al. 2010). In contrast to the work by KYPRIANIDIS & DÖLLNER (2008), the output of the edge enhancement is not combined with the quantized colour output (Fig. 5B/C). Instead, colour and outline can be blended at rendering time to stylize features of a specific class differently, e.g. to highlight important or prioritized information.

4.3 *Edge Enhancement of 3D Objects*

An edge enhancement of 3D objects highlights structural aspects of virtual 3D city models by emphasizing and separating features located in the background of an image. The edge enhancement in image-space is based on detecting discontinuities in depth, surface orientation and at boundaries of 3D objects (NIENHAUS & DÖLLNER 2003). Although the technique is feasible to stylize generic 3D scenes independently from their geometric complexity, it lacks the possibility to customize the design of enhanced edges (e.g. width, offset or stroke style). To this end, the edge enhancement in object-space provides an alternative approach by using textured, screen-aligned billboards (DÖLLNER & WALTHER 2003). A texture-based edge enhancement is a well-known method to emphasize 3D objects and communicate uncertainty, e.g. to aid city planning and reconstruction (SEMMO et al. 2011). Technically, textured strokes are used and parametrized at run-time for a flexible stylization (DÖLLNER & WALTHER 2003).

4.4 *Transparency of 3D Objects*

Transparency effects are a well-known method to improve comprehension and visibility of occluded model entities in 3D-space. Using transparency in the context of virtual 3D city models is not a new approach. For example, alpha blending can be used to enable blue print illustrations of 3D scenes (NIENHAUS & DÖLLNER 2004), e.g. to highlight construction states of buildings models (Fig. 6C). Further, transparency effects are used in web services and systems, such as Google Maps, to aid the perception of complex structures or the architecture of 3D building models.

4.5 *Real-Time Rendering*

The goal for an effective, interactive visualization is to render complex virtual 3D city models at real-time and to offer a flexible, fully configurable stylization. In addition to a manual classification of features, this includes a parametrization of styles at run-time.

For a semantics-based, cartography-oriented visualization, each feature needs to be mapped to its corresponding feature class. To this end, each feature is assigned – in addition to its semi-automatic classification – a unique identifier. In case a re-classification occurs at run-time, the assignments of feature identifiers to feature classes are updated, including style parametrizations.

For rendering complex 3D scenes, the graphics API OpenGL 3.2 and the OpenGL Shading Language (GLSL) are used. To parametrize the appearance of each single 3D feature dynamically, feature identifiers are assigned to fragments. This enables a mapping of fragments to feature classes and their defined stylization at real-time. For rendering transparency effects, a "*stencil routed A-buffer*" (Myers & Bavoil 2007) is used to buffer occluded model fragments. The fragments are then sorted dynamically in a post-process to synthesize the final image by image blending.

5 Application Examples

With a real-time, cartography-oriented visualization of virtual 3D city models, new possibilities emerge for the communication of thematic geoinformation. For example, regions-of-interest and structures of a city's image can be highlighted, such as landmarks, routes and districts (Lynch 1960), e.g. to ease orientation and navigation tasks. Further, the cartography-oriented visualization can aid a fast decision support in navigation systems by highlighting street networks and landmarks. Furthermore, thematic information can be visualized for analysis purposes by using qualitative and quantitative colour schemes (Brewer 1994), e.g. for visualization and exploration of crime maps (Wolff & Asche 2009). Finally, parametrized abstractions can be used to visually enhance city planning systems with focus on architectural aspects.

The visual quality of a stylization is a well known issue in aesthetics of computer graphics. It primarily builds on style configurations that adapt to a viewer's needs, technical background and cognitive qualities. Because the presented concepts and techniques enable a broad range of configuration possibilities, not all parametrizations are beneficial. Fig. 6 exemplifies configurations for three different scenarios:

1. a highlighting of feature classes that is based on a diverging colour scheme and an edge enhancement in image- and object-space to improve the perception of a city's structure (Fig. 6B),
2. a cartography-oriented visualization that communicates construction dates and states of building models by using a diverging colour scheme and transparency effects (Fig. 6C),
3. a visualization that communicates the distribution of green spaces, e.g. to analyze life quality within an urban environment (Fig. 6D).

The presented application examples cannot be completely expressed by current style description languages such as SLD3D and SE3D of the OGC. Whereas colourization and transparency effects can be defined by material properties, no language features exist to define non-photorealistic abstraction, edge enhancement, and global illumination. Moreover, the SLD3D and SE3D are located conceptually in the mapping stage of the visualization pipeline (section 3.2). Although this concept facilitates real-time rendering, an ad-hoc modification of the stylization is not possible for complex models because of their large data volume. In summary, current proposals for standardized style description languages cannot express advanced stylizations. This concludes that we cannot realize the benefits identified in section 3.3, and that the stylizations presented in this work remain a proprietary component of individual systems and implementations.

6 Conclusions and Future Work

The semantics-based, cartography-oriented visualization of virtual 3D city models represents a promising approach for improving the efficiency of 3D geoinformation communication. The identified approaches for a standardized visualization of cartography-oriented illustrations exemplify how high-quality geospatial visualization services can be technically implemented. For instance, mobile mapping services can benefit from a cartography-oriented visualization to improve the perception of information on small screens (Jobst & Döllner 2008). Further, the application examples show how cartography-oriented visualization can be used in GIS, urban planning and tourist information systems. But in order to demonstrate their effectiveness within these

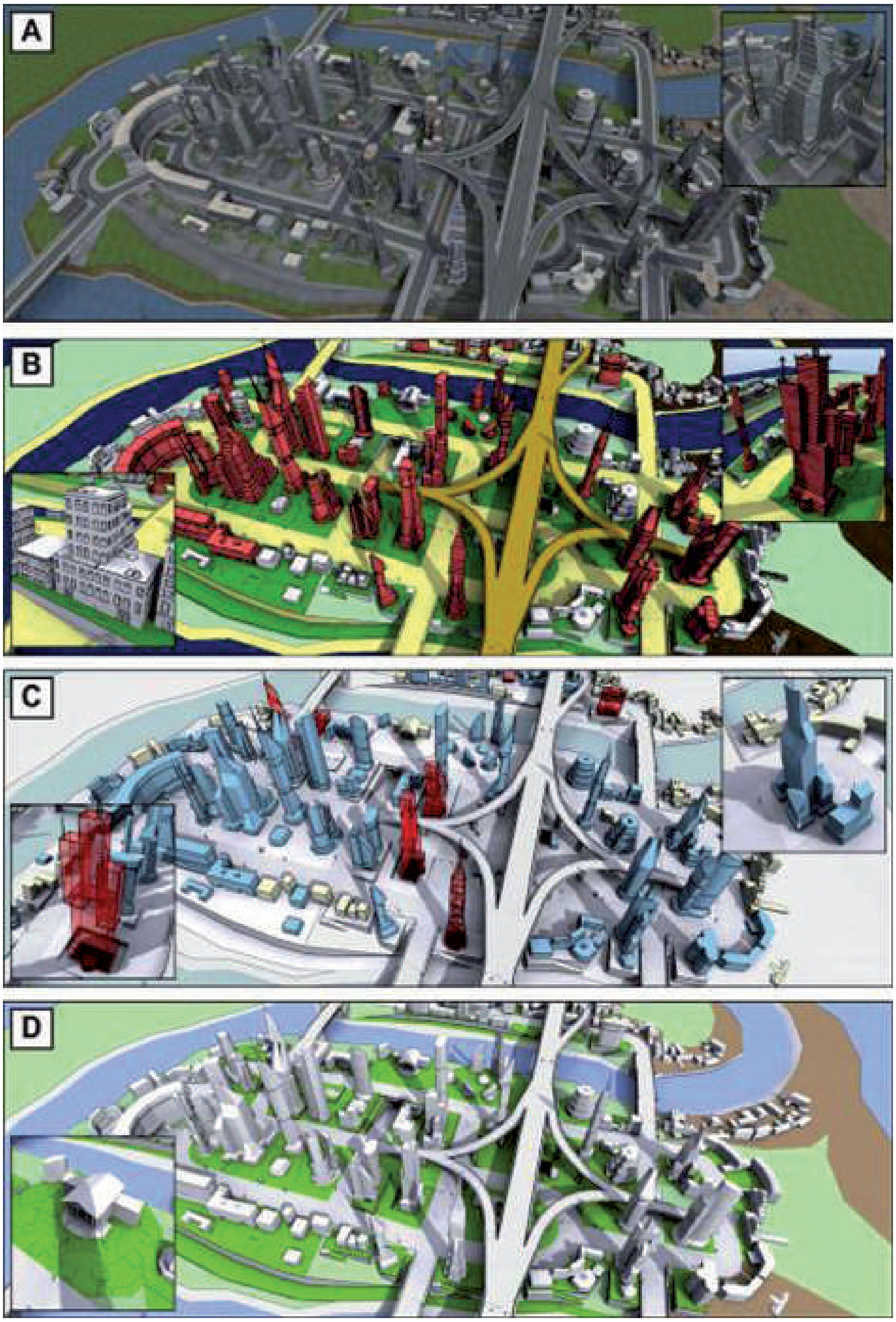


Fig. 6: Possible scenarios for a cartography-oriented visualization of a virtual 3D city model, in comparison to (A) a photorealistic version: (B) thematic overview of feature classes, (C) highlighting of construction states for building models, (D) highlighting the distribution of green spaces.

application domains, a validation as part of a user study is required.

To further improve the visualization quality, we see two general approaches: (1) the implementation of additional typographical characteristics of cartography for certain feature types, for example to communicate land use information or visualize street networks in an abstract style, and (2) the incorporation of interactive concepts for a dynamic, context-aware abstraction, for example to enable a saliency-guided visualization.

Currently, advanced stylizations as presented in this work are proprietary components of individual systems. To promote technical progress and to make these stylizations broadly available, future work needs to develop sufficiently expressive, standardized style description languages that are designed to improve the effectiveness and efficiency of geoinformation communication.

References

Agrawala, M. & Stolte, C., 2001: Rendering Effective Route Maps: Improving Usability Through Generalization. – Proc. ACM SIGGRAPH: 241–249.

Akenine-Möller, T. & Haines, E., 2008: Real-Time Rendering 3rd Edition. – A. K. Peters, Ltd., Natick, MA, USA.

Brewer, C.A., 1994: Color Use Guidelines for Mapping and Visualization. – Elsevier Science, Remote Sensing and Spatial Information Science, ch. **7:** 123–147.

Döllner, J. & Walther, M., 2003: Real-Time Expressive Rendering of City Models. – IEEE Information Visualisation: 245–250.

Döllner, J., Buchholz, H., Nienhaus, M. & Kirsch, F., 2005: Illustrative Visualization of 3D City Models. – Visualization and Data Analysis, SPIE: 42–51.

Foerster, T., Stoter, J.E. & Kobben, B., 2007: Towards a formal classification of Generalization operators. – Proc. ICC.

Glander, T. & Döllner, J., 2008: Techniques for Generalizing Building Geometry of Complex Virtual 3D City Models. – Advances in 3D Geoinformation Systems: 381–400.

Gooch, A., Gooch, B. & Costa-Sousa, M., 2011: Illustrative Visualization: The Art and Science of Non-Photorealistic Rendering. – Taylor & Francis Ltd.

Grünreich, D., 1992: ATKIS – A topographic information system as a basis for GIS and digital cartography in Germany. – From digital map series to geoinformation systems. – Geologisches Jahrbuch Reihe A.

Hildebrandt, D., 2011: Towards Service-Oriented, Standards- and Image-Based Styling of 3D Geovirtual Environments. – HPI Research School on Service-oriented Systems Engineering **46**, Universitätsverlag Potsdam.

Hildebrandt, D. & Döllner, J., 2010: Service-oriented, standards-based 3D geovisualization: Potential and challenges. – Computers, Environment and Urban Systems **34** (6): 484–495.

ISO 19101, 2002: Geographic information – Reference model. – ISO, Geneva, Switzerland.

Jobst, M. & Döllner, J., 2008: 3D city model visualization with cartography-oriented design. – REAL CORP: 507–400.

Kolbe, T.H., 2009: Representing and Exchanging 3D City Models with CityGML. – 3D Geoinformation Sciences: 15–31.

Kyprianidis, J.E. & Döllner, J., 2008: Image Abstraction by Structure Adaptive Filtering. – EG UK Theory and Practice of Computer Graphics: 51–58.

Lupp, M., 2007: Styled Layer Descriptor. – Profile of the Web Map Service Implementation Specification, Version 1.1.0. – Open Geospatial Consortium Inc.

Lynch, K., 1960: The image of the city. – MIT Press.

MacEachren, A.M., Edsall, R., Haug, D., Baxter, R., Otto, G., Masters, R., Fuhrmann, S. & Qian, L., 1999: Virtual environments for geographic visualization: potential and challenges. – Proc. NPIVM: 35–40.

Métral, C., Falquet, G. & Cutting-Decelle, A., 2009: Towards Semantically Enriched 3D City Models: An Ontology-Based Approach. – GeoWeb 2009 Academic Track – Cityscapes'.

Müller, M., 2006: Symbology Encoding Implementation Specification, Version 1.1.0. – Open Geospatial Consortium Inc.

Myers, K. & Bavoil, L., 2007: Stencil routed A-Buffer. – ACM SIGGRAPH 2007 Sketches.

Neubauer, S. & Zipf, A., 2009: 3D-Symbology Encoding Discussion Draft, Version 0.0.1. – Open Geospatial Consortium Inc.

Nienhaus, M. & Döllner, J., 2003: Edge-Enhancement – An Algorithm for Real-Time Non-Photorealistic Rendering. – Journal of WSCG **11** (2): 346–353.

Nienhaus, M. & Döllner, J., 2004: Blueprints – Illustrating Architecture and Technical Parts using Hardware-Accelerated Non-Photorealistic

Rendering. – Graphics Interface, AK Peters: 49–56.

SEMMO, A., KYPRIANIDIS, J.E. & DÖLLNER, J., 2010: Automated Image-Based Abstraction of Aerial Images. – AGILE International Conference on GI Science: 359–378.

SEMMO, A., TRAPP, M. & DÖLLNER, J., 2011: Ansätze zur kartographischen Gestaltung von 3D-Stadtmodellen – **31.** Wissenschaftlich-Technische Jahrestagung der DGPF: 473–482.

WARE, C., 2004: Information Visualization: Perception for Design. – Morgan Kaufmann Publishers, San Francisco.

WOLFF, M. & ASCHE, H., 2009: Geospatial Crime Scene Investigation – From Hotspot Analysis to Interactive 3D Visualization. – ICCSA: 285–299.

ZIPF, A., 2005: Using Styled Layer Descriptor (SLD) for the Dynamic Generation of User- and Context-Adaptive Mobile Maps – A Technical Framework. – W2GIS: 183–193.

Address of the Authors:

AMIR SEMMO, DIETER HILDEBRANDT, MATTHIAS TRAPP und Prof. Dr. JÜRGEN DÖLLNER, Hasso-Plattner-Institut, Universität Potsdam, Fachgebiet Computergrafische Systeme, D-14482 Potsdam, Tel.: +49-331-5509-170, Fax: +49-331-5509-172, e-mail: {amir.semmo} {dieter.hildebrandt} {matthias.trapp} {juergen.doellner}@hpi.uni-potsdam.de

Manuskript eingereicht: März 2012
Angenommen: Mai 2012

Deutsche Geodätische Kommission (DGK)

DGK Wissenschaftspreis 2012

Die Deutsche Geodätische Kommission (DGK) hat den Preisträger des DGK-Preises 2012 bekannt gegeben: Es ist Dr.-Ing. PETER STEIGENBERGER von der Technischen Universität München. Der DGK-Preis würdigt die national wie international sehr anerkannten Leistungen von PETER STEIGENBERGER (Jahrgang 1976) auf dem Gebiet der Langzeitanalyse von GPS-Daten sowie der Auswertung von GNSS-Signalen der Galileo GIOVE und der chinesischen Compass-Satelliten. Er hat in seinen Arbeiten wesentliche Beiträge zur Bestimmung der troposphärischen Signalverzögerungen, zur Evaluierung des Effektes von atmosphärischen Auflastdeformationen, zur Schätzung subtäglicher Erdrotations- und kurzperiodischer Nutationsparameter, zur Untersuchung lokaler Mehrwegeffekte und deren Einfluss auf lokale Verbindungsvektoren und zur Analyse der Genauigkeit von mit GPS bestimmten vertikalen Geschwindigkeiten geliefert.

