PFG

Photogrammetrie Fernerkundung Geoinformation

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

Jahrgang 2007, Heft 6

Hauptschriftleiter:
Prof. Dr.-Ing. Helmut Mayer

Schriftleiter:
Prof. Dr. rer. nat. Carsten Jürgens, Prof. Dipl.-Ing. Thomas P. Kersten, Prof. Dr. rer. nat. Lutz Plümer und Dr.-Ing. Eckhardt Seyfert

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

E. Schweizerbart'sche Verlagsbuchhandlung
(Nägele u. Obermiller) Stuttgart 2007

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

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

International Society
for Photogrammetry
and Remote Sensing

Deutsche Arbeits-
gemeinschaft für
Mustererkennung e. V.

GeoUnion
Alfred-Wegener-Stiftung

Herausgeber:
© 2007 Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF) e.V.
Präsident: Prof. Dr.-Ing. Thomas Luhmann, Fachhochschule Oldenburg Ostfriesland Wilhelmshaven, Institut für Angewandte Photogrammetrie und Geoinformatik, Ofener Str. 16, D-26121 Oldenburg, Tel.: +49 (0)441 7708-3172, e-mail: Praesident@dgpf.de, www.dgpf.de
Geschäftsstelle: Dr. Klaus-Ulrich Komp, c/o EFTAS Fernerkundung Technologietransfer GmbH, Oststraße 2–18, D-48145 Münster, e-mail: klaus.komp@eftas.com
Published by:
E. Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller), Johannesstraße 3 A,
D-70176 Stuttgart. Tel.: 0711 351456-0, Fax: 0711 351456-99, e-mail: mail@schweizerbart.de
Internet: http://www.schweizerbart.de
♾ Gedruckt auf alterungsbeständigem Papier nach ISO 9706-1994
All rights reserved including translation into foreign languages. This journal or parts thereof may not be reproduced in any form without permission from the publishers.
Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in dieser Zeitschrift berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.
Verantwortlich für den Inhalt der Beiträge sind die Autoren.
ISSN 1432-8364

Hauptschriftleiter: Prof. Dr.-Ing. Helmut Mayer, Institut für Photogrammetrie und Kartographie, Universität der Bundeswehr München, D-85577 Neubiberg, e-mail: Helmut.Mayer@unibw.de
Schriftleiter: Prof. Dr. rer. nat. Carsten Jürgens, Ruhr-Universität Bochum, Geographisches Institut, Gebäude NA 7/133, D-44780 Bochum, e-mail: carsten.juergens@rub.de, Prof. Dipl.-Ing. Thomas P. Kersten, HafenCity Universität Hamburg, Department Geomatik, Hebebrandstr. 1, D-22297 Hamburg, e-mail: thomas.kersten@hcu-hamburg.de, Prof. Dr. rer. nat. Lutz Plümer, Universität Bonn, Institut für Geodäsie und Geoinformation, Meckenheimer Allee 172, D-53115 Bonn, e-mail: Lutz.Pluemer@ikg.uni-bonn.de und Dr.-Ing. Eckhardt Seyfert, Landesvermessung und Geobasisinformation Brandenburg, Heinrich-Mann-Allee 107, D-14473 Potsdam, e-mail: eckhardt.seyfert@geobasis-bb.de
Erscheinungsweise: 7 Hefte pro Jahrgang.
Bezugspreis im Abonnement: € 122,– pro Jahrgang. Mitglieder der DGPF erhalten die Zeitschrift kostenlos.
Anzeigenverwaltung: Dr. E. Nägele, E. Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller), Johannesstraße 3A, D-70176 Stuttgart, Tel.: 0711 351456-0; Fax: 0711 351456-99.
e-mail: mail@schweizerbart.de, Internet: http://www.schweizerbart.de
Bernhard Harzer Verlag GmbH, Westmarkstraße 59/59a, D-76227 Karlsruhe, Tel.: 0721 944020, Fax: 0721 9440230, e-mail: Info@harzer.de, Internet: www.harzer.de
Printed in Germany by Tutte Druckerei GmbH, D-94121 Salzweg bei Passau

PFG – Jahrgang 2007, Heft 6
Inhaltsverzeichnis

Originalbeiträge DFG „China-Bündel“ – Interoperation of 3D Urban Geoinformation

Berichte und Mitteilungen

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

Photogrammetrie • Fernerkundung • Geoinformation 6/2007, S. 401–404

Editorial: Interoperation of 3D Urban Geoinformation

LIQIU MENG, München; CHRISTIAN HEIPKE, Hannover; HELMUT MAYER, Neubiberg & JUN CHEN, Beijing, China

1 Background

3D city models have grown in popularity during recent years, especially since the Google Earth revolution. They offer an intuitive organization of spatial objects from the real world, thus allowing a natural perception and understanding of urban information. Indeed, there is a huge potential to use 3D city models as communication language and working tools in a large number of fields, e. g., architectural construction, urban planning, virtual tourism, civil engineering, mobile telecommunication, navigation, facility management, disaster simulation, and mission rehearsal. In spite of the considerably different requirements with regard to spatial coverage, currentness, semantic information, and spatio-temporal accuracy these applications share the common demand for an easy access to detailed 3D urban information.

However, due to insufficient geometric and semantic granularity as well as the lack of interoperability, the available 3D city models are primarily treated as individual show pieces instead of urban information systems. More specifically, the state-of-the-art is characterized by

- separate and partly redundant acquisition of urban information and 3D city models mostly at a rather coarse granularity level,
- inconsistent modeling and updating of urban objects based on ad-hoc specifications for levels of detail,
- limited re-use of existing urban-related datasets due to incompatibility of data models,
- limited updating schemes that account for the rapid changes in the urban environment,
- incomplete, heterogeneous, and partly inaccurate coverage of 3D city models in large regions,
- little semantic information integrated with the geometries of individual 3D urban objects,
- visualization limited to the representation of the appearance of urban objects, and
- few analytical functions available in 3D urban information systems.

With the rapid development of high resolution sensors and multimedia technologies as well as the increasing computing performance, it is now possible to devise and implement methods spanning over the whole process from modeling, acquisition, visualization to analysis of highly detailed and current 3D urban information. Being aware of the potential of high-end techniques and the importance of establishing national and international geodata infrastructures, scientists from Germany and China in the field of photogrammetry, remote sensing, cartography, and geoinformatics jointly launched a bilateral bundle research project on the topic of "Interoperation of 3D Urban Geoinformation" in 2006. 14 research teams from the following institutions of both countries have been playing an active role in the joint bundle:

- Institute of Computer Engineering and Microelectronics, Technical University of Berlin
- Institute of Geodesy and Geoinformation, University of Bonn
- Institute of Cartography and Geoinformatics, Leibniz University of Hanover
- Institute of Photogrammetry and GeoInformation, Leibniz University of Hanover
- Institute of Photogrammetry and Cartography, Technical University of Munich

1432-8364/07/2007/0401 $ 1.00
© 2007 DGPF / E. Schweizerbart'sche Verlagsbuchhandlung, D-70176 Stuttgart

- Institute of Photogrammetry and Cartography, Bundeswehr University Munich
- China University of Information Engineering, Zhengzhou
- China University of Mining Technology, Beijing
- Hong Kong Polytechnic University
- Hong Kong Baptist University
- National Geomatics Center of China, Beijing
- Wuhan University

Being supported by the Deutsche Forschungsgemeinschaft (DFG) and the National Natural Science Foundation of China (NNSFC), the bundle project aims at promoting the seamless availability and accessibility of up-to-date 3D urban information with the finest possible geometric and semantic details. Trees, roads, land use areas and buildings of urban or suburban areas have been selected as test data. This special journal issue is devoted to the methods, techniques, case studies and the first results of six involved subprojects of the German side.

2 Research Challenges

The goal of making 3D urban geoinformation interoperable is associated with a number of research challenges. At first, the existing methods and techniques applied in two-dimensional GIS need a substantial extension to deal with the inherently complex 3D scenes in both, the geometric and the semantic sense. Second, urban information is of a highly dynamic nature and is related to densely populated areas. It is therefore by no means a trivial task to model the fine-grained 3D structures along with their temporal aspects and to develop efficient strategies for 3D data acquisition, updating, and retrieval. Finally, the realization of an interoperable GIS requires to (1) integrate data sets of different origins, scales, resolutions, formats, spatio-temporal scopes and semantic contents, (2) define quality measures for the integration, and (3) determine the quality of the integrated data.

As a whole, it is far more intricate to reconstruct, describe and render interoperable 3D urban objects than 2D footprints. Taking the object type "building" as an example, a 2D building is represented by its cadastral ground plan, while a 3D building model is characterized by a large number of surface elements. The meaning of a 2D building is usually expressed by a number of semantic attributes attached to its ground plan. In 3D space, however, every surface element of a building could be assigned a number of special semantic attributes.

The larger amount of information necessary to characterize a 3D building implies more effort for reliable data acquisition and maintenance. The relationship between two individual 2D buildings can be described by location, form, size, orientation, proximity, horizontal alignment, and semantic attributes of their ground plans, whereas two individual 3D buildings are additionally related by the vertical alignment, roof structures, wall elements, building heights, surface textures etc.

The visualization of a 2D building is conventionally confined to a plan view, while 3D rendering is determined by more parameters in terms of viewing distance, viewing angle, eye level, vision field etc. Similar arguments hold for other urban objects. Moreover, thematic overlay on a 3D presentation is difficult. Although many GIS vendors tend to expand their toolkits to include 3D interactions, users are often only allowed to change the visualization parameters such as camera position, illumination, color and texture, or conduct some simple calculations on visible objects. The individual 3D objects or object parts and their associated attributes are largely inaccessible, especially when they are occluded.

With regard to the above mentioned challenges, this special journal issue touches basic research questions on

- how to model and capture highly detailed geometric and semantic information of 3D urban objects,
- how to ensure the geometric, topological, attributive, and temporal quality of the available urban or sub-urban information,
- how to re-use existing 3D city models,

- how to harmonize 3D urban information from different sources,
- how to enrich 3D geometric models with additional semantic attributes, and
- how to communicate 3D urban geoinformation to diverse end users.

This special issue does not intend to cover all these questions in a balanced way. It rather tries to highlight some essential problems that are frequently encountered in the efforts toward the interoperation of 3D urban information system.

3 Contents Overview

Six papers are included in the special issue. They introduce a number of multidisciplinary methods and techniques that are rooted in mathematics, physics, computer science, and a number of engineering sciences such as electronic engineering, geodetic engineering, and geoinformatics. The contents reflect the following three focuses:

- modeling and reconstruction of highly detailed 3D urban objects such as facade elements, stairways, and structures of tree branches,
- updating and quality assurance of existing urban or suburban database, e. g., 3D building models and road databases, and
- geometric and semantic enrichment of 2D datasets (e. g., topographic objects of different resolutions) and 3D city models (e. g., facade textures in the infrared domain).

The work of Heinrichs et al. deals with a generic strategy for reconstructing architectural models using multiple mobile terrestrial video sequences from a trifocal sensor. In a prototype system, triangulation of spatial stereo is combined with tracking of temporal stereo. The trinocular rectification of uncalibrated images is solved in closed form with six degrees of freedom. An algorithm for image matching on the basis of trinocularly rectified images and semi-global optimization is elucidated and evaluated. The motion of the video sensor is estimated using temporal feature tracking and it is possible to generate dense point clouds. First experiments with image sequences from a real scene resulted in a highly detailed metric model of the scene, clearly demonstrating the feasibility of the proposed strategy.

Schmittwilken et al. report over a semantic model for the transition areas from building facades to the surrounding terrain which is termed "building collar". Their experiment is focused on stairways – one of the most frequently occurring features within a building collar. Stairways of different geometric complexities together with their semantic constraints and functional aspects are represented by Unified Modeling Language (UML) and Object Constraint Language (OCL) notation. It is then possible to derive an attribute grammar consisting of productions and semantic rules from UML/OCL. The paper introduces a procedure for mapping UML/OCL to an attribute grammar which is then used to specify a subset of stairs observed in reality. Further, a generic method to solve the constraints and to generate stair models is described. Although the currently implemented prototype is dedicated to instantiate a limited number of simple 3D stair models only, it opens up the way to formalize and reconstruct more complicated stair models and other 3D features in high resolution building models.