Der Wissenschaftspreis der Deutschen Geodätischen Kommission wurde im Jahre 2011 zur Förderung hochqualifizierter Wissenschaftlerinnen und Wissenschaftler ins Leben gerufen und wird auf der INTERGEO 2012 überreicht. Der DGK-Preis ist für junge Wissenschaftlerinnen und Wissenschaftler mit bereits ausgeprägt internationalem Profil gedacht. Als Grundlage für die Beurteilung der Preiswürdigkeit dient die wissenschaftliche Gesamtleistung einschließlich der nach der Promotion veröffentlichten wissenschaftlichen Publikationen. Der Preis wird alle zwei Jahre vergeben und besteht aus einer Urkunde und einem Preisgeld in Höhe von 2.000,– €.

32. Wissenschaftlich-Technische Jahrestagung 2012

Bericht zur Jahrestagung der DGPF, 14.–17. März 2012 am Hasso-Plattner-Institut (HPI) in Potsdam

Die diesjährige Wissenschaftlich-Technische Jahrestagung der DGPF stand unter dem Motto: *Erdblicke – Perspektiven für die Geowissenschaften*.

Die lokale Vorbereitung und Unterstützung der Tagung wurde von Prof. Dr. rer. nat. JÜRGEN DÖLLNER und seinem Team, insbesondere seiner Assistentin SABINE BIEWENDT, in vorbildlicher Art und Weise übernommen. Auf diesem Weg bedanken sich alle Teilnehmer der Jahrestagung nochmals ganz herzlich für die perfekt vorbereitete Tagung und das Exkursions- und Rahmenprogramm.

Tutorien, Exkursionen und Rahmenprogramm

Am 14.3.2012 wurde die Veranstaltung durch die Tutorien *„Möglichkeiten und aktueller Forschungsstand hyperspektraler Fernerkundung“* und *„Web-basierte Visualisierung massiver 3D-Stadtmodelle“* eingeleitet. Mit 19 bzw. 16 angemeldeten Teilnehmern waren beide Tutorien gut besucht und vermittelten den Gästen interessante Fakten und Zusammenhänge.

Das Rahmenprogramm umfasste am 14.3. eine kleine Exkursion zu den Bildflugzeugen der Firma BSF Swissphoto GmbH und zum Regionalflugplatz Schönhagen sowie am 17.3. eine große Exkursion zu den Referenzpunkten Deutschlands in Berlin und seiner Umgebung.

DOI: 10.1127/1432-8364/2012/0132 1432-8364/12/0132 $ 7.00

Beim zwanglosen Vorabendtreffen im Restaurant „Der Hammer“ im historischen Kutschstallensemble in Potsdam trafen sich bereits viele der Teilnehmer, um ebenso wie bei der Abendveranstaltung am 15. März in der Weinscheune im Krongut Bornstedt Erinnerungen aufzufrischen und neue Aktivitäten vorzubereiten.

Eröffnungsveranstaltung am 15. März 2012

Die Präsidentin der DGPF, Prof. Dr. Cornelia Glässer, bedankte sich bei den lokalen Organisatoren für die perfekte Vorbereitung und wünschte allen Teilnehmern eine erfolgreiche Tagung. Der Präsident des HPI, Prof. sc. nat. Christoph Meinel, begrüßte die Teilnehmer und erläuterte die Entstehungsgeschichte des Hasso-Plattner-Instituts. Der Name erinnert an ein Vorstandsmitglied der Firma SAP, der dieses Universitätsinstitut für die Entwicklung des IT-Nachwuchses gegründet hatte. Im Rahmen der Eröffnungsveranstaltung wurden die Preisträger des Karl-Kraus-Nachwuchsförderpreises für Photogrammetrie, Fernerkundung und Geoinformation geehrt. Der Preis wurde wieder gemeinsam von der SGPF, der OVG und der DGPF verliehen. Die Moderation übernahm Görres Grenzdörffer. Nähere Einzelheiten finden sich in seinem Bericht. Leider konnte der Hansa-Luftbild-Preis in diesem Jahr mangels geeigneter Artikel nicht verliehen werden. Die Präsidentin motivierte daher alle jungen Wissenschaftler zu Beiträgen in der PFG.

In seinem Plenarvortrag „*Echtzeitgeoinformation von Satelliten und fliegenden Robotern*“ informierte Wolfgang Förstner, Universität Bonn, Institut für Geodäsie und Geoinformation, über aktuelle Verfahren zur Berechnung von 3D Modellen aus Satellitenbildern, Luftbildern und Daten autonom fliegender Drohnen. Die automatisierte Bildorientierung erfolge inzwischen operationell für große homogene Bildverbände, Lösungsmöglichkeiten für beliebige Orientierungen und unter Nutzung visueller Navigationsmöglichkeiten werden jedoch noch gesucht. Die 3D Modellierung von Pflanzen sei weiterhin ein sehr gefragtes Thema.

Vortragsprogramm

Am 15. und 16. März wurden die Fachvorträge der Arbeitskreise parallel in drei Sitzungsräumen gehalten. Einzelheiten zu den Inhalten und Ergebnissen finden sich in den Berichten der Arbeitskreisleiter.

Fachfirmenausstellung und Posterbeiträge

Das Foyer bot den Raum für die Firmenausstellung, die wissenschaftlichen Poster, die zusätzlichen Präsentationen der Firmen und den Treffpunkt während der ganzen Tagung.

Mitgliederversammlung am 15. März 2012

Die ordentliche Mitgliederversammlung der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V. fand am 15. März von 17:45 bis 19:00 Uhr statt. Dort wurden folgende Tagesordnungspunkte behandelt:

1. Begrüßung und Genehmigung der Tagesordnung
2. Ehrungen
3. Entwicklung der Stiftung PFGeo
4. Geschäftsbericht des Vorstandes
5. Bericht der Kassenprüfer
6. Entlastung des Vorstands
7. Wahl des Präsidenten und der übrigen Mitglieder des Vorstandes
8. Wahl der Kassenprüfer
9. Verschiedenes

Vor Eintritt in die Tagesordnung gedachten die Anwesenden den verstorbenen Mitgliedern: Prof. Dr. Hans-Karsten Meyer, Prof. Dr. Rudolf Förstner und Dipl.-Ing. Wolfhard Elsner.

Die Tagesordnung wurde ohne Änderungswünsche angenommen.

Der Vorstand schlug der Mitgliederversammlung vor, Prof. Dr.-Ing. Dr. h.c. mult. Wolfgang Förstner zum Ehrenmitglied zu ernennen. Dieser Vorschlag wurde ohne Gegenstimmen mit einer Enthaltung angenommen.

Die Präsidentin berichtete über die positive Entwicklung der Stiftung PFGeo. Nach dem anfänglichen Aufbau des Stiftungskapitals konnte die zunächst unselbständige Stiftung nach Erreichen von mehr als 50.000 Euro Stiftungsvermögen am 22. Dezember 2011 von der Bezirksregierung Münster als selbständige Stiftung privaten Rechts anerkannt werden.

Die Präsidentin bedankte sich ausdrücklich für die Tätigkeit des ausgeschiedenen Hauptschriftleiters Helmut Mayer. Die Schriftleitung wird inzwischen von Wolfgang Kresse wahrgenommen. Er bedankte sich bei seinen Vorgängern für die ausgezeichnete Vorarbeit und bei den Gutachtern, die die Qualität der Zeitschrift PFG garantieren. Dem Schriftleiter Franz Rottensteiner gebührte ein besonderer Dank, weil gerade in diesem Jahr die PFG besonders viele Beiträge enthält, deren Begutachtung von ihm koordiniert wird. Der Dank des Vorstandes geht an die Schriftleitung und den Redaktionsbeirat, die erheblich zur hohen Qualität der Zeitschrift beigetragen haben.

Cornelia Glässer erläuterte die aktuelle Mitgliederstruktur. Die Mitgliederentwicklung der ordentlichen Mitglieder ist weiterhin leicht fallend. Daher müssen verstärkt junge Mitglieder geworben werden.

Sie informierte weiterhin über die Nutzung der Internetseiten der DGPF. Jährlich seien über 9000 Zugriffe auf die Homepage der Gesellschaft registriert worden. Für die Mitglieder werden inzwischen sowohl die digitalen Versionen der PFG Hefte als auch die Tagungsbände der Jahre 2001 bis 2011 als pdf-Dokument zum Herunterladen im Internet angeboten. Sie dankte Thomas Kolbe und seinem Team sowie Manfred Wiggenhagen für die Aufbereitung dieser Daten.

Vorstand der DGPF ab 2012: Wolfgang Kresse, Eberhard Gülch, Uwe Stilla, Thomas Kolbe, Herbert Krauß (von links)

Der Schatzmeister erläuterte den Geschäftsbericht der Gesellschaft. Die anwesenden Mitglieder stimmten dem Bericht ohne Enthaltungen und ohne Gegenstimmen zu. Der Kassenprüfer Karl Heiland bestätigte den einwandfreien Zustand der Kassenführung der Gesellschaft. Hans Mohl hatte gemeinsam mit Karl Heiland am 14.2.2012 die Kassenbücher der DGPF geprüft und keine Beanstandungen gefunden.

Zur Wahl des neuen Vorstandes übernahm Heinz-Jürgen Przybilla die Sitzungsleitung und bat um Vorschläge für den neuen Präsidenten. Aus der Mitgliederversammlung wurde Prof. Dr. rer. nat. Thomas Kolbe vorgeschlagen. Weitere Kandidaten wurden nicht genannt.

Thomas Kolbe stellte seine Kandidaten für den neuen Vorstand vor:

- Vizepräsident: Prof. Dr.-Ing. Uwe Stilla
- Sekretär: Prof. Dr.-Ing. Eberhard Gülch
- Hauptschriftleiter: Prof. Dr.-Ing. Wolfgang Kresse
- Schatzmeister: Dr.-Ing. Herbert Krauss

Thomas Kolbe wurde mit drei Enthaltungen ohne Gegenstimmen zum neuen Präsidenten gewählt. Die Wahl des Vorstandes erfolgte im Block mit einer Gegenstimme und einer Enthaltung.

Die Mitgliederversammlung stimmte dem Vorschlag zu, Cornelia Glässer in den Beirat des Vorstandes aufzunehmen. Thomas Kolbe dankte der bisherigen Präsidentin für ihre sehr erfolgreiche Arbeit und dem bisherigen Sekretär, Manfred Wiggenhagen, für seine 14-jährige sehr gewissenhafte und erfolgreiche Tätigkeit als Sekretär und Webmaster der Gesellschaft. Als neue Kassenprüfer wurden Karl Heiland und Karsten Jacobsen von den Anwesenden einstimmig ohne Gegenstimme mit einer Enthaltung gewählt. Cornelia Gläser dankte allen Anwesenden für die Teilnahme und Mitwirkung und schloss die Versammlung um 19:00 Uhr.

Die 33. Wissenschaftliche Jahrestagung der DGPF wird vom 27.2. bis 1.3.2013 an der Albert-Ludwigs-Universität Freiburg als Dreiländertagung D-A-CH durchgeführt.

Manfred Wiggenhagen, Hannover

Firmenausstellung

Die diesjährige Jahrestagung fand in den Räumlichkeiten des Hasso-Plattner-Institutes in Potsdam-Griebnitzsee statt. Insgesamt 13 Aussteller nutzten die Möglichkeit, ihre neuen Produktentwicklungen vorzustellen, auf bereits im Einsatz befindliche Gerätetechnik hinzuweisen oder aus ihrem Verlagsprogramm Beispiele vorzulegen.

Die Organisatoren platzierten die Firmenausstellung unmittelbar vor den Vortragsräumen und nicht weit von der Registrierung zur Tagung. Das ermöglichte jedem Besucher die Ausstellung täglich mehrfach zu durchqueren und dabei immer wieder neue Anknüpfungspunkte für Fachdiskussionen mit den Ausstellern zu finden.

Zur Tagung konnten die Stände der folgenden Firmen besucht werden: bsf swissphoto, DELPHI, eta AG engineering, IGI, ILV, MULTIROTOR, Hexagon (Leica, ZI-Imaging), Microsoft UltraCam, RapidEye, mit Plottechnik die Firmen AMCAD & Rauch GmbH, EPSON und die kartographische Firma TOPKART & TERRA NOSTRA.

Eckhardt Seyfert, Werder/Havel

Berichte der Arbeitskreise der DGPF

Optische 3D-Messtechnik

Arbeitsgebiete – Terms of Reference:

- Nahbereichsphotogrammetrie
- Terrestrisches Laserscanning
- Sensorintegration und Systemkalibrierung
- Aufnahme- und Auswertestrategien
- Prozessorientierte Auswertung und Automation
- 3D-Modellierung

Bericht von der Jahrestagung

An der 32. wissenschaftlich-technischen DGPF-Tagung an dem Hasso-Plattner-Institut der Universität Potsdam konnte sich der Arbeitskreis durch drei Sitzungen präsentieren. Die zwei normalen Sitzungen des Arbeitskreises fanden am 15. und 16. April 2012 statt, während eine spezielle Sitzung mit dem Thema *Automatische 3D-Objektrekonstruktion durch Pixel-basierte Matchingverfahren* zusammen mit dem AK Bildanalyse und Bildverstehen am 16. April organisiert werden konnte.

Die präsentierten acht Beiträge der beiden normalen Sitzungen hatten in diesem Jahr ihren Schwerpunkt vorwiegend in der Nahbereichsphotogrammetrie, doch mobiles und terrestrisches Laserscanning sowie optische Messtechnik in Schrägluftbildern wurden ebenfalls behandelt. Ca. 75 Teilnehmer nahmen an der ersten Sitzung über Nahbereichsphotogrammetrie teil, während bei der zweiten Sitzung, die als letzte Sitzung der Tagung am Freitagnachmittag lief, nur mehr ca. die Hälfte der Teilnehmer erschienen.

Die erste Sitzung wurde mit dem 1. Preisträger des Karl-Kraus-Nachwuchspreises eröffnet. Frank Liebold (TU Dresden) präsentierte seine Abschlussarbeit mit dem Thema *Entwicklung einer Methode zur Bestimmung und Korrektur der Verzerrungen bei der Projektion auf eine doppelt gekrümmte Fläche*. Für einen interaktiven Fahrsimulator wurde ein 3D-Modell von dessen gekrümmter Projektionsfläche erstellt sowie die innere und äußere Orientierung der sich im Inneren befindlichen sechs Projektoren bestimmt. Mit Hilfe einer Differentialentzerrung kann ein für den Betrachter unverzerrt erscheinendes und zusammenhängendes Panoramabild auf die Leinwand projiziert werden. In einem weiteren Beitrag der TU Dresden stellte Katja Richter ein voll-reflektives Kamerasystem für Nahbereichsanwendungen vor. Mit dem dafür entwickelten geometrischen Modell kann die stark asymmetrische Verzeichnung des ausschließlich auf Spiegeln basierenden Objektivs modelliert und die Kamera für Anwendungen, welche die Abbildung eines breiten Spektralbereichs erfordern, eingesetzt werden. Anschließend zeigte John Moré (BSF Swissphoto, Berlin) anhand von Beispielen, wie man aus orientierten Schrägluftbildern unter Annahme von Bedingungen die 3D-Geometrie monoskopisch erfassen kann. Zum Abschluss der ersten Sitzung wurde von Norbert Haala (Uni Stuttgart) das neue DGPF-Projekt „Mobile Mapping Test" vorgestellt, das von den beiden Arbeitskreisen Sensoren und Plattformen sowie Optische 3D-Messtechnik in Zusammenarbeit mit den beiden deutschsprachigen Partnergesellschaften SGPF und OVG federführend bearbeitet wird. In einer kontrollierten Umgebung (möglicherweise Testfeld Vaihingen/Enz) sollen automatisch 3D-Daten durch Mobile-Mapping-Fahrzeuge erfasst und anschließend analysiert werden. Die technischen Spezifikationen des Tests bzw. die Themen der Arbeitsgruppen müssen noch präziser definiert und die Testumgebung muss noch detailliert erfasst werden. Systemanbieter und Anwender werden eingeladen, an dem Test teilzunehmen.