Huang & Mayer address the extraction of the 3D branching structure of unfoliaged deciduous trees from wide-baseline urban image sequences. By combining the descriptive power of Lindemeyer-Systems (L-Systems) for trees with statistical sampling based on Markov Chain Monte Carlo (MCMC) and cross correlation, individual 3D trees are generatively modeled and extracted. The approach allows the reconstruction of realistic branch structures even when they are partly occluded or mixed up with noisy background. This potential has been confirmed by the results of first experiments. The statistical analysis of tree parameters could provide useful information for ecological applications and the detailed 3D branching structure allows for animated visualizations, e. g., for windy days. The authors have started to extend the current

work to open L-systems that will make it possible to recognize different types of trees.

A two-stage strategy of using high resolution aerial orthoimages to verify road databases for suburban areas is described by GROTE et al. In the first stage, the Normalized Cuts algorithm, a graph-based approach, is applied to segment the image on the basis of pixel similarities reflected by colour and edge properties. After grouping the segments containing road parts are then separated from the non-road segments by means of shape criteria. The second stage is devoted to the verification of the road database by comparing the road parts in the image segments with their counterparts in the database. The preliminary experiments have proved the suitability of the segmentation approach. Further, an iterative approach is introduced which should overcome the existing shortcomings of the Normalized Cuts algorithm concerned with the necessity of indicating the number of segments a priori. As to the implementation of the verification approach, an available platform developed for the quality assurance of rural road information is in the process of being adapted for road information in suburban areas.

KIELER et al. attempt to improve the interoperability between different geospatial datasets on the basis of semantic integration with data matching as one of the core techniques. Starting from simple cases with known semantic relationships between the datasets to be matched, they use the semantic correspondences to enrich similar objects or object classes. For the case of unknown semantic relationships, they first establish the geometric correspondences between two different datasets and subsequently infer the semantic correspondences. For this latter task, the possible differences with respect to geometric and semantic granularity are considered. The experience gained so far reveals good matching results for data sets with different granularity levels and similar currentness. The matching process will be refined and extended with additional attribute information. Meanwhile, the authors strive for an automatic analysis of object correspondences with the help of spatial data mining techniques.

Methods of extracting textures of building facades from low resolution infrared (IR) image sequences and integrating the results with a 3D database are demonstrated in HOEGNER et al. This work aims at enriching the currently available 3D building models with the information from non-visible domains such as thermal textures. The images and the given 3D model are first matched with two alternative strategies. One is based on the correction of the camera parameters using three corresponding points, while the other relies on the estimation of planes in the image sequence using homographies. According to the principles of ray casting, partial textures can subsequently be generated for every visible surface in each single image of a sequence. Different partial textures belonging to the same facade are then combined. The resulting textures are finally integrated with a 3D database in a format conformal to CityGML which will likely become an open source standard. The enriched 3D building models serve as the starting point for the subsequent multi-purpose visualization and interactive analysis.

4 Acknowledgements

We wish to share our experiences, learnt lessons, and gained insight with the readers of this special issue. We gratefully acknowledge the generous support of the Deutsche Forschungsgemeinschaft (DFG) and the National Natural Science Foundation of China (NNSFC) to the ongoing bilateral scientific cooperation. Finally, we would like to express our sincere thanks to the Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF) and everybody involved in preparing this special issue for allowing us to publish the first results of this ongoing Sino-German bundle project in the PFG.

Prof. Dr.-Ing. LIQIU MENG, Prof. Dr.-Ing. CHRISTIAN HEIPKE, Prof. Dr.-Ing. HELMUT MAYER, Prof. Jun CHEN, e-mail: meng@bv.tum.de, Heipke@ipi.uni-hannover.de, Helmut.Mayer@unibw.de, Chen-Jun@nsdi.gov.cn

Photogrammetrie • Fernerkundung • Geoinformation 6/2007, S. 405–414, 6 figs.

ARTIST: Architectural Model Refinement Using Terrestrial Image Sequences from a Trifocal Sensor

MATTHIAS HEINRICHS, OLAF HELLWICH & VOLKER RODEHORST, Berlin

Keywords: Close-range Photogrammetry, Architectural Model Refinement, Trifocal Video Sensor, Trinocular Rectification, Spatial Stereo, Semi-global Matching, Temporal Tracking

Summary: This paper proposes a high resolution video sensor for the 3D reconstruction of architectural models from multiple image sequences. The hybrid system unifies triangulation methods of spatial stereo with tracking methods of temporal stereo. We describe an efficient spatial image matching algorithm, which is based on trinocular image rectification and semi-global optimization. The motion of the video sensor is estimated using temporal feature tracking and allows the integration of dense point clouds. First experimental results are shown for images of a real scene.

Zusammenfassung: *Verfeinerung von Gebäudemodellen mittels terrestrischer Videosequenzen eines Trifokalsensors*. In diesem Beitrag wird ein hochauflösender Videosensor für die 3D Rekonstruktion von Architekturmodellen aus mehreren Bildfolgen vorgestellt. Das hybride System vereinheitlicht Triangulationsmethoden des räumlichen Stereos mit Verfolgungsmethoden des zeitlichen Stereos. Wir beschreiben ein effizientes Bildzuordnungsverfahren, das auf einer trinokularen Bildkorrektur und semi-globaler Optimierung basiert. Die Bewegung des Videosensors wird durch zeitliche Merkmalsverfolgung geschätzt und erlaubt die Integration dichter Punktwolken. Erste experimentelle Ergebnisse für Bilder einer realen Szene werden vorgestellt.

1 Introduction

Some interesting applications of urban 3D geographic information systems (GIS) require a level of detail (LOD), which is currently not available using airborne data. Therefore, an approach to acquire and/or refine architecture models from terrestrial image sequences is proposed. We develop a fully automated prototype system to recover 3D models of several buildings based on three moving video cameras (trifocal sensor). In general, digital video cameras provide dense image sequences that contain a high potential for photogrammetric application which is presently not fully used. Video sequences for scene modeling and various possible applications are treated by (AKBARZADEH et al. 2006, POLLEFEYS et al. 2004, KOCH 2003). When dense video sequences are used for object reconstruction the correspondence problem does not have to be solved by wide-baseline matching any longer but tracking and motion estimation methods such as affine feature tracking (SHI & TOMASI 1994), visual odometry (NISTÉR 2006), simultaneous localization and mapping SLAM (MONTEMERLO 2003) and optical flow estimation gain importance.

Reconstructing a three-dimensional model from a single video sequence is often conducted with the structure-from-motion SFM approach. In close-range photogrammetry systems have been mounted on vans in order to acquire GIS data semi-automatically (TAO 2000). First attempts on hybrid

1432-8364/07/2007/0405 $ 2.50
© 2007 DGPF / E. Schweizerbart'sche Verlagsbuchhandlung, D-70176 Stuttgart

algorithms unifying triangulation methods of spatial stereo with tracking methods of temporal stereo are presented by (Neumann & Aloimonos 2002). They propose multi-resolution subdivision surfaces for spatio-temporal stereo. Algebraic projective geometry (Hartley & Zisserman 2003, Faugeras & Luong 2001) provides an effective mathematical framework to obtain geometrically precise information from partially calibrated cameras with varying parameters.

The main novelty of our approach is an exhaustive integration of feature extraction, image matching, orientation for video sequences, as well as modeling of surfaces with their reflectance characteristics. We combine calibrated and relatively oriented trifocal image geometry with temporal tracking in video sequences to generate a photogrammetric model (Zheng et al. 2007). In this scenario, the photogrammetric model is partially reconstructed from the neighboring images of the triplet, partially from the preceding and following images of the sequence. Due to two facts the trifocal video system allows generating a reliable photogrammetric model for each image triplet with little computational effort. First, each candidate triplet of corresponding points has a high potential to be correct as the matching between images of the triplet is stabilized by a tracking approach for points in each of the video sequences. Second, the trifocal tensor allows checking each triplet based on the relative orientation of the cameras. Thus, the system basically acquires a three-dimensional image which is used in a tracking procedure.

Practically useful results can only be derived when the accuracy achieved fulfills photogrammetric standards. At the same time, only an automatic processing would ensure to make use of the full potential of video sequences. However, the computational requirements to deal with hand-held markerless video streams exceed the capabilities of real-time systems. Therefore, the proposed approach is designed for off-line processing of real-time recorded digital video.

The organization of the paper is as follows. In section 2, the trifocal video sensor is briefly introduced. The efficient spatial image matching algorithm, which is based on trinocular image rectification and semi-global optimization, is explained in the subsequent sections 3 and 4. The estimation of the camera motion and the registration of the 3D point clouds using temporal feature tracking are discussed in section 5. First experimental results for images of a real scene are presented in section 6 and finally we conclude and state possible improvements.

2 System Overview

In the actual stage the trinocular stereo rig consists of three Scorpion color cameras from Point Grey Research (PTGrey). The image sensor is based on a progressive scan CCD with square pixels. Each camera is able to acquire and transmit a high resolution digital video sequence (1384×1038 pixels) with up to 19 bayer tiled full frames per second using firewire (IEEE 1394a). The cameras are synchronized with an accuracy of less than 1 ms. The effective data rate of the three cameras is around 80 MB and exceeds the bandwidth of one firewire channel as well as the writing performance of a regular harddrive. Therefore, a desktop PC with three independent 1394a-channels and a RAID-0 array consisting of four SATA disks was assembled. The system is designed to capture video with a maximum data rate up to 200 MB/sec for more than three hours using a battery based power source (see Fig. 1).

For a flexible image acquisition, we selected CCTV-lenses with variable principal distance between 6 mm and 12 mm. We discovered a significant radial lens distortion up to 80 pixels. Therefore, we model radial errors with 3 additional coefficients and resample all images using bicubic interpolation. The undistorted images simplify the geometric imaging model to a line-preserving pinhole camera. In addition to our stationary control point field we developed a mobile calibration rig (see Fig. 2). It allows an on-site calibration in few seconds due to an automatic marker detection and fitting

Fig. 1: a) Trinocular stereo rig with mounted Scorpion cameras and b) the mobile image acquisition system.

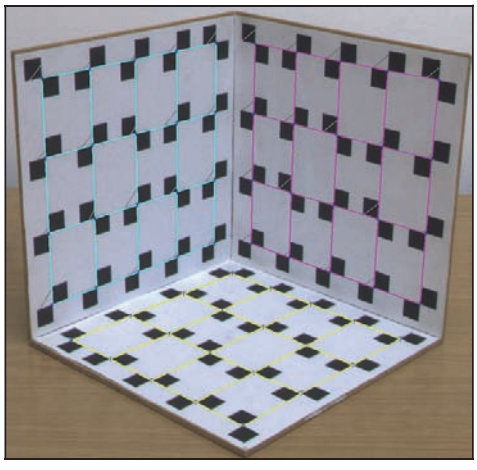

Fig. 2: Image of the mobile calibration rig with color coded results of the 3D model fitting.

algorithm for parameterized 3D models (Lowe 1991).

3 Trinocular Rectification

This section describes the geometric transformation of an uncalibrated image triplet to the stereo normal case (Heinrichs & Rodehorst 2006). In case of a trinocular rectification, the images are reprojected onto a plane, which lies parallel to the projection centers. The proposed trinocular rectification method requires an image triplet with more or less L-shaped camera alignment. The camera configuration is arbitrary, but each projection center must be invisible in all other images. This condition is necessary, since otherwise the epipoles lie in the image and mapping them to infinity will lead to unacceptable distortion of the images.

Furthermore, we assume non-degenerate camera positions, where the camera centers are not collinear, because collinear setups can be rectified by chaining a classical binocular rectification approach. Additionally, a common overlapping area and at least six corresponding image points are necessary, so that the trifocal tensor, the fundamental matrices and the epipoles can be determined. The result consists of three geometrically transformed images, in which the epipolar lines run parallel to the image axes. The resampling due to radial distortion and for trinocular rectification is processed in one step to minimize image blur.

3.1 *Camera Setup*

A given image triplet consists of the original images b (base), h (horizontal) and v (vertical). Subsequently, we denote the rectified images $\tilde{b}$, $\tilde{h}$ and $\tilde{v}$. The rectification tries to fit any image triplet to an L-configuration. This setup has the following properties:

- The epipolar lines of image $\tilde{b}$ and image $\tilde{h}$ correspond with their image rows.
- The epipolar lines of image $\tilde{b}$ and image $\tilde{v}$ correspond with their image columns.