In der zweiten Sitzung deckten die Beiträge Themen aus der Nahbereichsphotogrammetrie und des mobilen Laserscannings ab. Zuerst präsentierte Carsten Götz (TU München) Untersuchungen zum Genauigkeitspotential des räumlichen Rückwärtsschnitts mit zwei Kameras in einem räumlichen Testfeld, in dem die Bewegung einer und zweier Kameras entlang eines Pfades durch eine bekannte Umgebung simuliert wurde. Im zweiten Vortrag stellte Nalani Hetti Arachchige (TU Dresden) ein Verfahren vor, mit dem man Gebäudefassaden in Punktwolken von mobilen Laserscannern automatisch durch Klassifikation von Boden- und Objektpunkten und anschließender fortgeschrittener Klassifikation von Gebäudepunkten detektieren kann. Anschließend zeigte Dirk Stallmann (New Class Consulting Ltd., Paphos, Zypern) das Potential und die Probleme bei der automatischen Texturierung von 3D-Modellen für Anwendungen in Architektur und Archäologie auf. Bei der Texturierung bedeutet Automation höhere Effizienz, aber noch geringe Qualität im Vergleich zur manuellen Texturierung mit entsprechender Software wie Cinema 4D. Zum Abschluss des Tagungsprogrammes präsentierte Ludwig Hoegner (TU München) das geometrische Potential eines Kinect-Sensorsystems bei der Tiefengenauigkeit, in dem er den Sensor verschiedenen Testszenarien unterzog. Dabei zeigte sich, dass die Genauigkeit ab einer Distanz von 4 m schlechter als 1 cm wurde.

Die mit 80 Teilnehmern gut besuchte, gemeinsame Sitzung der Arbeitskreise Bildanalyse und Bildverstehen sowie Optische 3D-Messtechnik hatte sich in vier Beiträgen, in denen Weltraumaufnahmen, Luftbilder und terrestrische Aufnahmen verwendet wurden, mit der automatischen 3D-Objektrekonstruktion durch Pixel-basierte Matchingverfahren befasst. Klaus Gwinner (DLR, Berlin) zeigte im ersten Vortrag, wie das Least-Squares Matching mit polynomialer lokaler Bildtransformation als flächenbasierte Stereo-Rekonstruktion Oberflächenmodelle vom Planeten Merkur generierte. Eine Kombination von ungeordneten terrestrischen und UAV-gestützten Bilddaten auch mit großer Basis benutzte Helmut Mayer (Universität der Bundeswehr, München), um automatisch dichte 3D-Punktwolken und 3D-Modelle von Landschaftsbereichen bzw. von Gebäuden durch semi-globale Matching-Algorithmen abzuleiten. Thomas Kersten (HCU Hamburg) zeigte anhand von verschiedenen praktischen Beispielen Möglichkeiten und Potentiale auf, die open-source Software und Webservices für die automatische 3D-Objektrekonstruktion aus digitalen Bilddaten im Vergleich zu terrestrischem Laserscanning für Anwendungen in Architektur, Denkmalpflege und Archäologie bieten. Im letzten Vortrag dieser speziellen Sitzung präsentierte Sandra Beckmann (BSF Swissphoto, Berlin) das Potential der Dense Image Matching Algorithmen (Socet Set NGATE und Inpho Match AT) am Beispiel von UltraCAM-Luftbildern, die mit Referenzdaten verglichen wurden. In diesen vier Beiträgen konnte deutlich gezeigt werden, dass durch das Pixel-basierte Matchingverfahren eine effiziente Konkurrenz für das Laserscanning bei der automatischen 3D-Datenerfassung entsteht.

Insgesamt war das Vortragsprogramm des Arbeitskreises wieder sehr abwechslungsreich. Trotz der Oldenburger 3D-Tage, die als jährliches Highlight des Arbeitskreises sechs Wochen vorher stattfanden, waren die drei Sitzungen mit zwölf Beiträgen vollständig belegt. Die Sitzungen waren gut besucht und die Beiträge wurden sehr intensiv diskutiert.

Weitere und geplante Aktivitäten

- Am 13. und 14. Februar 2013 werden die 12. Oldenburger 3D-Tage stattfinden, bei denen der Arbeitskreis Optische 3D-Messtechnik als Mitveranstalter mitwirken wird (www.jade-hs.de/3dtage).
- Die beiden Arbeitskreise Sensoren und Plattformen sowie Optische 3D-Messtechnik bereiten das DGPF-Projekt *Mobile Mapping Test* vor.

Thomas Kersten, Hamburg

Sensoren und Plattformen

Arbeitsgebiete – Terms of Reference

- Sensorkonzepte für Anwendungen in der Photogrammetrie und Fernerkundung auf unterschiedlichen Plattformen vom Satelliten bis hin zum Nahbereich
- Kooperation mit den Anwendern/Datennutzern zur anwendungsbezogenen Optimierung der Sensorsysteme und Sensorkomponenten sowie zur Ableitung/Definition neuer Aufgabenstellungen für Sensorsysteme
- Modellierung/Simulation der Sensoren und Aufnahmesysteme auch unter Berücksichtigung der Auswertealgorithmen zur Objekt- und der Dateninterpretation
- Kalibrierung und Validierung von Sensorsystemen und erfassten Daten einschließlich der Dokumentation der jeweiligen Charakteristiken im Hinblick auf die vorgesehenen Anwendungen

Bericht von der Jahrestagung

Die Nutzung unbemannter Luftfahrzeuge für photogrammetrische Anwendungen ist mittlerweile ein fast schon traditioneller Schwerpunkt des Arbeitskreises Sensoren und Plattformen. Entsprechend wurde die Sitzung des Arbeitskreises anlässlich der 32. wissenschaftlich-technischen DGPF-Tagung am Hasso-Plattner Institut der Universität Potsdam durch einen Beitrag zum Genauigkeitspotenzial der photogrammetrischen Bildauswertung für Daten unbemannter Luftfahrzeuge (UAVs) eröffnet. Michael Cramer (Universität Stuttgart) diskutierte dabei das Genauigkeitspotential der photogrammetrischen Datenauswertung für UAVs und demonstrierte die dabei notwendige Anpassung der photogrammetrischen Auswertekette am Beispiel verschiedener Testflüge.

Ein weiterer Themenkomplex des Arbeitskreises ist neben dem Design großformatiger digitaler Luftbildkameras die notwendige Kalibrierung dieser Sensorsysteme. Rongfu Tang, Universität Stuttgart, schlug hierfür den Einsatz von Legendre Polynomen vor. In seinem Beitrag zu New Mathematical Self-Calibration Models in Aerial Photogrammetry zeigte er, dass die Verzeichnung digitaler Flächenkameras durch zusätzliche Parameter auf Basis von Legendre Polynomen sehr gut kalibriert werden kann. Die Leistungsfähigkeit des vorgestellten Ansatzes, der eine Verallgemeinerung der bekannten Parametermodelle von Ebner und Grün darstellt, wurde anhand von Testdaten der DMC II, UltraCamXp und DigiCam demonstriert. Der geometrischen Kalibrierung großformatiger Luftbildkameras widmete sich auch der anschließende Beitrag zum Potential neuer großformatiger digitaler Kameras von Karsten Jacobsen von der Leibniz Universität Hannover. Am Beispiel der neuen großformatigen digitalen Kameras DMC II 140, 230 und 250 sowie der UltraCam Eagle wurde zunächst die Steigerung der Wirtschaftlichkeit durch den Einsatz dieser Systeme zur Bildaufnahme diskutiert. Zur Verifikation dieses Potentials untersuchte der Beitrag die Spezifikationen der Kameras hinsichtlich der geometrischen Genauigkeit und effektiven Pixelgröße. Hierfür wurden die genannten Systeme anhand mehrerer Befliegungen validiert, um unter anderem basierend auf der möglichen Objekterkennbarkeit eine Grundlage für Projektplanungen zu geben.

Die Übernahme des GIS-Anbieters Intergraph durch den Hexagon-Konzern, führte zum Zusammenschluss der Sensoranbieter Leica Geosystems und Z/I Imaging. GERHARD LAUENROTH (Hexagon Geospatial Solutions Division) stellte in seinem Beitrag die daraus folgenden Auswirkungen auf das angebotene Produkt- und Anwendungsspektrum vor, wobei insbesondere das Portfolio der großformatigen Flächen- und Zeilensensoren, sowie die neuen Mittelformat-Kameras zur Aufnahme von hoch aufgelösten 4-Band-Bildern diskutiert wurden. Zum Abschluss der Sitzung erläuterte ANDREAS BRUNN ein Verfahren zur Kalibrierung des aus fünf baugleichen Erdbeobachtungssatelliten bestehenden Satellitensystems RapidEye. Im Rahmen des so genannten Side-Slither-Manövers werden dabei im Orbit die Detektorzeilen der Kameras durch eine 90° Drehung des Satelliten parallel zur Flugrichtung ausgerichtet. Nach diesem Manöver bildet jeder Detektor der Zeile den identischen Bodenpunkt ab, so dass anhand der nahezu gleichbleibenden Eingabe eine Kalibrierung der Kamera erfolgen kann.

Weitere und geplante Aktivitäten

- Die beiden Arbeitskreise Sensoren und Plattformen sowie Optische 3D-Messtechnik bereiten das DGPF-Projekt *Mobile Mapping Test* vor.

NORBERT HAALA, Stuttgart

Bildanalyse und Bildverstehen

Arbeitsgebiete – Terms of Reference

- Geometrische Bildverarbeitung
- Merkmalsextraktion und –nutzung
- Semantische Modellierung

Bericht von der Jahrestagung

Im Rahmen der 32. wissenschaftlich-technischen DGPF-Tagung am Hasso-Plattner Institut der Universität Potsdam fanden zwei Sitzungen des Arbeitskreises statt. Bei der ersten Sitzung handelte es sich um eine reguläre Sitzung des Arbeitskreises mit verschiedenen Beiträgen zur Bildanalyse. Die zweite Sitzung mit dem Thema *Automatische 3D-Objektrekonstruktion durch Pixel-basierte Matchingverfahren* wurde zusammen mit dem AK Optische 3D-Messtechnik abgehalten (s.o.).

Die vorgestellten fünf Beiträge der gut besuchten regulären Arbeitskreissitzung spannten einen thematisch breiten Fächer auf. Zu Beginn der Sitzung stellte der dritte Preisträger des Karl-Kraus-Nachwuchspreises ALESSANDRO BECK (ETH Zürich) die Ergebnisse seiner Bachelorarbeit *Automatische Detektion geodätischer Ziele* vor. Im Rahmen dieser wurde untersucht, inwieweit geodätische Ziele automatisch mittels moderner Klassifizierungsverfahren wie Support Vector Machines (SVM) erkannt werden können. Der zweite Beitrag *Automatische Interpretation von Bebauungsstrukturen in topographischen Karten unter Nutzung von Bildanalyse und maschinellem Lernen* wurde von ROBERT HECHT (Leibniz-Institut für ökologische Raumentwicklung, Dresden) gehalten. Dieser Beitrag beschäftigte sich ebenfalls mit dem Einsatz von Klassifizierungsverfahren wie SVM zur Auswertung von Rasterkarten, die eine Datengrundlage auch für zurückliegende Zeitschritte ermöglichen. Der dritte Beitrag von GILDARDO LOZANO-VEGA (Fachhochschule Mainz) stellte ein Projekt zur Analyse von Pollenbelastungen (*Image based pollen detection and quantitative analysis*) vor. Die Reihe der Präsentationen wurde von MARTIN DRAUSCHKE (Universität der Bundeswehr, München) mit dem Thema *Interpretation von Fassadenbildern* fortgesetzt. In dem Beitrag wurde ein Überblick über drei aktuelle Arbeiten zur datengetriebenen Fassadeninterpretation gegeben. Der Beitrag *Linienbasiertes Matching von 3D-Gebäudemodellen mit IR-Luftbildsequenzen zur automatischen Texturgewinnung* von DOROTA IWASZCZUK (TU München), der die Verbesserung der Äußeren Orientierung (GPS/INS) zum Ziel hatte, beschloss die Sitzung.

Die zweite Sitzung mit dem Schwerpunkt auf Matchingverfahren zusammen mit dem AK Optische 3D-Messtechnik umfasste Beiträge, die sich mit der Verbesserung von bestehenden Matchingverfahren sowie Verfahren zur Ableitung dichter Punktwolken und Oberflächen aus Bildern und deren praktischem Einsatz. Details hierzu sind im Bericht des Arbeitskreises Optische 3D-Messtechnik gegeben (s.o.).

Zusammenfassend ist festzuhalten, dass das Vortragsprogramm des Arbeitskreises wieder sehr abwechslungsreich war und insbesondere auch die gemeinsame themenorientierte Durchführung von Sitzungen zu intensiven Diskussionen führte.

Uwe Weidner, Karlsruhe

Hyperspektrale Fernerkundung

Arbeitsgebiete – Terms of Reference

- Innovative Methoden und Verfahren in der hyperspektralen Fernerkundung
- Qualitative und quantitative Probleme der Datenerhebung
- Abbildende und nicht-abbildende Spektrometer
- Multimediale Wissenskommunikation und Visualisierung von Forschungsergebnissen

Bericht von der Jahrestagung

Der Arbeitskreis Hyperspektrale Fernerkundung hatte in der diesjährigen Jahrestagung in Potsdam zwei Sitzungen, die optimale Bedingungen für Diskussionen boten. Die Anzahl und Diversität der Anwendungsgebiete der eingereichten Beiträge zeigt die zunehmende Nutzung hyperspektraler Fernerkundungsdaten in Forschungsvorhaben insbesondere in Deutschland. Die diesjährigen Beiträge zeigten Forschungsergebnisse aus den Bereichen Forst, Landwirtschaft, Pflanzenzüchtung, Bodenwissenschaften, Ökologie sowie Sensortechnik.

Der erste Vortrag der Sitzung „*Hyperspektrale Fernerkundung I*“ war ein Beitrag von Florian Beyer von der Martin-Luther-Universität Halle-Wittenberg, der ein Verfahren zur Ableitung geochemischer Parameter aus dem Bedeckungsgrad von Pioniervegetation renaturierter Tagebauflächen vorstellte. Dabei wurde auf das Verfahren der linearen spektralen Entmischung näher eingegangen. Anne Clasen von der TU Berlin stellte danach eine Plattform für kontinuierliche Spektralmessungen über einem Waldbestand vor, wobei von einem 40 m hohen Kran im DLR-Testfeld DEMMIN abbildende sowie nicht-abbildende Spektralmessungen von Baumkronen durchgeführt werden. Der dritte Vortrag der Sitzung stellte Forschungsergebnisse eines Verfahrens zur Ableitung des Blattflächenindex von Weizenbeständen dar, präsentiert von Bastian Siegmann (Universität Osnabrück). Dabei wurden Daten des Flugzeug getragenen AISA-DUAL Systems verwendet, um mittels spektraler Analysen im Feld erhobener Daten auf Pflanzenparameter zu schließen. Der abschließende Vortrag der ersten Sitzung war ein Beitrag von Nicole Richter zum Einsatz von hyperspektraler Messtechnik in der Pflanzenzüchtung. Dabei wurde ein am Julius-Kühn-Institut entwickeltes Spektrometersystem Tri-Spek verwendet, um Pflanzenparameter von Roggenbeständen mittels partial-least-square Regressionsmodellen abzuleiten.

Thomas Jarmer (Universität Osnabrück) hat die zweite Sitzung des Arbeitskreises begonnen und berichtete über die Erfassung des organischen Kohlenstoffs landwirtschaftlicher Böden. Er hat besonderen Wert auf den experimentellen Teil seiner Forschung gelegt und die Wichtigkeit und Problematik der fernerkundlichen Methoden für die praxisnahe Landwirtschaft betont. András Jung (Universität Ulm) hat seine Ergebnisse im Rahmen des SpecTour Projektes erläutert und einen Überblick über die nicht-spektralen Aspekte des Projektes und über die typischen Anomalien der Messungen berichtet. René Michels (Universität Ulm) hat die Entwicklung einer neuartigen hyperspektralen Framekamera vorgestellt und deren anwendungsorientierte Möglichkeiten diskutiert.

Wir bedanken uns bei allen Teilnehmern für ihre Beiträge, ihr Interesse an den Aktivitäten des Arbeitskreises sowie die lebhaften Diskussionen.

Weitere und geplante Aktivitäten

Das Projekt SpecTour (früher White Reference Tour) wird weiterhin gepflegt und erweitert. Nach häufigen Nachfragen wird das Projekt seinen Fokus erweitern und auch bildgebende Systeme einbeziehen.