- The epipolar lines of image $\tilde{h}$ and image $\tilde{v}$ have a slope of minus unity.

The last property has the advantage, that the disparities between corresponding points in $\tilde{b} \leftrightarrow \tilde{h}$ and $\tilde{b} \leftrightarrow \tilde{v}$ are equal. The basic idea of rectification is to map the epipoles e between the images b, h and v to infinity.

3.2 *Linear Rectification*

The initial task is to determine the relative orientation of the images. The fully automatic approach uses interest point locations from a modified FÖRSTNER operator (RODEHORST & KOSCHAN 2006) in combination with the SIFT descriptor (LOWE 2004) for matching. We have implemented a robust estimation of the trifocal tensor **T**, which describes the projective relative orientation of three uncalibrated images. It is based on a linear solution derived from six points seen in three views (HARTLEY & ZISSERMAN 2003, MAYER 2003) followed by a non-linear bundle adjustment over all common points. To handle the large number of high resolution images the computationally intensive RANSAC algorithm for robust outlier detection has been replaced by a faster evolutionary approach called Genetic Algorithm Sampling Consensus GASAC (RODEHORST & HELLWICH 2006). Note that the fundamental matrices derived from this tensor are not independent and have only 18 significant parameters in total. Let $\mathbf{H}_b$, $\mathbf{H}_h$ and $\mathbf{H}_v$ be the unknown 3×3 homographies between the original and rectified images. These primitive rectifying homographies can linearly be determined from the compatible fundamental matrices with 6 degrees of freedom left. A detailed derivation is given in (HEINRICHS & RODEHORST 2006).

3.3 *Imposing Geometric Constraints*

We recommend to calculate values of the remaining 6 degrees of freedom in the following order:

- Finding proper shearing values
- Determine a global scale value
- Finding right offset values

The shearing of images can be minimized by keeping two perpendicular vectors in the middle of the original image perpendicular in the rectified one. This results in quadratic equations which have two solutions. The result with the smaller absolute value is preferred. On the one hand, the global scale should preserve as much information as possible, but on the other hand produce small images for efficient computation. Therefore, we adjust the length of the diagonal line through $\tilde{b}$ to the original length in b. The

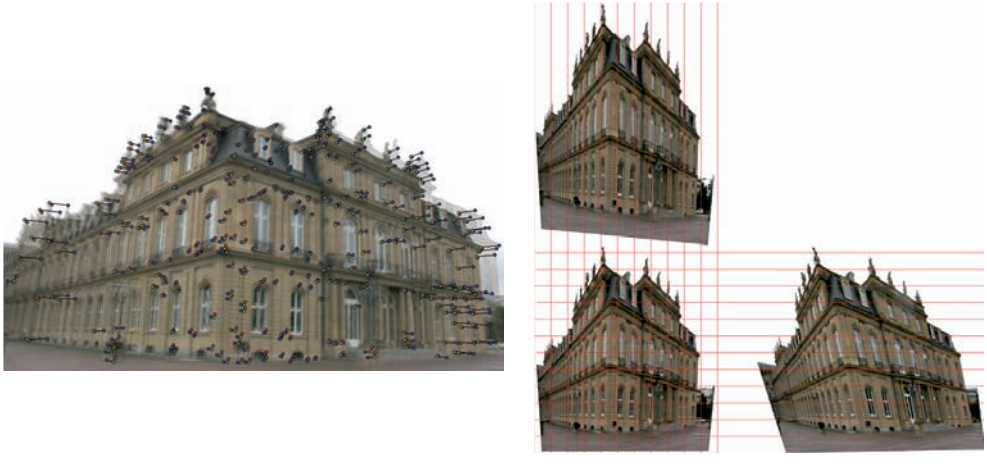

Fig. 3: a) Overlaid image triplet of Stuttgart palace with robust feature matches and b) rectified normal images.

offsets shift the image triplet in the image plane. To keep the absolute coordinate values small, the images should be shifted to the origin.

4 Trinocular Image Matching

After geometric transformation of the given image triplets using the proposed rectifying homographies, the correspondence problem must be solved using dense stereo matching. To find corresponding image points which arise from the same physical point in the scene, we suggest a modified semi-global matching (SGM) technique (Heinrichs et al. 2007, Hirschmüller 2005/2006). The normal images substantially simplify and accelerate the time-consuming computation. After rectification,

$$\begin{aligned} &\tilde{b}(x,y) \approx \tilde{h}(x + D(x,y), y) \quad \text{and} \\ &\tilde{b}(x,y) \approx \tilde{v}(x, y + D(x,y)) \end{aligned} \tag{1}$$

hold where x is the column coordinate, y the image row coordinate and D is called disparity map. The disparity at the current position (x, y) is inversely proportional to the depth of the scene. In addition to that, we assume that an estimation of the smallest and largest displacement is roughly given. This defines the search range $[d_{\min}, d_{\max}]$ for a reference point in $\tilde{b}$ along the corresponding rows and columns in $\tilde{h}$ and $\tilde{v}$ respectively.

4.1 *Local Similarity Measures*

Area-based matching is a widely used method for dense stereo correspondence. The similarity is computed statistically on the rectangular neighborhood (matching window) around the examined pixel. The algorithm searches at each pixel in reference image $\tilde{b}$ for maximum correlation in the horizontal image $\tilde{h}$ and the vertical image $\tilde{v}$ by shifting a small window pixel-by-pixel along the corresponding epipolar line. The normalized cross correlation NCC measures the linear relation between two image windows a and b normalizing over all intensity changes. The modified NCC (MNCC)

$$\rho_{MNCC}(a,b) = \frac{2 \cdot \sigma_{ab}}{\sigma_a^2 + \sigma_b^2} \tag{2}$$

handles homogeneous areas better by adding the two denominator variances instead of multiplying them (Egnal 2000). We precalculate the means and the means of squared intensities to accelerate the computation significantly. Finally, we transform the correlation coefficient range to $[0, 1]$. Due to the proposed trinocular rectification, the disparities in the horizontal and vertical image pairs are identical. Thus, the position of a matching candidate in image $\tilde{b}$ and $\tilde{h}$ is linked exactly to a position in $\tilde{v}$. Now, the local matching costs $1 - \rho$ can simply be averaged. This increments the computational effort for the additional third image only by a linear factor. Additionally, the matching is more robust, because the linked cost function has less local minima than the individual cost functions for the image pairs.

4.2 *Semi-Global Optimization*

Generally, matching based on local costs only is ambiguous and therefore a piecewise smoothness constraint must be added. In (Hirschmüller 2005/2006) a very simple and effective method of finding minimal matching costs is proposed. SGM tries to determine a disparity map D such that the energy function

$$\begin{aligned} E(D) = &\sum_{x,y \in I} ((1 - \rho(a(x,y), b(x + D(x,y), y))) \\ &+ Q_1 \sum_{i,j=-1}^{1} \mathrm{T}[|D(x,y) - D(x+i, y+j)| = 1] \\ &+ Q_2 \sum_{i,j=-1}^{1} \mathrm{T}[|D(x,y) - D(x+i, y+j)| > 1]) \\ &\text{for } i \neq j \end{aligned} \tag{3}$$

is minimal. The first term calculates the sum of all local matching costs using the inverse correlation coefficient ρ of the image windows a and b around the current position (x, y) and the related disparity in D. The subsequent terms require a Boolean function T that return 1 if the argument is true and 0 otherwise. Explained intuitively, $E(D)$ accumulates the local matching costs with a

small penalty $Q_1 = 0.05$, if the disparity varies by one from the neighbored disparities. If the disparity differs by more than 1, a high penalty $Q_2 \in [0.06, 0.8]$ is added. The actual value of Q_2 depends on the intensity gradient in the original image. Long gradients result in a low Q_2 while short gradients result in a high Q_2. This prevents depth changes in homogeneous regions. There are only two different penalties for the depth changes. First, Q_1 ensures that regions with a slightly changing depth are not penalized too hard. Second, if depth changes in the scene occur, the size of the discontinuity is not correlated to the penalty.

Computing the minimum energy of $E(D)$ leads to NP-hard complexity, which is difficult to solve efficiently. Following (Hirschmüller 2005), a linear approximation over possible disparity values $d \in [d_{\min}, d_{\max}]$ is suggested by summing the costs of several 1D-paths L through the search space towards the actual image location (x_1, y_1). A path L with $i = n \ldots 1$ steps is recursively defined and the number of accumulated paths should be at least eight. We introduce a threshold for the local matching costs, to penalize dissimilar candidates. If the correlation coefficient ρ is lower than a certain threshold the local matching costs is set to a high constant value. If the minimal costs for the best matching candidate is higher than this value the match is marked as invalid.

One disadvantage of SGM is the required space for all correlation values, which is needed to compute all non-horizontal trails L. The memory for this buffer is $O(n^3)$ depending on the image width, height and disparity search range. We save memory by reducing the length of the trails. Since the influence of previous L after a disparity discontinuity is very low, we need the complete path only for homogeneous areas. Except for trails along the epipolar lines, we limit the length of L to a small value (e. g. five). Therefore, the buffer size reduces to $O(n^2)$, which allows to process larger images.

An important issue for image matching is the stability of the found correspondence. If a correspondence is unstable, this is either an occlusion or the image significance is very low, e. g. in homogeneous regions or periodic patterns. To enforce stability, we check the left/right consistency (LRC) of the bidirectional correspondence search. A robust matching process should produce a unique result. On the one hand, LRC detects most stereo errors and depends not critically on thresholds. On the other hand, LRC does not report an error if the two matching directions mistakenly agree and one extra matching process is required. Nevertheless, the computational expense is tolerable for many applications. LRC leads to two disparity maps D_i, one for each image permutation. If the matched point in the second image points back to the original one in the first image the match is validated

$$D_1(x, y) + D_2(x + D_1(x, y), y) \leq 1. \qquad (4)$$

Otherwise it is invalidated or, in case of multi-image matching, other permutations of the disparity map must verify this match. In addition, using the reverse direction guaranties that all matched points are one-to-one correspondences, because doubly matched points can verify only one location.

4.3 *Hierachical Approach*

Based on the original image resolution a number of reduced images are computed using a scale factor f_i. The search range can be scaled as well by f_i, so that the computational complexity drops dramatically from the actual scale level to the next smaller one. The images are processed from the lowest resolution to the highest one. Only image points of the first layer have to be checked at every possible location within the search range. To reduce the number of candidates in the succeeding layers the potential information of the previous layer is used and refined. If displacement information from a previous layer is available, the number of candidates can be reduced by restricting the possible range. The valid candidates fulfill at least one of three criteria:

- **Accuracy improvement:** The information from the previous layer has an accuracy of $\pm s \cdot f_i$, where s is the distance from one

candidate to the next one and f_i is the scale factor from the previous layer to the actual one. Possible matches within this accuracy range must be checked.

- **Unmatched points:** Points in the target image, which are already matched in a previous layer, should be excluded from further matching to avoid double matches.
- **Edge preservation:** If points of the previous layer lie on a surface edge, the depth value of the associated points in the actual layer is bounded by the depths of the two neighboring points. It might happen that not all of this information is available.

Fig. 4 illustrates this technique. The diagonal strip represents the search space of the original layer. Every column is the search space for a pixel position. The red line represents the selected correspondence. The thin diagonal stripe around the red line is the accuracy improvement from the first criterion. Vertical lines are unmatched positions in the previous layer. Therefore, the search space at these positions has to be analyzed completely to find possible new matches.

Horizontal lines represent unmatched points from the second criterion. The small vertical strips are caused by the edge preservation of the third criterion. Candidates in the black area are excluded by the hierarchical approach. This shows the efficiency of the proposed method. The search space is reduced to approximately 25 % of the original size. After calculating the local costs for each candidate our modified version of SGM calculates the best match for the given position.