András Jung, Halle
Jonas Franke, München

Aus- und Weiterbildung

Arbeitsgebiete – Terms of Reference

- Beobachtung von Aus- und Weiterbildungsangeboten
- Beobachtung und Bewertung neuer Lehr- und Lernmethoden
- Förderung des wissenschaftlichen Nachwuchses
- Nationale und internationale Kooperationen

Bericht von der Jahrestagung

Im Gegensatz zu den vergangenen Jahren, in denen die Arbeitskreissitzung von Einzelbeiträgen getragen wurde, ist in diesem Jahr eine Podiumsdiskussion abgehalten worden. Das Thema bzw. Motto der Diskussion war „*Qualitätspakt Lehre - was kommt an*?“ Als Referenten konnten vier Experten gewonnen werden, die nicht nur aus dem Kreis der DGPF, sondern auch aus der universitären Verwaltung stammten und sich intensiv mit dem Thema Qualitätspakt und der Förderung in der Lehre beschäftigen. Die Teilnehmerrunde sah wie folgt aus:

Dr. Eva Rossmanith, Potsdam Graduate School, Universität Potsdam
Dr. Patrick Thurian, Technische Universität Berlin, Gruppenleiter Bereich Strategisches Controllings (Qualitätsmanagement, Studienreform und Kennzahlen)
Robert Eckardt, Friedrich-Schiller-Universität Jena, E-Learning für SAR-Auswertung (Potentiale neuer Lehrmethoden)
Prof. Michael Breuer, Beuth Hochschule Berlin

Vor einer sehr überschaubaren Kulisse von nur ca. 10 Teilnehmern leitete Görres Grenzdörffer die 90 minütigen Podiumsdiskussion. Zuerst stellte der Moderator den „Qualitätspakt Lehre“ der Bundesregierung vor. Dieses Programm soll zwischen 2011 und 2020 die Lehre an den Hochschulen und Universitäten verbessern. Mit der Förderung der Hochschulen im Umfang von immerhin 2 Mrd. € sollen Maßnahmen zur Verbesserung der Personalausstattung, z.B. vorgezogene oder zusätzliche Berufungen, die Einstellung von zusätzlichem Personal für Lehre oder auch die Organisation und Studentenbetreuung und Tutorien zur Betreuung in kleinen Lerngruppen unterstützt werden. Von den ca. 240 deutschen Hochschulen haben in zwei Runden immerhin 186 einen Zuwendungsbescheid bekommen.

Eva Rossmanith stellte zuerst das Förderprogramm „Teaching Professionals“ vor. Seit Mai 2010 bietet die Universität Potsdam sowohl Promovierenden (Junior Teaching Professionals) als auch PostDocs und Juniorprofessorinnen und Juniorprofessoren (Senior Teaching Professionals) eine gezielte Weiterqualifizierung in der akademischen Lehre an. 20 Nachwuchswissenschaftler(innen), die sich eine neue Lehrveranstaltung zutrauen und diese entwickeln wollen, lernen durch die Teilnahme an Programmworkshops methodischen Grundlagen und die notwendigen Fähigkeiten, um Lehrveranstaltungen didaktisch zu optimieren und sich fundierte und zertifizierte Lehrkompetenzen anzueignen. Da der mit der Lehre verbundene Zeitaufwand nicht unerheblich ist, bekommen Mitarbeiter mit einem Zeitvertrag eine zusätzliche Verlängerung um sechs Monate.

Robert Eckardt stellte anschließend das Projekt SAR-EDU vor, das vom DLR initiiert wurde, um den Fachkräftemangel im Bereich der Radarfernerkundung zu lindern (http://sar-edu.uni-jena.de/). Geplant sind in erster Linie PPT-Foliensätze mit dazugehörigen Skripten zur Erläuterung der Folien, Literaturhinweise usw. Ergänzend soll es praktische Übungen, Testdatensätze, und eventuell Videotutorials für jede Lehreinheit geben. Der Zeithorizont der Fertigstellung ist momentan noch nicht geklärt. Der Großteil der Module soll jedoch schon Ende 2012 fertig sein. In diesem Zusammenhang äußerte Robert Eckhardt die Bitte, dass externe Lehrende die Qualität des Materials evaluieren und natürlich auch in die eigene Lehre einbauen mögen. Dies kann von kleineren Aspekten wie der Integration einer praktischen Übung in die bereits existierende Lehre bis hin zu größeren Adaptionen wie dem vollständigen Aufbau der SAR-Lehre auf SAR-EDU Material gehen.

Michael Breuer berichtete über seinen Einstieg als Professor an der Beuth Hochschule Berlin. Dabei machte er deutlich, dass von verschiedenen Seiten, d.h., von den Studenten, der Forscher Community und der Verwaltung hohe Anforderungen an einen Professor gestellt werden. Alle wünschen volle Aufmerksamkeit und Zeit, was manchmal nur schwer zu realisieren ist. Die Durchführung von Lehrveranstaltungen ist für viele neu berufenen Professoren eine große Herausforderung, die durchaus den Einsatz professioneller Hilfe erforderlich machen kann, wenn auch der damit verbundene Zeitaufwand immer einen Spagat mit den sonstigen Aufgaben bedeutet.

Patrick Thurian rundete die interessante Diskussion mit seinem Beitrag zur Erstellung eines Antrages zum Qualitätspakt Lehre der Technischen Universität Berlin ab. Dabei konnte er sehr anschaulich verdeutlichen, wie ein solcher Antrag entsteht und was sich die TUB mit den Maßnahmen zur Verbesserung der Lehre im Bachelorbereich verspricht.

Weitere und geplante Aktivitäten

Neben der DGPF-Jahrestagung ist die mittlerweile 8. GIS-Ausbildungstagung am 12./13. Mai 2012 in Potsdam ein weiteres Highlight, bei dem sich viele Mitglieder des Arbeitskreises treffen und austauschen werden.

Görres Grenzdörffer, Rostock

Radarfernerkundung und Flugzeuglaserscanning

Arbeitsgebiete – Terms of Reference

- Synthetic Aperture Radar (SAR)
- SAR-Interferometrie und -Polarimetrie
- 3D-Objekterkennung aus Laserscannerdaten
- Vegetationsanalyse aus Laserscannerdaten

Bericht von der Jahrestagung

In diesem Jahr wurde vom Arbeitskreis eine Sitzung mit fünf Vorträgen organisiert, die in ausgewogenem Umfang die Themenbereiche Radarfernerkundung und Flugzeuglaserscanning behandelten. Den Anfang machte Michael Schäfer vom Institut für Geotechnik und Markscheidewesen der TU Clausthal, der über seine jüngst in einer Dissertation gemündeten Arbeiten zum Einfluss der Atmosphäre bei der Erfassung von Höhenänderungen mit differentieller Radarinterferometrie vortrug. Dies ist eine wichtige Fragestellung der SAR-Interferometrie, weil veränderter Wasserdampfgehalt in der Troposphäre sowohl die Analyse von Geländebewegungen als auch die Genauigkeit von Höhenmodellen aus nicht gleichzeitig erfassten SAR-Bildern beeinträchtigt. Danach berichtete Lukas Schack (Leibniz Universität Hannover) über Untersuchungen zur Nutzung von Nachbarschaftsinformationen zur Validierung der SAR-Tomographie. Letzteres ist eine Technik zur Unterscheidung mehrerer Objekte, die durch das SAR-Messprinzip zwar in dieselbe Entfernungs/Azimut-Auflösungszelle abgebildet werden, jedoch tatsächlich unterschiedliche Positionen in Elevationsrichtung haben. Die Tomographie ermöglicht somit eine echte 3D-Abbildung, wodurch zum Beispiel die Anzahl der für die Beobachtung von Geländebewegungen in Städten wichtigen starken Reflektoren erhöht werden kann.

Claudia Hilbert vom Lehrstuhl für Fernerkundung der Friedrich-Schiller-Universität Jena stellte eine Methode zur Erfassung der vertikalen und horizontalen Waldstruktur mittels Radarbildern und Full-Waveform Laserprofilen vor. Hierzu wurden Daten der Satellitensensoren TerraSAR-X und GLAS (ICESat) für ein Testgebiet im Thüringer Wald ausgewertet.

Die letzten beiden Vorträge waren dem Flugzeuglaserscanning gewidmet. Zunächst berichtete Sanka Perera vom Institut für Photogrammetrie und Fernerkundung der TU Dresden über seine Arbeiten zur Erkennung von Dachkonturen aus dichten Laserpunktwolken. Die Topologie der Dachstruktur wird mittels graphbasierten Verfahren aus Mengen von 3D-Linien rekonstruiert, die in einem Vorverarbeitungsschritt durch Hough-Transformation segmentiert worden sind. Zum Abschluss gab Sven Jany von der Firma MILAN Geoservice GmbH, Kamenz, zuerst einen Überblick über die vielfältigen operationellen Einsatzgebiete des Laserscannings in der Praxis, um im zweiten Teil des Vortrags auf das Projektbeispiel Muskauer Faltenbogen zu fokussieren.

Weitere und geplante Aktivitäten

Der Sensorverbund der vom DLR betriebenen Schwestersatelliten TerraSAR-X und TanDEM-X hat nunmehr die Landmassen der Erde ein erstes Mal vollständig im Einmalüberflugmodus abgedeckt (d.h., die Bilder einer Szene werden praktisch gleichzeitig gewonnen). Gegenwärtig erfolgt eine Zweiterfassung mit veränderter Basiskonfiguration. Den Abschluss sollen nächstes Jahr weitere Messungen bilden, die auf schwierige Flächen wie Gebirgslagen fokussieren. Man darf gespannt sein, welche Qualität schließlich das aus all diesen Daten prozessierte globale digitale Oberflächenmodell haben wird.

Beim Flugzeuglaserscanning ist ein Trend zu Bathymetrie-Anwendungen zu beobachten. Bislang waren nur wenige solcher speziellen Lasersysteme verfügbar, welche meist für militärische Anwendungen zum Einsatz kamen und kommen. Nunmehr wird auch in Deutschland in einigen Pilotprojekten die Eignung dieser Technik für zivile Aufgaben untersucht, wie etwa die Vermessung von küstennahen Bereichen in der Ostsee zu Zwecken der Seeschifffahrt. Des Weiteren hat jüngst auch die österreichische Firma Riegl ein solches Sensorsystem vorgestellt, das speziell für Flachwasserbereiche ausgelegt wurde.

Uwe Sörgel, Hannover

Standardisierung und Qualitätssicherung

Arbeitsgebiete – Terms of Reference

- Standardisierung
- Qualitätssicherung

Bericht von der Jahrestagung

Die Session des Arbeitskreises Standardisierung und Qualitätssicherung bestand aus zwei Beiträgen zu den Themenbereichen Normung und INSPIRE.

Sven Baltrusch stellte die aktuellen DIN-Normen und Normvorhaben des DIN-Normungsausschusses NA005-03-02 AA „Photogrammetrie und Fernerkundung“ vor. Der Normungsausschuss hat die Federführung für die Normungsreihen 18716 und 18740. Die neu veröffentlichte Norm zu den Anforderungen an die Klassifizierung von optischen Fernerkundungsdaten sowie die aktuell in Bearbeitung befindlichen Normen zu Anforderungen an digitale Höhenmodelle und an Pan-Sharpening Verfahren wurden vorgestellt.

Im Rahmen des INSPIRE-Prozesses werden zu Themen mit photogrammetrischem Bezug (Orthophotos, Landbedeckung, Höhendaten, usw.) Datenspezifikationen erstellt, die einen interoperablen Austausch und die Nutzung europäischer Geodaten gewährleisten sollen. Der Themenbereich der Modellierung von Höheninformationen hinsichtlich der Erdoberfläche bzw. des Meeresbodens (Bathymetrie) wurde von Michael Hovenbitzer vorgestellt. In seinem Beitrag gab er einen allgemeinen Überblick über die Ziele des INSPIRE-Prozesses und stellte das Datenmodell des Themenbereichs „Höhendaten“ vor. Weiterhin wurden geplante Veränderungen beschrieben, die sich durch die erst kürzlich abgeschlossene Kommentierungsphase ergeben haben. Da die Diskussion innerhalb des Expertenteams noch nicht abgeschlossen ist, konnten nur Zwischenergebnisse der Arbeit an den Spezifikationen dargestellt werden. Der Ausblick beschäftigte sich mit dem Zeitplan des INSPIRE-Prozesses und dessen Auswirkungen auf den Produktbereich Höheninformation.

Sven Baltrusch, Schwerin

Auswertung von Fernerkundungsdaten

Arbeitsgebiete – Terms of Reference

Innerhalb des fachlichen Spektrums der DGPF fokussiert der Arbeitskreis auf die Bearbeitung, Auswertung, Analyse und Interpretation von Fernerkundungsdaten. Methoden- und Algorithmenentwicklung unter unterschiedlichsten Anwendungsaspekten, thematischen Analysen und Interpretationen auch unter Nutzung multi-sensoraler und multi-temporaler Daten, Genauigkeitsuntersuchungen und Qualitätsanalysen von Fernerkundungsdatenauswertungen sowie die Integration von Zusatzdaten sind Schwerpunkte dieses Arbeitskreises.

Der Arbeitskreis versteht sich dabei als Diskussionsplattform für Wissenschaftler, Studierende und Anwender von Fernerkundungsdaten und soll insbesondere den Gedankenaustausch zwischen Forschung, Entwicklung und Anwendung unterstützen. Zur Mitarbeit im Arbeitskreis sind aber auch Interessenten, Fachwissenschaftler und Anwender aus allen angrenzenden geowissenschaftlichen Disziplinen und potenziellen Anwendungsbereichen aufgerufen, um eine breite Basis für die Diskussionen zu schaffen.

Im vergangen Jahr hat sich der AK schwerpunktmäßig mit neuen Methoden für die Analyse von Fernerkundungsdaten beschäftigt. Dazu wurde im Oktober an der Uni Halle ein Workshop mit einem breiten Vortragsspektrum u.a. zu Methoden wie Conditional Random Fields, Support Vector Machines und Data Mining angeboten. Neben den acht Vorträgen hatten die über 40 deutsch-österreichischen Teilnehmer ausreichend Zeit zur Diskussion. Einige der auf dem Workshop gehaltenen Vorträge sind auf der Webpage der DGPF abrufbar.

Bericht von der Jahrestagung

Auf der diesjährigen Jahrestagung der DGPF richtete der Arbeitskreis drei Sitzungen aus. Die erste Session wurde dabei speziell dem Projekt "DeCOVER2" gewidment. Das Forschungs- und Entwicklungsprojekt DeCOVER2 hat zur Aufgabe, durch die Entwicklung von Methoden der optischen und Radar-Fernerkundung Fachbehörden in ihren Verwaltungsaufgaben zu unterstützen, um aktuelle Informationen zur Landbedeckung zu liefern. Im Kern geht es dabei darum, vorhandene Datensätze zur Landbedeckung auf der Basis von Fernerkundungsdaten mit aktuellen und bei Bedarf auch ergänzenden Informationen zu aktualisieren und anzureichern. Im Fokus der Entwicklungen steht die Unterstützung der bestehenden Geobasisdatenlandschaft sowie die Erleichterung spezifischer Fachprozesse bei der Umsetzung der Flora-Fauna-Habitat-Richtlinie (FFH) zur Überwachung des europäischen Schutzgebietsnetzwerks Natura2000 sowie der Erfüllung von Monitoringaufgaben und der Fortführung von Fachdaten im Bereich der Landwirtschaft. Neben einem Übersichtsbeitrag wurden der Arbeitsstand zum FFH-Monitoring sowie ein Erfahrungsbericht zu den Synergieeffekten durch die Verwendung von RapidEye und TerraSAR-X Daten für die Veränderungsdetektion von Landnutzungsinformationen präsentiert. Die Herausforderung semantischer und geometrischer Aspekte der Interoperabilität in dem Überführungsprozess stellte der letzte Beitrag der Session vor.

In der zweiten Session trug u.a. Robert Eckhardt, der diesjährige Gewinner des Karl Kraus-Nachwuchsförderpreises, zum Thema „Cloud removal from multispectral satellite imagery using multifrequency SAR data" vor. Weitere Themen waren der Vergleich der radiometrischen und geometrischen Auflösung von RapidEye und SPOT 5 sowie ein Beitrag zum Change Detection mit hochauflösenden Satellitendaten zur Extraktion von Gebäudeschäden nach einem Erdbeben.

Eingeleitet wurde die dritte Session von einem Beitrag zur Luftbildarchäologie römischer Siedlungen im Rheinland. Der thematische Schwerpunkt der weiteren drei Beiträge lag auf dem Gebiet der forstlichen Vegetationskunde. Es wurden dazu zum einen Methoden zur Change Detection zur Erfassung von Vitalitätsveränderungen in Auenwäldern und zum anderen Methoden zur Ableitung von forstlichen Kenngrößen aus Laserscandaten und Luftbildern vorgestellt.

Weitere und geplante Aktivitäten

Auch in diesem Jahr ist im Herbst wieder ein thematischer Workshop geplant. Dieser wird erstmalig gemeinsam mit dem AK „Fernerkundung" der Deutschen Gesellschaft für Geographie an der Uni Bochum vom 4. bis 5. Oktober unter dem Motto „Fernerkundung – von der wissenschaftlichen Entwicklung zur Praxisreife" ausgerichtet. Die zweitägige Veranstaltung beginnt mit zwei praxisnahen Workshops einerseits zur objektorientierten Klassifikation und andererseits zur Change Detection Analyse. Neben regulären Vorträgen werden zwei Keynote-Präsentationen und eine interaktive Postersession angeboten. Natürlich sind wie immer alle interessierten Kolleginnen und Kollegen und Einrichtungen herzlich eingeladen sich an dem Workshop zu beteiligen. Mehr Infos finden Sie unter www.ak-fernerkundung.de.

Birgit Kleinschmit, Berlin
und Horst Weichelt, Brandenburg

3D Stadtmodelle

Arbeitsgebiete – Terms of Reference

Dreidimensionale Modelle von Städten und Regionen spielen heute eine wichtige Rolle in zentralen Aufgabenstellungen der Architektur, Stadt- und Raumplanung, Vermessung, mobilen Telekommunikation und des Facility Managements. Im Umweltbereich ermöglichen 3D-Stadtmodelle u. a. die Simulation von Lärm- und Abgasausbreitungen sowie Voraussagen über mögliche Veränderungen des Stadtklimas. In Katastrophensituationen wie z. B. bei Hochwasser kann bei Vorliegen von 3D-Landschaftsmodellen schnell ermittelt werden, welche Gebiete und Gebäude betroffen sein werden, so dass entsprechende Maßnahmen frühzeitig eingeleitet werden können.

Die DGfK und die DGPF haben am 6. März 2009 eine gemeinsame Kommission bzw. einen gemeinsamen Arbeitskreis „3D-Stadtmodelle“ gegründet, um diesem wichtigen Thema Raum zu geben und gleichzeitig die Zusammenarbeit zwischen beiden Gesellschaften zu intensivieren.

Bericht von der Jahrestagung

Der gemeinsame AK von DGfK und DGPF war bei der Jahrestagung mit einer Session vertreten, in der vier Themen mit dem Schwerpunkt 3D-Stadtmodelle vorgestellt wurden.

Dabei ging es um den Einsatz von GML-Datenbanken in der 3D-Simulation (Fa. CPA) oder um die immersive Visualisierung von 3D-Stadtmodellen für die Stadtplanung (Hasso-Plattner-Institut). Die TU Berlin stellte ein Integratives Entscheidungswerkzeug als Planungshilfe am Beispiel des Energieatlasses vor, und von der Hochschule Bochum kam ein Vorschlag für eine alternative Gebäudemodellierung, die insbesondere für ein 3D-Informationssystem in der Bauphysik Verwendung finden soll.

Auch der AK Geoinformation hat das Thema in seinen Sessions mit mehreren Vorträgen gestreift, so dass die Teilnehmer unter einer ganzen Reihe von Vorträgen zum Thema 3D-Stadtmodelle auswählen konnten.

Weitere und geplante Aktivitäten

Anfang Mai trafen sich die Mitglieder des AK in der HTW Dresden zu einer weiteren Sitzung. Ein intensiver Erfahrungsaustausch im Zuge eines AK-Treffens an der HTW Dresden über Software-Anwendungen und die Laufendhaltung der bestehenden 3D-Stadtmodelle brachte insbesondere bei mehreren Anwendern bestehende Probleme zum Vorschein, die durch die schnelle Versionierung der Software hervorgerufen wird. Ein Ziel des AK ist es daher – wie auch die Workshop-Teilnehmer diskutierten – eingehend Möglichkeiten zu untersuchen, wie solchen Herausforderungen entgegengewirkt bzw. entgangen werden kann.

Der AK plant für die kommenden Jahre eine Unterstützung von Studenten, die sich dem Thema 3D-Stadtmodelle widmen. So soll erstmals einer kleinen Anzahl von Studenten die Teilnahme an der 26. „International Cartographic Conference“ im August 2013 in Dresden ermöglicht werden. Die Auswahlkriterien wurden eingehend diskutiert; sie werden in Kürze an Hochschulen, Verbände und wichtige Gremien verteilt. Inzwischen liegt auch eine Vorab-Version eines umfangreichen Fragebogens vor, mit dessen Hilfe der AK 3D-Stadtmodelle, die TU Berlin und der AK Geoinformation des Deutschen Städtetags einen Überblick über den Aufbau und die Nutzung von 3D-Stadtmodellen in den Kommunen Deutschlands erhalten wollen. Letzte Arbeiten wurden diskutiert, insbesondere um eine möglichst breite Akzeptanz zu erzielen. Der Auswertung sehen die Beteiligten mit Spannung entgegen.

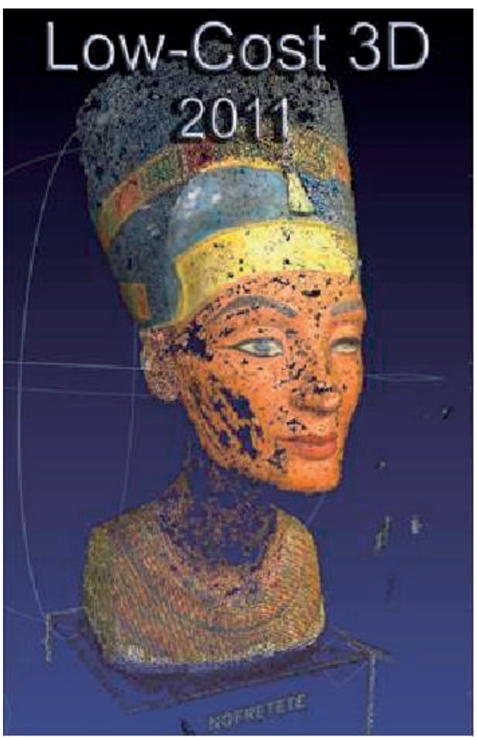
Mit Hochdruck wird derzeit am Programm für den Workshop 3D-Stadtmodelle im November in Bonn gearbeitet, den der AK inzwischen zur einer festen Größe im Terminkalender etablieren konnte. Am 6. und 7. November 2012 wird im Universitäts-Club Bonn wieder eine Reihe von interessanten Vorträgen rund um das Thema 3D-Stadtmodelle zu hören sein. Die Organisatoren gehen bei der Schwerpunktsetzung auf die Themen ein, die in den Fragebögen des letztjährigen Workshops als Wunschthemen genannt wurden. Weitere Hinweise auf den Workshop und auf das detaillierte Programm erscheinen auf der Homepage des AK: www.3D-stadtmodelle.org

Bettina Petzold, Wiesbaden

Berichte von Veranstaltungen

Workshop „Low-Cost 3D“, 6.–7. Dezember 2011 in Berlin

Das Institut für Geodäsie und Geoinformationstechnik der Technischen Universität Berlin und das Institut für Informatik (Fachbereich Computer Vision) der Humboldt-Universität (HU) zu Berlin veranstalteten am 6. und 7. Dezember 2011 den Workshop *Low-Cost 3D – Sensoren, Algorithmen, Anwendungen* an der TU Berlin. Der Workshop bestand aus Live-Demonstrationen unterschiedlicher Systeme, die am Nachmittag und Abend des 6. Dezember stattfanden und einer ganztägigen Vortragsveranstaltung am darauf folgenden Tag.



Nofretete als Symbol und Aufnahmeobjekt der Veranstaltung Low-Cost 3D.

Die Organisation des Workshops lag in den Händen von Prof. Dr.-Ing. Frank Neitzel (TU Berlin), Prof. Dr. rer. nat. Ralf Reulke (HU Berlin) und den jeweiligen Mitarbeitern. Mitveranstalter des Workshops waren das DLR Berlin-Adlershof, das Fraunhofer Heinrich-Hertz-Institut und die Internationale Gesellschaft für Photogrammetrie und Fernerkundung, vertreten durch die Arbeitsgruppe (WG V/5) *Image Sensor Technology*. Der Workshop Low-Cost 3D war die erste Veranstaltung im deutschsprachigen Raum, die sich ausschließlich der speziellen Thematik *Low-Cost 3D-Rekonstruktionsverfahren* in all ihren Ausprägungen widmete. Die Veranstaltung richtete sich speziell an Forscher, Entwickler und Anwender mit Interesse an Low-Cost 3D-Rekonstruktionsverfahren und deren Anwendungsmöglichkeiten und -potential. Trotz der kurzfristigen Ankündigung waren 80 Teilnehmer aus dem In- und Ausland zur Veranstaltung gekommen, was auch das steigende Interesse an diesen neuen Verfahren widerspiegelt. Mit 67% stellten die Hochschulen und Forschungseinrichtungen erwartungsgemäß den größten Teilnehmerkreis, während mit 28% bereits fast ein Drittel der Interessierten aus der freien Wirtschaft und mit 5% ein kleiner Teil aus Behörden bzw. von Verwaltungen (Landesforst, Archäologie) kamen.

Dass 3D-Rekonstruktion nicht teuer sein muss, zeigten Live-Demonstrationen verschiedener Systeme und 13 Referenten in ihren Präsentationen. Dass man Live-Demos von verschiedenen Systemen und die Icebreaker-Party zur gleichen Zeit unter einem Dach veranstalten kann, wurde bereits am Abend vor den Präsentationen auf dem Geodätenstand der TU Berlin gezeigt. Bei gutem Blick über die Stadt in der Abenddämmerung demonstrierten die TU Berlin (Objektrekonstruktion mit den Programmen „Bundler“ und „PMVS2“), die Ruhr-Universität Bochum (Low-Cost Navigationssystem mit einer PMD-Kamera), die TU Clausthal-Zellerfeld (Höhlenvisualisierung), die Firma fiagon GmbH (Navigationssystem für die Zahnheilkunde) und die Firma gfai tech GmbH (3D Laserscanner FinalScan LR-50) ihre Systeme, während sich die Teilnehmer gleichzeitig kulinarisch erfreuen konnten.

Am darauf folgenden Tag wurden in fünf Sitzungen und in 13 Vorträgen Aufnahmesysteme, Algorithmen, geometrische Untersuchungen und verschiedene Anwendungen auf Englisch präsentiert. Nach einer kurzen Begrüßung durch die beiden Organisatoren stellte Frank Neitzel ein Low-Cost UAV System (Optokopter mit Canon IXUS 100 IS) für das Mobile 3D Mapping eines Parkplatzes (zur geometrischen Genauigkeitsuntersuchung) und einer Deponie (zur Bestandsaufnahme) vor, bei der überwiegend frei verfügbare (kostenlose) Software für die Datenverarbeitung eingesetzt wurde. Anschießend präsentierte Konrad Wenzel (Universität Stuttgart) ein Multi-Kamera-System, bestehend aus zwei Kameras im sichtbaren Licht und zwei im nahen Infrarot mit einem Kinect-Sensor als Muster-Projektor, für die Datenerfassung im Nahbereich. Damit wurde in Amsterdam in zehn Tagen von 2200 Standpunkten ein Relief aufgenommen, dessen Oberfläche nach dem Dense Image Matching durch 10 Punkte pro mm^2 mit sehr hoher Auflösung beschrieben wurde. Etwas exotischer wurde es dann im dritten Vortrag, in dem Yuan Xu (Humboldt-Universität zu Berlin) den Kinect-Sensor als Low-Cost Bewegungsaufnahmesystem zur Simulationssteuerung eines Fußball spielenden Roboters als Vorbereitung auf den internationalen Robo-Cup vorstellte. Für die Validierung des Systems wurden allerdings RGB-Kameras verwendet.

In der zweiten Sitzung wurden Untersuchungen von Low-Cost Sensoren gezeigt. Jan Böhm (University College London) stellte seine Genauigkeitsuntersuchungen von Natural User Interface Sensors (Kinect, PrimeSense, Asus) vor, bei denen es sich zeigte, dass nicht alle Sensoren aus der gleichen Serie die geometrischen Anforderungen auf sehr kurze Entfernungen für Innenraumanwendungen, z.B. Bekleidungsindustrie, erfüllten. Fabio Remondino (FBK Trento, Italien) präsentierte nicht nur geometrische Genauigkeitsuntersuchungen verschiedener Low-Cost Sensoren, sondern zeigte auch Ergebnisse verschiedener Auswertesoftware (Photosynth, Agisoft, Apero/MicMac) bei der automatischen Bildorientierung und der bildbasierten Erstellung von Punktwolken. Während der Kinect-Sensor noch Systematiken in der Punktwolke aufweist, zeigt die Stereo-Kamera Fujifilm Real 3D W1 mit einer Stereoaufnahme bessere Resultate als der Kinect-Sensor, wenn zuvor eine Sensorkalibrierung erfolgt. Beim Vergleich von drei Softwarepaketen zeigte Apero (automatische Bildtriangulation) und MicMac (Dense Image Matching) bei Standardkörpern (Würfel und Kugel) die besten Resultate.

In der dritten Sitzung demonstrierte Sebastian Vetter (Fokus GmbH, Leipzig) in seinem Vortrag und in einer Live-Demo die Vielseitigkeit der Software metigo 3D bei diversen Aufgaben in der Objektdokumentation aus Bildverbänden. Anschließend stellte Günther Pomaska (FH Bielefeld) das Vorgehen mit der im Internet verfügbaren Microsoft Photosynth (Grundlage ist die Bundler-Software) in Zusammenarbeit mit Google SketchUp 3D für die Bauwerksmodellierung vor. Im dritten Vortrag dieser Sitzung präsentierte Nesrine Grati (i3mainz) eine Strategie, um die 3D-Rekonstruktion von urbanen Szenen in Punktwolken mit 2D-Bilddaten zu verbessern.

In der vierten Sitzung stellte Mohammed Abdel-Wahab (Universität Stuttgart) seinen optimierten Arbeitsablauf und Datenfluss vor, mit dem die Daten aus dem Amsterdam-Projekt durch bildpaarweises Dense Matching prozessiert wurden, deren Aufnahme Konrad Wenzel am Morgen präsentierte. Im Anschluss entführte Wilhelm Hannemann (TU Clausthal-Zellerfeld) die Zuhörer virtuell unter Tage, in dem er die Zustandsdokumentation ausgedehnter untertägiger Hohlraumsysteme mit Hilfe der Programme Bundler und PMVS2 an verschiedenen Beispielen darstellte, die für Anwender interaktiv mit WebGL präsentiert werden können. Thematisch bildete der Vortrag von Thomas Kersten (HCU

Hamburg) den Abschluss der Veranstaltung. Er zeigte das Potential automatischer 3D-Rekonstruktionsverfahren anhand verschiedener Softwarepakete (Bundler/PMVS2, Photofly/123D Catch und Microsoft Photosynth) und unterschiedlicher Daten (Architektur, Denkmalpflege und Archäologie) auf und verglich die Ergebnisse mit Daten vom terrestrischen Laserscanning. Dabei stellte sich allerdings die Frage, welcher Datensatz überhaupt genauer ist und als Referenz gegenüber dem anderen dienen kann. Die neuen photogrammetrischen Verfahren stellen heute aufgrund der steigenden Rechnerleistungen (z.B. sehr schnelles Rechnen auf der GPU – Graphics Processing Unit) eine sehr kostengünstige Alternative zu terrestrischen Laserscannern dar. „Photogrammetry Reloaded!“ könnte das neue Motto heißen.

Den Abschluss des Workshops bildete die fünfte Sitzung mit zwei Firmenvorträgen, die das Thema Low-Cost 3D aus einer etwas anderen Sichtweise betrachteten. Dirk Kowalewski (navXperience GmbH, Berlin) stellte die Frage „What is the real accuracy of a GNSS antenna?" und beantwortete diese Frage durch einen Vergleich der Antenne navXperience 3G+C mit den Produkten von Mitbewerbern am Markt. Andreas Rose (fiagon GmbH, Berlin) präsentierte ein optisches Navigationssystem für die Zahnheilkunde, wobei ausgehend von den klinischen Grundlagen die Anforderungen an das System aufgezeigt wurden.