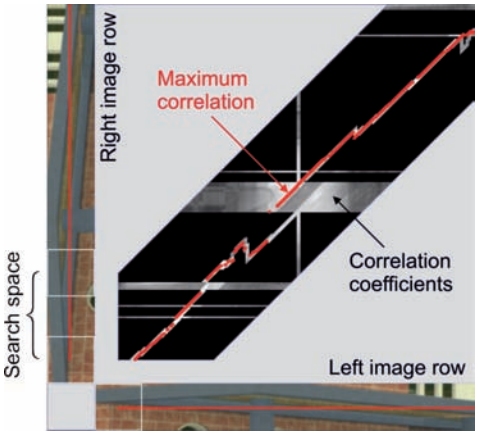

Fig. 4: Sample of the reduced search space of corresponding image rows.

5 Camera Path Estimation and 3D Point Cloud Registration

The temporal image correspondences $\mathbf{x} \leftrightarrow \mathbf{x}'$ are determined using the KLT tracker (Shi & Tomasi 1994). The feature-based approach uses local similarity measures and the temporal epipolar geometry. With a hierarchical approach using image pyramids the estimation of the orientation on a coarse level allows to improve the matching on a finer level. We extend the approach by filtering outliers using a temporal trifocal geometry (see Section 3.2). After the matching process we are able to orient the images fully automatically. Furthermore, the minimal 5-point algorithm (Nistér 2004) computes the essential matrix $\mathbf{E}$ of a camera pair even from correspondences on critical surfaces, i. e. planes. However, the presence of false correspondences in the tracking data and the unstable computation of eigen vectors (Batra et al. 2007) require a robust computation of the essential matrix via GASAC. This procedure results in a set of succeeding camera pairs with a uniform base length.

The registration of temporal image pairs using the estimated camera path is realized following (Fitzgibbon & Zisserman 1998). We are given temporal image pairs, each with an estimated essential matrix $\mathbf{E}$ and a set of homologous image points $\mathbf{x} \leftrightarrow \mathbf{x}'$. The goal is to register the spatial reconstructions into the same coordinate system by determining a spatial homography $\mathbf{H}$ which results in the best overlap of the two reconstructions. Since a spatial homography has 15 degrees of freedom and a projection matrix only 11, two corresponding cameras $\mathbf{P}$ and $\mathbf{P}'$ for a common image can be exactly registered by $\mathbf{P} = \mathbf{P}'\mathbf{H}^{-1}$. This constrains eleven parameters of $\mathbf{H}$. The remaining four

parameters can be found by minimizing the algebraic distance of the triangulated object points $d(\mathbf{X}, \mathbf{HX'})$ subject to the constraint $\mathbf{PH} = \mathbf{P'}$. The solution is a member of the 4-parameter family of homographies

$$\mathbf{H(v)} = \mathbf{P}^{+}\mathbf{P'} + \mathbf{hv}^{\mathsf{T}} \tag{5}$$

where $\mathbf{h}$ is the nullvector and $\mathbf{P}^{+}$ the pseudoinverse of $\mathbf{P}$. A direct solution for $\mathbf{v}$ can be obtained using a system of 3 equations per object point:

$$\mathbf{bX'}^{\mathsf{T}}\mathbf{v} = \mathbf{c} \tag{6}$$

where the 3-vectors $\mathbf{b}$ and $\mathbf{c}$ are defined by $b_k = h_k X_4 - h_4 X_k, \quad c_k = X_k a_4 - X_4 a_k$ and $\mathbf{a} = \mathbf{P}^{+}\mathbf{P'X'}$. The direct solution minimizes an algebraic error with no direct geometric or statistical meaning, so we refine the reprojection error using bundle adjustment.

6 Experimental Results

First results of the proposed method are presented for the Stuttgart palace in Germany. Figs. 5 and 6 show that it is possible to compute a realistic 3D architecture model utilizing high resolution video sequences. For

Fig. 5: a) The resulting disparity map of the Stuttgart palace using the modified SGM and b) top-view on the point cloud of the reconstructed corner.

Fig. 6: Available architecture model of the Stuttgart palace from aerial images improved by manual texture mapping (top) and 3D point clouds acquired automatically with the proposed trifocal sensor system using one image triplet (bottom).

these experiments, we use partially calibrated image triplets where radial distortion has been eliminated in advance. The hierachical approach reduces the computational costs to 25 percent without significant loss in accuracy. The dense matching of one image triplet with 1.4 mega pixels was completed in 2 minutes on a 2.4 GHz dual core CPU.

7 Conclusions and Future Work

In this paper, we introduce a high resolution video sensor for the 3D reconstruction of architectural models. The hybrid system unifies triangulation methods of spatial stereo with tracking methods of temporal stereo. We presented a linear method for trinocular rectification of uncalibrated images, which can be solved in closed form with 6 degrees of freedom. In a post processing stage, proper geometric constraints are selected to minimize projective distortion. The proposed dense matching algorithm is a fast and effective adaption of the SGM for image triplets. Combining the local costs of an image triplet to a single value stabilizes the matching, especially in regions with repetitive patterns like bricks, grids or stripes. The motion of the video sensor is estimated using temporal feature tracking and allows the integration of dense point clouds. First experimental results for image sequences of a real scene demonstrate the potential performance. The resulting photogrammetric models are combined as a metric model of the scene. The approach proposed is generic from a methodological point of view and allows various applications. Nevertheless, architectural models consist of planes, polyhedrons and freely formed surfaces. At this occasion geometric primitives, like planes and polyhedrons, should be fitted to the large point cloud to increase the accuracy and spare memory. Such systematic generation of a photogrammetric model from several trifocal views gives a digital video system its full potential.

Acknowledgements

This work was partially supported by grants from the German Research Foundation DFG. We would like to thank Hongwei Zheng and Nicole Bouvier for the textured model of the palace as well as Dirk Mehren for the acquisition of the trifocal images in Stuttgart.

References

Akbarzadeh, A., Frahm, J.-M., Mordohai, P., Clipp, B., Engels, C., Gallup, D., Merrell, P., Phelps, M., Sinha, S., Talton, B., Wang, L., Yang, Q., Stewenius, H., Yang, R., Welch, G., Towles, H., Nistér, D. & Pollefeys, M., 2006: Towards Urban 3D Reconstruction From Video. – 3rd Int. Symp. on 3D Data Processing, Visualization and Transmission (3DPVT).

Batra, D., Nabbe, B. & Hebert, M., 2007: An Alternative Formulation for Five Point Relative Pose Problem. – IEEE Workshop on Motion and Video Computing: 21–26.