Mit dem Workshop *Low-Cost 3D – Sensoren, Algorithmen, Anwendungen* zeigten die beiden Kollegen aus Berlin, Frank Neitzel und Ralf Reulke, ein sehr gutes Näschen für ein hochaktuelles Thema in der 3D-Vermessung. Als Fazit kann der Teilnehmer mitnehmen, dass das Potential aktiver und passiver bildbasierter Aufnahmesysteme für sehr viele Anwendungen enorm ist. Dabei steht man erst am Anfang einer Entwicklung, die sich mittlerweile auch im kommerziellen Markt durchzusetzen beginnt. Wir dürfen gespannt sein, was im zweiten Workshop in Berlin-Adlershof Ende 2012 präsentiert wird. Dem TU- und HU-Team um Frank Neitzel und Ralf Reulke gilt besonderer Dank für die hervorragende Organisation und Durchführung dieses Workshops. Die Vorträge werden als PDF-Dateien unter www.lc3d.net im Internet zur Verfügung gestellt.

Thomas Kersten, Hamburg

EuroCOW 2012: EuroSDR / ISPRS European Calibration and Orientation Workshop, February 8–10, 2012, Castelldefels, Spain

The 4th EuroCOW workshop, biennially co-organized by EuroSDR and ISPRS took place February 8–10, 2012 at the Institute of Geomatics (IG) in Castelldefels, Spain, 20 km south of downtown Barcelona. More than 70 participants from 18 countries registered for this event, 34 authors presented their latest findings in oral presentations with about 30 minutes each.

On the day before the official workshop opening a one day pre-workshop tutorial on camera orientation based on IGN France's open source photogrammetric software (APERO) was offered by the IGN.

The workshop itself was opened by the conference keynote given by Craig Glennie, University of Houston, who presented the open challenges for mobile mapping systems. He focussed on the technical challenges including the better trajectory estimation, more rigorous calibration and integration of lower cost sensors. From a commercial adoption perspective, software to analysis, process or extract from a dataset is clearly lagging far behind the hardware to collect it.

Naser el-Sheimy, University of Calgary and currently president of the ISPRS Commission 1, was the second keynote speaker. He gave an introduction to the integrated navigation topic of the workshop. His keynote was entitled: MEMS (Microelectromechanical system) Sensors – State of the art and future trends in mapping and navigation applications. He summarizes the talk with his main conclusion: With the technology push and the market pull, MEMS inertial systems is going to reach the performance of tactical grade IMU soon.

The scientific program was structured in eight different sessions: three on calibration and orientation in general, two on special calibration and orientation for UAV, as well as one each on GNSS, INS and radiometry.

Announcement of the next EuroCOW in 2014, engraved at a champaign glass.

The proceedings (and presentations) can be accessed via the IdeG web site: http://www.ideg.es/page.php?id=1360 . Selected papers have already been made available in PFG's issue 2/2012.

The technical program was continued with a remarkable evening social program, where the "1st Wine Orientation and Calibration Seminar" happened in the historical castle of Castelldefels which should be explicitly mentioned. Thus, an almost 24hour / 3days program, was giving lots of opportunities, not only for the discussion of the technical issues but also for deepening social contacts and meeting old and finding new friends!

As in the years before, the EuroCOW 2012 was a very valuable workshop, because this event has preserved its "workshop character" and did not develop towards a too big conference. This allowed manifold personal communication like meeting the presenters in the coffee breaks to discuss their presentations. The program was structured as a single track only. Thus, the meeting had its clear scientific focus on orientation and calibration issues, but also addressed "side-disciplines".

In the meantime EuroCOW is very well established and accepted and has its own history. It has taken place already four times (2006, 2008, 2010, and 2012) and originated as a follow-up of the ISPRS workshops on "Integrated Sensor Orientation" held 1995, 1999 and 2003 at the ICC (Institut Cartogràfic de Catalunya) and IG. To conclude, the workshop was perfectly organized and had a very good "input/output-ratio" mainly due to Ismael Colomina and his team at the Institute of Geomatics!

For sure, EuroCOW will be continued. The 5th edition is already scheduled for February 12–14, 2014, again in Castelldefels.

Michael Cramer, Stuttgart

4. RESA-Workshop, 21.–22. März in Neustrelitz

Am 21. und 22. März 2012 fand in Neustrelitz am DLR–DFD der vierte RESA Nutzer Workshop statt. RESA steht für das RapidEye Science Archive. In diesem Jahr stand die Tagung unter dem Leitthema „*Daten für die Wissenschaft - Vom Algorithmus zum Produkt*".

Nach der Eröffnung durch Dr. Erik Borg, Leiter und Organisator der Veranstaltung, wurden 13 Fachvorträge und 12 extended abstracts in verschiedenen thematischen Blöcken vorgestellt.

Zu Beginn der Vortragsreihe informierte Michael Oxfort, Chief Technology Officer bei RapidEye AG, über technische und organisatorische Entwicklungen bei der Firma RapidEye. Er bekräftigte das Interesse des Unternehmens, weiterhin eng mit dem DLR und insbesondere mit dem RapidEye Science Archive zusammenzuarbeiten. Jedoch sprach er auch über die ökonomischen Perspektiven, die ein kommerziell arbeitendes Unternehmen immer stark im Fokus haben muss. In Zukunft soll die Anzahl komplexer Anfragen für kleine Flächen vermieden werden. Damit soll die Systemeffizienz erhöht und die Standzeit verlängert werden.

Die ersten Fachvorträge behandelten das Thema „*Geologie und Hydrologie*" Sehr interessant war der Beitrag von Barbara Theilen-Willige von der TU Berlin über ein Projekt zum Katastrophenschutzmanagement, welches auch den Aufbau adäquater Referenzdatenbanken zum Ziel hatte. Bemerkenswert ist vor allem, dass dank der guten Arbeitsabläufe bei RapidEye bereits nach dreieinhalb Stunden erste FE-Daten zur Auswertung eines simulierten Erdbebens bereitstanden. Dabei wurden auch temporäre Feuchtgebiete (nach Regenfall) detektiert, welche als Fallout-Risikogebiete für atomare Katastrophen ausgewiesen werden können. Die folgenden Referate zu „*Vegetation und Landschaftsräumen*" erörterten Potenziale der RapidEye-Bilddaten beim Vegetationsmonitoring, bei der operationellen Forstinventur und dem Erstellen von Ertragskarten für landwirtschaftliche Anbauflächen. Vielfach wurden dabei multisensorale FE-Daten in die Prozessierung integriert. Im Rahmen des Forschungsprojektes ENVILAND 2 wurden beispielsweise optische RapidEye und SAR-Daten des TerraSAR-X Satelliten synergetisch genutzt, um mit einem eigens entwickelten Workflow weitestgehend automatisch Informationen zur Landbedeckung abzuleiten. Tobias Benedikt Hank von der LMU München zeigte in seinem Vortrag, dass es mit multisensoralen und multitemporalen Fernerkundungsdaten möglich ist, die Variation von Ernteerträgen mit verlässlicher Genauigkeit zu modellieren. Nach zwei weiteren Vorträgen zum Themenblock „*Ozeanographie und Kryosphäre*" folgten die kürzeren Referate (extended abstracts). Besonders hervorzuheben ist der Vortrag von Herrn Illiger vom Helmholtz-Zentrum für Umweltforschung Halle, der in seinem Vortrag die deutlichen Vorteile der hochauflösenden RapidEye Daten bei der hydrologischen Modellierung gegenüber den in zunehmendem Umfang genutzten Daten mit reduzierter räumlicher Auflösung in einem Vergleich mit MODIS Daten hervorhob.

Das pünktliche Einreichen aller Workshop-Beiträge ermöglichte den rechtzeitigen Druck des Tagungsbandes vorab, wofür an Sören Hese von der Universität Jena ein Preis verliehen wurde. Für die perfekte Organisation und nicht zuletzt für das leibliche Wohl muss an dieser Stelle noch einmal ausdrücklich Dank an die Veranstalter am DLR-DFD Standort Neustrelitz ausgesprochen werden. Namentlich geht der Dank an Erik Borg und Holger Daedelow, verbunden mit dem Wunsch, diese erfolgreiche Veranstaltungsreihe in 2013 fortzuführen.

Stefan Jentsch,
Robert-Walter Wildgrube,
Matthias Möller, Berlin

Message from DGPF

Data of the DGPF Project on the Evaluation of Digital Airborne Cameras made available for Research Purposes

In 2008, the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) initiated a test project to evaluate the geometrical and radiometric potential of commercial digital airborne camera systems. For that purpose, digital imagery from various camera systems at different image scales and multiple-pulse airborne laserscanner (ALS) data were acquired in 2008 in the test field Vaihingen / Enz near Stuttgart in the federal state of Baden-Württemberg, Germany. This test field is maintained by the Institute of Photogrammetry (ifp) at the University of Stuttgart. The data were distributed to different research groups; use of the data was restricted to the DGPF test itself.

In the meantime, the project has been finished successfully. Its outcomes were reported in issue 2/2010 of PFG as well as at the DGPF meetings in Jena (spring 2009) and Vienna, Austria, in summer 2010. Nevertheless, the data are still of interest to the international research community. They provide coverage by modern ALS and imaging sensors, in the latter case including multiple-view scenarios and multiple geometrical resolutions, so that they are potentially useful for research beyond the immediate goals of the DGPF test. In particular, they can be extremely useful to provide a test scenario for automated object detection techniques. In this context, the DGPF gave permission to Working Group III/4 of the International Society of Photogrammetry and Remote Sensing (ISPRS), "*Complex Scene Analysis and 3D Reconstruction*" to distribute a subset of the Vaihingen / Enz data for an international test on "*Urban Classification and 3D Building Reconstruction*" (itc.nl/isprs_wgiii4/tests_datasets.html). These test data also include ground truth for various object detection tasks and for 3D building reconstruction. The use of the data was restricted to the purposes of the ISPRS test. A summary of the outcomes is presented at the ISPRS Congress in Melbourne.

The high quality of the data and the demand of the international research community for benchmark datasets have convinced the DGPF that the Vaihingen / Enz data should be made available for general research purposes, thus to lift some of the restrictions for the use of the data. After the ISPRS Congress in Melbourne, the data can be accessed via the following URL:

dgpf.de/neu/projekte.htm

The Vaihingen / Enz dataset can be accessed via a link from that URL. In addition, the subset used in the ISPRS test can still be accessed, either via a link from that URL or directly via itc.nl/isprs_wgiii4/tests_datasets.html. The web site for the ISPRS test will remain accessible and will be complemented by an access to the test results and to an online evaluation system where new results can be uploaded and evaluated on-line using the reference data.

In both cases, a user registration is required. After registration, the user will receive an e-mail containing the login information required to download the data. The data may be used subject to the following conditions:

1. The data must not be used for other than research purposes. Any other use is prohibited.
2. The data must not be distributed to third parties. Any person or institution interested in the data must register at one of the sites mentioned above.
3. Any scientific paper whose results are based on the Vaihingen/Enz test data must cite (Cramer 2010) and must contain the following acknowledgement: "The Vaihingen / Enz dataset was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) (Cramer 2010): www.ifp.uni-stuttgart.de/dgpf/dkep-allg.html."
4. The Executive Board of the DGPF should be informed about any published papers whose results are based on the Vaihingen /

Enz test by an e-mail to the secretary of the society; currently Eberhard Gülch (eberhard.guelch@hft-stuttgart.de).

We hope that the dataset will provide a test bed for many applications in Photogrammetry, Remote Sensing and Computer Vision. Furthermore, we expect the subset used in the ISPRS test to become a standard benchmark dataset for urban object extraction, which on the one hand will make different techniques more comparable, but on the other hand also trigger new research by identifying common problems of existing techniques or promising strategies to be followed in the future.

References

Cramer, M., 2010: The DGPF Test on Digital Aerial Camera Evaluation – Overview and Test Design. – PFG **2010** (2): 73–82.

Franz Rottensteiner, Hannover

Hochschulnachrichten

Technische Universität München

Frau M.Sc. Xiaoxiang Zhu wurde am 26.5.2011 an der Fakultät für Bauingenieur- und Vermessungswesen der Technischen Universität München mit der Arbeit „*Very High Resolution Tomographic SAR Inversion for Urban Infrastructure Monitoring – A Sparse and Nonlinear Tour*“ zur Dr.-Ing. promoviert.

Referent: Prof. Dr.-Ing. habil. Richard Bamler, Technische Universität München, Koreferenten: Prof. Dr.-Ing. Uwe Sörgel, Leibniz Universität Hannover, Prof. Dr.-Ing. habil. Alberto Moreira, Karlsruher Institut für Technologie.

Kurzfassung: Synthetik Apertur Radar (SAR) ist ein kohärentes Abbildungsverfahren und arbeitet mit Wellenlängen typischerweise im Bereich 3 – 25 cm. Daher ist es das einzige Weltraum gestützte Verfahren, das es erlaubt, Deformationen der Erdoberfläche mit Zentimeter- und sogar Millimeter-Genauigkeit zu erfassen. Mit dem Start neuer SAR-Systeme wie den deutschen TerraSAR-X und TanDEM-X sowie den italienischen COSMO-Skymed-Satelliten hat die SAR-Fernerkundung einen wichtigen Schritt in die Zukunft getan. Diese Satelliten liefern SAR-Daten mit sehr hoher räumlicher Auflösung (VHR für englisch „very high resolution“) von bis zu 1 m. Sie eröffnen deshalb erstmals die Möglichkeit, SAR zur 2D, 3D und 4D (raum-zeitlichen) oder sogar höherdimensionalen Abbildung urbaner Infrastruktur und einzelner Gebäude vom Weltraum aus einzusetzen, d.h., die 3D Form von Gebäuden sowie deren Deformationen oder Absenkungen abzuleiten.

Eine einzelne SAR-Aufnahme liefert nur kartographische Information in den beiden nativen Koordinaten „Azimut“ und „Range“. Um die tatsächliche 3D Lokalisation, also einschließlich der dritten Koordinate „Elevation“, und die Bewegungsparameter der streuenden Objekte zu schätzen, werden moderne interferometrische SAR-Verfahren benötigt. Diese nutzen Stapel komplexwertiger SAR-Daten, aufgenommen von unterschiedlichen Orbitpositionen und zu unterschiedlichen Zeiten. Tomographische SAR-Inversion, also SAR-Tomographie und differentielle SAR-Tomographie, stellt das derzeit fortschrittlichste 4D SAR-Abbildungskonzept dar. Diese relativ neue Methode wird allerdings noch kaum mit den nun verfügbaren VHR-Daten zur Abbildung urbaner Gebiete genutzt. Diese Dissertation hat zum Ziel, die tomographische Technik weiter zu entwickeln, um das Potential der VHR SAR-Systeme zur raum-zeitlichen Abbildung urbaner Infrastruktur auszuschöpfen. Die Arbeit trägt dazu vier neue Aspekte und Lösungen bei:

VHR tomographische SAR-Inversion wird mit TerraSAR-X Spotlight-Daten demonstriert. Es wird gezeigt, dass eine 3D und 4D Karte eines gesamten Innenstadtgebiets mit Hochhäusern abgeleitet werden kann, einschließlich der Auflösung von Layover-Bereichen, d.h. der Trennung von mehrfachen Streuern in einem Bildpunkt. Für einzelne Gebäude werden mit einer modifizierten Version des bekannten Singulärwert-Zerlegungs-Algorithmus und nachgeschalteter Modelselektion hohe Zweifachstreuer-Dichten – bis zu 20% – detektiert.

Wegen des sehr engen Orbitschlauchs moderner SAR-Satelliten ist die Elevationsapertur klein und damit die Auflösung in Elevationsrichtung ca. 50 mal schlechter als in Azimut und Range. Diese starke Anisotropie des 3D Auflösungselements verlangt nach Überauflösungsverfahren in Elevation, ohne die Azimut- und Range-Auflösung zu verschlechtern. Andererseits garantiert diese anisotrope Punktantwort auch, dass das Signal als „sparse“ in Elevation angesehen werden kann. Somit werden nur wenige diskrete Streubeiträge pro Azimut-Range-Pixel erwartet. Dies legt die Nutzung von Compressive Sensing Verfahren nahe. Zur tomographischen Inversion von VHR SAR-Daten wurde ein solcher Compressive Sensing basierter Algorithmus entwickelt mit dem Namen „Scale-down by L_1 norm Minimization, Model selection, and Estimation Reconstruction“ (SL1MMER, ausgesprochen wie „slimmer“). SL1MMER kombiniert die Vorteile der Compressive Sensing Rekonstruktion, also Überauflösungsfähigkeit und genaue Positionsbestimmung, mit der Amplituden- und Phasengenauigkeit linearer Schätzer. Der Algorithmus gibt daher verlässliche Schätzungen der Anzahl der Streuer sowie deren Elevationspositionen, Amplituden und Phasen. Die Überauflösungsfähigkeit von SL1MMER wird mit realen TerraSAR-X-Daten demonstriert und zeigt eine beachtliche Steigerung der Anzahl der Zweifachstreuer auf bis zu 38%.

Eine systematische Analyse der Leistungsfähigkeit des vorgeschlagenen SL1MMER-Algorithmus zeigt dessen Elevations-Schätzgenauigkeit, Überauflösungsfähigkeit und Robustheit auf. Sowohl numerische wie auch analytische Näherungen der Elevations-Genauigkeit werden mit der Cramér-Rao-Grenze verglichen. Es wird gezeigt, dass SL1MMER ein effizienter Schätzer ist. Die erreichbare Überauflösung wird durch extensive Simulationen bestimmt. Diese Überauflösungsfaktoren stellen zugleich absolute und grundsätzliche Grenzen für Spektralschätzer dar. Für den typischen Parameterbereich in der SAR-Tomographie werden Überauflösungsfaktoren von 1.5–25 erreicht. Schließlich wird die Mindestanzahl von Datensätzen zur robusten topographischen Rekonstruktion hergeleitet.

Bisherige tomographische Inversionsverfahren berücksichtigen nur lineare Bewegung, obwohl Deformationen von Gebäuden oder der Erdoberfläche meist nichtlinear sind, z.B. periodisch, beschleunigt, oder abrupt. Daher wurde eine generalisierte „Time Warp“ Methode entwickelt, mit der die Aufnahmezeiten auf beliebige Basisfunktionen abgebildet werden, so dass nichtlineare multi-modale Bewegungen geschätzt werden können.

Alle hier entwickelten Verfahren und Algorithmen wurden sowohl in Simulationen wie auch durch Verarbeitung großer realer TerraSAR-X-Datenstapel validiert.

Die mit dieser Dissertation präsentierte Arbeit möge ein substantieller Beitrag sein zu der Vision in Bezug auf ein dynamisches Stadtmodell, das Form und Deformation jedes Gebäudes zeigt.

Die Dissertation ist in der Kollektion „Elektronische Prüfungsarbeiten“ der Fakultät für Bauingenieur- und Vermessungswesen als Online-Veröffentlichung erschienen (Urn: http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20110526-1071713-1-0). Gleichzeitig ist die Arbeit in der Reihe C der Deutschen Geodätischen Kommission bei der Bayerischen Akademie der Wissenschaften (ISBN 978-3-7696-5078-5) unter der Nr. 666 online veröffentlicht (www.dgk.badw.de).

Humboldt-Universität zu Berlin

Herr M.Sc. Karsten Kozempel wurde am 10.2.2012 an der Mathematisch-Naturwissenschaftlichen Fakultät II der Humboldt-Universität zu Berlin mit der Arbeit „*Entwicklung und Validierung eines Gesamtsystems zur Verkehrserfassung basierend auf Luftbildsequenzen*“ zum Dr. rer. nat. promoviert.

Referent: Prof. Dr. rer. nat. Ralf Reulke, Humboldt-Universität zu Berlin, Koreferenten: Prof. Dr.-Ing. Stefan Hinz, Karlsruher

Institut für Technologie, Prof. Dr. sc. nat. i.R. Hans-Dieter Burkhard, Humboldt-Universität zu Berlin.

Kurzfassung: Die Dissertation leistet einen Beitrag zur Weiterentwicklung der luftgestützten Verkehrslageerfassung. Als Plattform dafür dient ein flugzeuggetragenes Kamerasystem, welches mit einem Fluglageerfassungssystem gekoppelt ist. Vorgestellt werden hauptsächlich bildverarbeitende Algorithmen, die beginnend von der Bildvorverarbeitung bis hin zur Ermittlung der wichtigsten verkehrstechnischen Kenngrößen zum Einsatz kommen.

Die Arbeit behandelt im Wesentlichen vier Aspekte:

1. Die Kalibrierung der Offsetwinkel zwischen Kamera und Inertialsystem
2. Einen mehrstufigen Ansatz zur schnellen und präzisen Fahrzeugerkennung
3. Die Bestimmung der Durchschnittsgeschwindigkeiten durch Assoziation in Folgebildern
4. Die Entwicklung eines optischen Orientierungsalgorithmus zur Kompensation des Inertialsystems

Nach kurzer Skizzierung der verwendeten Hardware wird zunächst die Kalibrierung der Kameraeinbauwinkel durch Testflüge erläutert und auf ihre Genauigkeit hin untersucht. Es wird gezeigt, dass die Orientierungsdaten zwar nur moderate Genauigkeiten erreichen, diese für die Verkehrslageerfassung jedoch ausreichend sind.

Anschließend an Algorithmen für die Bildaufbereitung wie Orthobildgenerierung und Ableitung von Suchregionen für die Eingrenzung der verkehrsaktiven Flächen wird zur Ermittlung der Fahrzeugdichte ein zweistufiger Fahrzeugerkennungsalgorithmus entwickelt. Aufbauend auf einer Bank an Kantenfiltern werden sehr effizient Hypothesen für Fahrzeuge erstellt. Diese werden in einer zweiten Phase durch eine Support Vector Machine überprüft, wobei ein Großteil der Fehlhypothesen verworfen wird. Die Erkennung erreicht bei guten Voraussetzungen Vollständigkeiten bis zu 90 Prozent bei einem sehr geringen Anteil von Fehldetektionen.

Anschließend wird ein auf Singulärwertzerlegung basierender Tracking-Algorithmus verwendet, um Fahrzeughypothesen in benachbarten Bildern zu assoziieren und die mittleren Geschwindigkeiten zu ermitteln. Die erhaltenen Geschwindigkeiten unterscheiden sich um weniger als 10 km/h von den manuell erhobenen.

Abschließend wird eine alternative Orientierungsmethode vorgestellt, welche auf Basis von GPS-Positionen und Bildinformationen automatisch die Fluglage ermittelt. Dies geschieht durch die Extraktion und das Matching von Straßensegmenten sowie zusätzliche Passpunktverfolgung. Die Ergebnisse weisen Genauigkeiten von etwa 0,1 bis 0,2 Grad auf.

Der Anspruch der Arbeit besteht bei der Fahrzeugerkennung in der Kombination aus schneller Vorselektion und zuverlässiger Verifizierung der Objekthypothesen. Weiterhin sind die automatische Extraktion von Straßen als Landmarken zur Orientierungsbestimmung und die Fusion mit relativen Orientierungsparametern zur Stabilisierung neuartig.

Die Dissertation wurde am 22.3.2012 auf dem elektronischen Dokumentenserver der Humboldt-Universität zu Berlin im Fachgebiet Informatik veröffentlicht (http://edoc.hu-berlin.de/).

Technische Universität Dresden

Herr Khalid Guma Biro Turk, M.Sc., promovierte am 5.12.2011 am Institut für Kartographie der Technischen Universität Dresden mit der Arbeit *„Development of a Multi-Temporal Satellite Data for Landuse/Landcover Change Analysis and its Impacts on Soil Properties in Gadarif Region, Sudan"* zum Dr. rer. nat.

1. Gutachter: Prof. Dr. Manfred Buchroithner, TU Dresden, Institut für Kartographie
2. Gutachter: Prof. Dr. Volker Hochschild, Universität Tübingen, Fachbereich Geowissenschaften

Kurzfassung: Mehrere Jahrzehnte intensiven Trockenfeldbaus in der Region von Gadarif, welche sich im östlichen Teil des Sudans befindet, führten hauptsächlich aufgrund von landwirtschaftlicher Expansion, politischen Beschlüssen der Regierung und Naturkatastrophen wie Trockenheit zu einer raschen Ver-

änderung der Landnutzung und Landbedeckung. Das wesentliche Ziel dieser Dissertation war es, die Degradation des Landes sowie die Auswirkungen von landwirtschaftlicher Expansion auf die Landbedeckung, den Boden und den Pflanzenbau im Untersuchungsgebiet, welches Teile der afrikanischen Sahelzone beinhaltet, abzuschätzen.

Zur Analyse und Beobachtung der Veränderungen der Landnutzung und Landbedeckung wurden multi-temporale Landsat-Daten der Jahre 1979, 1989 und 1999 sowie ASTER-Daten aus dem Jahr 2009 genutzt, welche eine Fläche von schätzungsweise 1200 km^2 abdecken. Um Veränderungen von Landnutzung und Landbedeckung aus Satellitenbilddaten zu bestimmen, wurde ein auf Post-Klassifikation basierendes Vergleichsverfahren angewandt. Sechs Landnutzungs- und Landbedeckungsklassen, welche die Namen bewirtschaftetes Land, brach liegendes Land, Waldgebiet, Ödland, besiedeltes Land und Wasserfläche tragen, wurden während des Klassifikationsprozesses bestimmt. Für die vier Aufnahmezeitpunkte der Satellitendaten lag die allgemeine Klassifikationsgenauigkeit zwischen 86% und 92%. Während des dreißigjährigen Untersuchungszeitraums fand eine beträchtliche Veränderung der Landnutzungs- und Landbedeckungsstruktur statt. Bewirtschaftete Flächen nahmen in ihrem Anteil signifikant zu und bedeckten innerhalb des Zeitraums von 1979 bis 2009 81% der früheren Waldgebiete. Der Anteil von brach liegendem Land nahm lediglich während des Zeitraums von 1989 bis 1999 zu. Besiedelte Gebiete breiteten sich über die drei Jahrzehnte kontinuierlich aus und wuchsen innerhalb des Zeitraums von 1979 bis 1989 um eine Fläche von 23 km^2, sowie um 21 km^2 zwischen 1989 und 1999 und um 27 km^2 im Zeitabschnitt 1999–2009.

Eine detaillierte Karte zur Landnutzung und Landbedeckung des Untersuchungsgebiets wurde mittels der Nutzung dual polarisierter (HH und HV) TerraSAR-X Daten aus dem Jahr 2009 erzeugt. Die verschiedenen Landnutzungen und Landbedeckungen im Beobachtungsgelände wurden durch die Anwendung eines objektorientierten Klassifikationsansatzes analysiert. Um Bildobjekte zu erzeugen, wurde für diesen Zweck die auf einer mehrfachen Auflösung basierende Segmentierung der Software Definiens genutzt. Das Werkzeug Feature Space Optimisation wurde für die Optimierung der Attribute der TerraSAR-X Bilder angewandt, damit eine ideale Unterscheidungsfähigkeit entlang der Klassen für die Kartierung der Landnutzungen und Landbedeckungen erreicht werden kann. Zusätzlich zu jenen Klassen, welche mittels optischer Daten abgeleitet wurden, ergaben sich aus SAR-Daten noch die nachfolgenden Landnutzungen und Landbedeckungen: Abgeerntetes Land, Fels, Besiedlung 1 (Gebäude mit landestypischer Bedachung) und Besiedlung 2 (Gebäude mit Betondach).

Seit über sechs Jahrzehnten wird in der Region Gadarif maschinenbetriebener Trockenfeldbau ausgeübt. In Folge dessen fand eine beträchtliche Abholzung und Überweidung sowie eine schwerwiegende Bodendegradation aufgrund des stetigen konventionellen Feldbaus statt. Um die Auswirkungen der Veränderung von Landnutzung und Landbedeckung auf die ausgewählten Bodenbeschaffenheiten auszuwerten, wurden drei Haupttypen der Landnutzung und Landbedeckung für die weitere Untersuchung ausgewählt: Bewirtschaftetes Land, brach liegendes Land, und Waldgebiet.

Auf dem Trockenland-Vertisolboden der Region Gadarif im Sudan wurde mehr als ein Drittel der nationalen Hirseproduktion erwirtschaftet – dem Haupternährungserzeugnis des Landes. Bodenverdichtung erwies sich als eines der weltweiten Hauptprobleme für den Pflanzenbau. Bodenfestigkeit und Versickerungsrate sind wichtige Variabeln, um Bodenprozesse verstehen und vorhersagen zu können. Die Auswirkungen der drei verschiedenen Landnutzungssysteme auf die Bodenverdichtung und Versickerungsrate wurden an zwei Standorten im Beobachtungsgebiet untersucht

Herr Ali Gaber Muhamad Mahmoud, M.Sc., promovierte am 10.1.2012 am Institut für Kartographie der Technischen Universität Dresden mit der Arbeit *„Plot-Based Land-Cover and Soil-Moisture Mapping Using X-/L-Band SAR Data. Case Study Pirna South, Saxony, Germany“* zum Dr. rer. nat.

1. Gutachter: Prof. Dr. Manfred Buchroithner, TU Dresden, Institut für Kartographie
2. Gutachter: Prof. Dr. Richard Gloaguen, TU Bergakademie Freiberg, Institut für Geologie

Kurzfassung: Landwirtschaftliche Produktion erlangt mit weltweit steigender Nahrungsmittelnachfrage zunehmende Bedeutung. Zahlreiche Faktoren bedrohen die landwirtschaftliche Produktion wie beispielsweise die globale Klimaveränderung einschließlich ihrer indirekten Nebenwirkungen. Somit ist das Monitoring der Produktion selbst und der wesentlichen Produktionsparameter eine zweifelsfrei wichtige Aufgabe. Die vorliegende Studie widmet sich in diesem Kontext zwei Schlüsselinformationen, der Aufnahme landwirtschaftlicher Kulturen und den Bodenfeuchteverhältnissen, jeweils unter Nutzung von Satellitenbilddaten von Radarsensoren mit Synthetischer Apertur, die im X- und L-Band operieren.

Landnutzungskartierung spielt eine essentielle Rolle für zahlreiche agrarische Anwendungen; genannt seien hier nur Bewässerungsmaßnahmen, Ernteschätzung und Fördermittelkontrolle. In der vorliegenden Arbeit wurde ein Modell entwickelt, welches auf Grundlage einer Texturanalyse der genannten SAR-Daten für variable Richtungen und Distanzen eine Klassifikation landwirtschaftlicher Nutzungsformen ermöglicht. Das Modell wurde als zusätzliche Funktionalität für die ArcGIS-Software implementiert. Es bindet dabei Klassifikationsverfahren ein, die aus dem Funktionsschatz der Sprache „R“ entnommen sind.

Zum Konzept: Ein Bündel von Texturparametern wird durch das vorliegende Programm auf Schlagbasis berechnet und in einer Polygonattributtabelle der landwirtschaftlichen Schläge abgelegt. Auf diese Attributtabelle greift das nachfolgend einzusetzende Klassifikationsmodul zu. Die Software erlaubt nun die Suche nach „aussagekräftigen“ Teilmengen innerhalb des umfangreichen Texturmerkmalsraumes. Im Klassifikationsprozess kann aus verschiedenen Ansätzen gewählt werden. Genannt seien „Support Vector Machine“ und künstliche neuronale Netze.

Das Modell wurde für einen typischen mitteleuropäischen Untersuchungsraum mit landwirtschaftlicher und gartenbaulicher Nutzung getestet. Er liegt südlich von Pirna im Freistaat Sachsen. Zum Test lagen für den Untersuchungsraum Daten von TerraSAR-X und ALOS/PALSAR (HH/HV) aus identischen Aufnahmetagen vor. Die Untersuchungen beweisen ein hohes Potenzial der Texturinformation aus hoch aufgelösten SAR-Daten für die landwirtschaftliche Nutzungserkennung. Auch die erhöhte Dimensionalität durch die Kombination von zwei Sensoren erbrachte eine Verbesserung der Klassifikationsgüte.

Kenntnisse der Bodenfeuchteverteilung sind u.a. bedeutsam für Bewässerungsanwendungen und hydrologische Modellierung. Die oben genannten SAR-Datensätze wurden auch zur Bodenfeuchteermittlung genutzt. Eine Verifikation wurde durch synchrone Feldmessungen ermöglicht. Initial musste der SAR-typische „Speckle“ in den Bildern durch Filterung verringert werden. Verschiedene Filtertechniken wurden getestet und das beste Resultat genutzt. Die Bodenfeuchtebestimmung erfolgte in Abhängigkeit vom Nutzungstyp über Regressionsanalyse. Auch die Resultate für die Bodenfeuchtebestimmung bewiesen das Nutzpotenzial der genutzten SAR-Daten für offene Ackerböden und Stadien, in denen die Kulturpflanzen noch einen geringen Bedeckungsgrad aufweisen.

Herr M.Sc. Pablo Iván Azócar Fernández promovierte am 9.2.2012 am Institut für Kartographie der Technischen Universität Dresden mit der Arbeit *„Paradigmatic Tendencies in Cartography: A Synthesis of the Scientific-Empirical, Critical and Post-Representational Perspectives“* zum Dr. rer. nat.

1. Gutachter: Prof. Dr. Manfred Buchroithner, TU Dresden, Institut für Kartographie
2. Gutachter: Prof. Dr. Zsolt G. Török, Eötvös Lórand-Universität Budapest, Ungarn

Kurzfassung: Im Entwicklungsprozess verschiedener Gesellschaften sind Karten immer wichtige Elemente visueller Darstellung gewesen. Karten wurden meist aus einer praktischen und utilitaristischen Sicht betrachtet. Das heißt, dass sich Kartographen oder Kar-

tenmacher gezielt auf die technischen Aspekte kartographischer Produkte fokussiert haben, und Kartographie sich nur wenig mit den theoretischen Komponenten und philosophischen oder epistemologischen Aspekten auseinandergesetzt hat. Diese Arbeit verfolgt das Ziel, diese Sichten zu analysieren.

Diese Studie untersucht die verschiedenen kartographischen Denkrichtungen, die während des Positivismus/Empirismus, des Neo-Positivismus und der Post-Strukturalismusperioden entstanden sind und analysiert Kartographie der Moderne und post-moderner Perioden. Argumente von Philosophen wie Ludwig Wittgenstein und Karl Popper werden untersucht als wichtige Beiträge zu unserem Verständnis der Entwicklung der Kartographie. Diese Arbeit berücksichtigt auch das Konzept des Paradigmas, welches aus dem Gebiet der wissenschaftlichen Epistemologie adaptiert wurde. Dies eröffnet die Möglichkeit, Kartographie hinsichtlich eines Paradigmenwechsels analysieren zu können.

Wenn man die Tendenzen der zeitgenössischen Kartographie – von der zweiten Hälfte des zwanzigsten Jahrhunderts bis heute – studiert, muss der theoretische Rahmen einer wissenschaftlichen Disziplin (Forschungsobjekt, Forschungsziel, Arbeitsmethodik und Ergebnisse) berücksichtigt werden. Dies erlaubt es, das gesammelte Wissen der Kartographie zu ermitteln. Ebenfalls wichtig ist die Berücksichtigung des epistemologischen Kontexts, in dem diese Tendenzen entstanden: Positivismus/Empirismus, Realismus/Strukturalismus und Idealismus/Hermeneutik.

Unter Berücksichtigung dreier epistemologischer Ebenen – Essenzialisten/ontisch (wissenschaftlich), dekonstruktiv (soziologisch) und ontologisch (emergent) – werden ausgewählte paradigmatische Tendenzen postuliert. Die erste Ebene ergibt Tendenzen wie die kartographische Kommunikation, die kartographische Semiotik, die analytische Kartographie und die kartographische Visualisierung, die alle zu der wissenschaftlich-empirischen Perspektive gehören. Zur zweiten Ebene gehört die kritische Kartographie, welche der kritischen Perspektive zugeordnet ist und die wissenschaftliche Standpunkte konfrontiert. Die so genannte post-repräsentative Kartographie entsteht aus der dritten Ebene im offenen Widerstand zur traditionellen repräsentativen Kartographie.

Herr Dipl.-Ing. Patrick Westfeld promovierte am 8.5.12 an der Fakultät Forst-, Geo- und Hydrowissenschaften der TU Dresden mit der Arbeit „*Geometrische und stochastische Modelle zur Verarbeitung von 3D-Kameradaten am Beispiel menschlicher Bewegungsanalysen*“ zum Dr.-Ing.

1. Gutachter: Prof. Dr. sc. techn. habil. Hans-Gerd Maas, Technische Universität Dresden
2. Gutachter: Prof. Dr. rer. nat. Ralf Reulke, Humboldt-Universität zu Berlin

Kurzfassung: 3D-Kameras verbinden die Handlichkeit einer Digitalkamera mit dem Potential der dreidimensionalen Datenakquisition etablierter Oberflächenmesssysteme. Sie stellen trotz der noch vergleichsweise geringen räumlichen Auflösung als monosensorielles System zur Echtzeit-Tiefenbildakquisition eine interessante Alternative für Aufgabenstellungen der 3D-Bewegungsanalyse dar. Die Schwerpunkte der Dissertation liegen dabei auf Modell- und Verfahrensentwicklung zur Sensorkalibrierung und zur 3D-Bewegungsanalyse am Beispiel der Untersuchung des zwischenmenschlichen Bewegungsverhaltens.

Eine 3D-Kamera stellt aufgrund ihres inhärenten Designs und Messprinzips gleichzeitig Amplituden- und Entfernungsinformationen zur Verfügung, welche aus einem Messsignal rekonstruiert werden. Die simultane Einbeziehung aller 3D-Kamerainformationen in jeweils einen integrierten Ansatz ist eine logische Konsequenz und steht im Vordergrund der Verfahrensentwicklungen. Zum einen stützen sich die komplementären Eigenschaften der Beobachtungen durch die Herstellung des funktionalen Zusammenhangs der Messkanäle gegenseitig, wodurch Genauigkeits- und Zuverlässigkeitssteigerungen zu erwarten sind. Zum anderen gewährleistet das um eine Varianzkomponentenschätzung erweiterte stochastische Modell eine vollständige Ausnutzung des heterogenen Informationshaushalts. Die entwickelte integrierte Bündelblockausgleichung ermöglicht die Bestimmung der exakten 3D-Kamerageometrie sowie die Schätzung der distanzmessspezifischen Korrekturparameter zur Modellierung linearer, zyklischer und signalwegeffektbedingter Fehleranteile einer 3D-Kamerastreckenmessung. Die integrierte Kalibrierroutine gleicht in beiden Informationskanälen gemessene Größen gemeinsam, unter der automatischen Schätzung optimaler Beobachtungsgewichte, aus. Die Methode basiert auf dem flexiblen Prinzip einer Selbstkalibrierung und benötigt keine Objektrauminformation, wodurch insbesondere die aufwendige Ermittlung von Referenzstrecken übergeordneter Genauigkeit entfällt. Die durchgeführten Genauigkeitsuntersuchungen bestätigen die Richtigkeit der aufgestellten funktionalen Zusammenhänge, zeigen aber auch Schwächen aufgrund noch nicht parametrisierter distanzmessspezifischer Fehler. Die Adaptivität und die modulare Implementierung des entwickelten mathematischen Modells gewährleisten aber eine zukünftige Erweiterung. Die Qualität der 3D-Neupunktkoordinaten kann nach einer Kalibrierung mit 5 mm angegeben werden. Für die durch eine Vielzahl von meist simultan auftretenden Rauschquellen beeinflusste Tiefenbildtechnologie ist diese Genauigkeitsangabe sehr vielversprechend, vor allem im Hinblick auf die Entwicklung von auf korrigierten 3D-Kameradaten aufbauenden Auswertealgorithmen.

Für die Bildzuordnung wurde ein 2,5D *Least Squares Tracking* (LST) Verfahren entwickelt, das sowohl Grau- als auch Entfernungswerte in einem integrierten Modell verknüpft. Die aus der Fehlerrechnung resultierenden Genauigkeits- und Zuverlässigkeitsangaben können als Entscheidungskriterien für die Integration in einer anwendungsspezifischen Verarbeitungskette Verwendung finden. Die Validierung des Verfahrens zeigte, dass die Einführung komplementärer Informationen eine genauere und zuverlässigere Lösung des Korrespondenzproblems bringt, vor allem bei schwierigen Kontrastverhältnissen in einem Kanal. Die Genauigkeit der direkt mit den Distanzkorrekturtermen verknüpften Maßstabs- und Neigungsparameter verbesserte sich deutlich. Darüber hinaus brachte die Erweiterung des geometrischen Modells insbesondere bei der Zuordnung natürlicher, nicht gänzlich ebener Oberflächensegmente signifikante Vorteile.

Das Potential des 3D-Kamerazuordnungsansatzes wurde an zwei Anwendungsszenarien der menschlichen Verhaltensforschung demonstriert. 2,5D-LST kam zur Bestimmung der interpersonalen Distanz und Körperorientierung im erziehungswissenschaftlichen Untersuchungsgebiet der Konfliktregulation befreundeter Kindespaare ebenso zum Einsatz wie zur Markierung und anschließenden Klassifizierung von Bewegungseinheiten sprachbegleitender Handgesten. Die Arbeit zeigt, dass die Methoden der Photogrammetrie auch für bewegungsanalytische Aufgabenstellungen auf dem bisher noch wenig erschlossenen Gebiet der Verhaltensforschung wertvolle Ergebnisse liefern können. Damit leistet sie einen Beitrag für die derzeitigen Bestrebungen in der automatisierten videographischen Erhebung von Körperbewegungen in dyadischen Interaktionen.

Veranstaltungskalender

2012

4.–6. Juli: **AGIT** 2012 – Symposium und Fachmesse für Angewandte Geoinformatik in **Salzburg**, Österreich. agit.at

22.–27. Juli: **IGARSS 2012**, International Geoscience and Remote Sensing Symposium, in **München**. igarss12.org

24. August–3. September: **XXII ISPRS Congress** 2012 in **Melbourne**, Australien. isprs2012-melbourne.org

4.–5. Oktober: 1. Gemeinsames Arbeitskreistreffen des **AK „Fernerkundung" der DGfG** und des **AK „Interpretation von Fernerkundungsdaten" der DGPF** in **Bochum**. ak-fernerkundung.de

7.–13. Oktober: **ECCV 2012**, European Conference on Computer Vision, in **Florenz**, Italien. eccv2012.unifi.it

9.–11. Oktober: **INTERGEO 2012** und **60. Deutscher Kartographentag** und **3. Europäischer Kongress der CLGE** (Comité de Liaison des Géomètres Européens) in **Hannover**. intergeo.de

16.–18. Oktober: 9th **Symposium on Location-Based Services** in **München**. lbs2012.tum.de

6.–7. Dezember: **3D-NordOst 2012**, 15. Anwendungsbezogener Workshop zur Erfassung, Modellierung, Verarbeitung und Auswertung von 3D-Daten am GFaI in Berlin-Adlershof. 3d-nordost.de.

2013

17.–23. Februar: 17. **Internationale Geodätische Woche** in **Obergurgl**, Österreich. uibk.ac.at/vermessung/veranstaltung/obergurgl.html

27. Februar–1. März: Dreiländertagung D-A-CH der DGPF, ÖVG, SGPF in Freiburg. dgpf.de

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

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

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

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

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

Korporative Mitglieder

Firmen

AEROWEST GmbH
AICON 3D Systems GmbH
aphos Leipzig AG
Becker GeoInfo GmbH
Bernhard Harzer Verlag GmbH
Blom Deutschland GmbH
Brockmann Consult GmbH
bsf swissphoto GmbH
Büro Immekus
CGI Systems GmbH
con terra GmbH
DB Netz AG
DELPHI IMM GmbH
Deutsches Bergbau-Museum
EFTAS Fernerkundung Technologietransfer GmbH
ESG Elektroniksystem- und Logistik-GmbH
Esri Deutschland GmbH
EUROPEAN SPACE IMAGING
Eurosense GmbH
fokus GmbH
g.on experience gmbh
GAF GmbH
GeoCart Herten GmbH
GeoContent GmbH
Geoinform. & Photogr. Engin. Dr. Kruck & Co. GbR
geoplana Ingenieurgesellschaft mbH
GEOSYSTEMS GmbH
GGS - Büro für Geotechnik, Geoinformatik, Service
Hansa Luftbild AG
IGI - Ingenieur-Gesellschaft für Interfaces mbH
ILV Ingenieurbüro für Luftbildauswertung und Vermessung
Imetric 3D GmbH
Infoterra GmbH
INVERS - Industrievermessung & Systeme
ITT Visual Information Solutions Germany
J. Linsinger ZT-GmbH
Jena-Optronik GmbH
Leica Geosystems GmbH
Luftbilddatenbank-Würzburg
Messbildstelle GmbH
Microsoft Photogrammetry
MILAN Geoservice GmbH
M.O.S.S. Computer Grafik Systeme GmbH
PHOENICS GmbH
PMS - Photo Mess Systeme AG
RIEGL Laser Measurement Systems GmbH
RWE Power AG, Geobasisdaten/Markscheidewesen
technet GmbH
TRIGIS Vermessung + Geoinformatik GmbH
Trimble Germany GmbH
trimetric 3D Service GmbH
Wichmann, VDE Verlag GmbH
Z/I Imaging Ltd.

Behörden

Amt für Geoinformationswesen der Bundeswehr
Bayerische Landesanstalt für Wald und Forstwirtschaft
Bundesamt für Kartographie und Geodäsie
Bundesmininisterium für Ernährung, Landwirtschaft und Verbraucherschutz
Hessisches LA für Bodenmanagement und Geoinformation
Innenministerium NRW, Gruppe Vermessungswesen
Institut für Umwelt- und Zukunftsforschung
LA für Geoinformation und Landentwicklung, BW
LA für Vermessung und Geoinformation, Bayern
LB Geoinformation und Vermessung, Hamburg
LB für Küstenschutz, Nationalpark und Meeresschutz, SH
Landesvermessung und Geobasisinformation Niedersachsen
Märkischer Kreis, Vermessungs- und Katasteramt
Regierungspräsident Tübingen, Abt. 8 Forstdirektion
Regionalverband Ruhr
Staatsbetrieb Sachsenforst Pirna
Stadt Bocholt, Fachbereich 31
Stadt Düsseldorf, Vermessungs- und Katasteramt
Stadt Köln, Amt für Liegenschaften, Vermessung und Kataster
Stadt Wuppertal, Vermessung, Katasteramt und Geodaten
Thüringer LA für Vermessung und Geoinformation

Hochschulen

BTU Cottbus, Lehrstuhl für Vermessungskunde
FH Frankfurt a.M., FB 1, Studiengang Geoinformation
FH Mainz, Institut für Raumbezogene Informations- und Messtechnik
Jade Hochschule, Institut für Angewandte Photogrammetrie und Geoinformatik
HCU HafenCity Universität Hamburg, Geomatik
HfT Stuttgart, Vermessung und Geoinformatik
HS Bochum, FB Vermessung und Geoinformatik
HS Karlsruhe, Fakultät für Geomatik
HTW Dresden, FB Vermessungswesen/Kartographie
LUH Hannover, Institut für Kartographie und Geoinformatik
LUH Hannover, Institut für Photogrammetrie und Geoinformation
MLU Halle, FG Geofernerkundung
Ruhr-Uni Bochum, Geographisches Institut
RWTH Aachen, Geodätisches Institut
TU Bergak. Freiberg, Institut für Markscheidewesen und Geodäsie
TU Berlin, Computer Vision & Remote Sensing
TU Berlin, Institut für Geodäsie und Geoinformationstechnik
TU Braunschweig, Institut für Geodäsie und Photogr.
TU Clausthal, Institut für Geotechnik und Markscheidewesen
TU Darmstadt, Institut für Photogrammetrie und Kartographie
TU Dresden, Institut für Photogrammetrie und Fernerkundung
TU München, FG Photogrammetrie und Fernerkundung
TU Wien, Institut für Photogrammetrie und Fernerkundung
Uni Bonn, Institut für Photogrammetrie
Uni Göttingen, Institut für Waldinventur und Waldwachstum
Uni Heidelberg, IWR Interdisziplinäres Zentrum für Wissenschaftliches Rechnen
Uni Kassel, FB Ökologische Agrarwissenschaften
Uni Kiel, Geographisches Institut
Uni Stuttgart, Institut für Photogrammetrie
Uni Würzburg, Geographisches Institut
Uni zu Köln, Geographisches Institut