Egnal, G., 2000: Mutual Information as a Stereo Correspondence Measure. – Computer and Information Science MS-CIS-00-20, University of Pennsylvania, PA, USA.

Faugeras, O.D. & Luong, Q.-T., 2001: The geometry of multiple images. – The MIT Press, Cambridge, Massachusetts.

Fitzgibbon, A.W. & Zisserman, A., 1998: Automatic Camera Recovery for Closed or Open Image Sequences. – European Conference on Computer Vision: 311–326.

Hartley, R. & Zisserman, A., 2003: Multiple view geometry in computer vision. – Cambridge University Press, 2nd edition.

Heinrichs, M. & Rodehorst, V., 2006: Trinocular Rectification for Various Camera Setups. – Photogrammetric Computer Vision PCV'06, Bonn, 43–48.

Heinrichs, M., Hellwich, O. & Rodehorst, V., 2007: Efficient Semi-Global Matching for Trinocular Stereo. – Photogrammetric Image Analysis PIA'07, Munich, 185–190.

Hirschmüller, H., 2005: Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. – Computer Vision and Pattern Recognition **2:** 807–814.

Hirschmüller, H., 2006: Stereo Vision in Structured Environments by Consistent Semi-Global Matching. – Computer Vision and Pattern Recognition **2:** 2386–2393.

Koch, R., 2003: 3D-scene modeling from image sequences. – Photogrammetric Image Analysis PIA'03: 3–9.

Lowe, D., 1991: Fitting parameterized three-dimensional models to images. – IEEE Transactions on Pattern Analysis and Machine Intelligence **13**(5): 441–450.

Lowe, D., 2004: Distinctive image features from scale invariant keypoints. – International Journal of Computer Vision **60**(2): 91–110.

Mayer, H., 2003: Robust Orientation, Calibration, and Disparity Estimation of Image Triplets. – Pattern Recognition – DAGM'03, Springer: 281–288.

Montemerlo, M. 2003: FastSLAM – A Factored Solution to the Simultaneous Localization and Mapping Problem with Unknown Data Association. – PhD thesis, CMU-RI-TR-03-28, Robotics Institute, Carnegie Mellon University.

Neumann, J. & Aloimonos, Y., 2002: Spatio-Temporal Stereo using multi-resolution subdivision surfaces. – International Journal of Computer Vision **47**(1/2/3): 181–193.

Nistér, D., 2004: An efficient solution to the five-point relative pose problem. – IEEE Transactions on Pattern Analysis and Machine Intelligence **26**(6): 756–777.

Nistér, D., Naroditsky, O. & Bergen, J., 2006: Visual odometry for ground vehicle applications. – Journal of Field Robotics **23**(1): 3–20.

Pollefeys, M., Gool, L.V., Vergauwen, M., Verbiest, F., Cornelis, K., Tops, J. & Koch, R., 2004: Visual modeling with a hand-held camera. – International Journal of Computer Vision **59**(3): 207–232.

Rodehorst, V. & Hellwich, O., 2006: Genetic Algorithm SAmple Consensus (GASAC) – A Parallel Strategy for Robust Parameter Estimation. – Int. Workshop "25 Years of RANSAC" in conjunction with CVPR'06, New York.

Rodehorst, V. & Koschan, A., 2006: Comparison and Evaluation of Feature Point Detectors. – Proc. of 5th Turkish-German Joint Geodetic Days, Berlin.

Shi, J. & Tomasi, C., 1994: Good features to track. – Computer Vision and Pattern Recognition: 593–600.

Tao, C.V., 2000: Semi-Automated object measurement using multiple-image matching from mobile mapping image sequences. – Photogrammetric engineering and remote sensing **66**(12): 1477–1485.

Zheng, H., Rodehorst, V., Heinrichs, M. & Hellwich, O., 2007: Improvement of the Fidelity of 3D Architecture Modeling Combining 3D Vector Data and Uncalibrated Image Sequences, ISPRS Workshop on Updating Geospatial Databases with Imagery and on DMGISs, Urumqi, 127–134.

Address of the Authors:

Dipl.-Ing. Matthias Heinrichs, Prof. Dr.-Ing. Olaf Hellwich, Dr.-Ing. Volker Rodehorst, Computer Vision & Remote Sensing, Berlin University of Technology, Franklinstr. 28/29, Sekr. FR 3-1, D-10587 Berlin, e-mail: matzeh@cs.tu-berlin.de, hellwich@cs.tu-berlin.de, vr@cs.tu-berlin.de

Manuskript eingereicht: Mai 2007
Angenommen: Juni 2007

Photogrammetrie • Fernerkundung • Geoinformation 6/2007, S. 415–427, 3 figs., 6 tabs.

A Semantic Model of Stairs in Building Collars

JÖRG SCHMITTWILKEN, JENS SAATKAMP & WOLFGANG FÖRSTNER, Bonn; THOMAS H. KOLBE, Berlin; LUTZ PLÜMER, Bonn

Keywords: 3D City Models, Stairs, Ontology, Stochastic Attribute Grammars, Constraint Solver

Summary: The automated extraction of high resolution 3D building models from imagery and laser scanner data requires strong models for all features which are observable at a large scale. In this paper we give a semantic model of stairs. They play a prominent role in the transition from buildings to the surrounding terrain or infrastructure. We name the transition area between terrain and building collar, and the focus is on stairs in building collars. Simple and complex stairways are represented by UML class diagrams along with constraints reflecting semantic and functional aspects in OCL. A systematic derivation of an attribute grammar consisting of production and semantic rules from UML/OCL is presented. Finally, we show how hypotheses with comprehensive predictions may be derived from observations using mixed integer/real programming driven by grammar rules.

Zusammenfassung: *Ein semantisches Modell für Treppen in Gebäudekragen*. Die automatisierte Extraktion hoch aufgelöster 3D-Gebäudemodelle aus Bildern und Laserscanner-Daten erfordert starke Modelle für alle Objekte, die im großmaßstäbigen Bereich beobachtbar sind. In diesem Beitrag konzentrieren wir uns auf die Entwicklung eines semantischen Modells für Treppen. Sie spielen eine wichtige Rolle beim Übergang vom Gebäude zum umgebenden Gelände und der Infrastruktur. Diesen Übergangsbereich von Gelände zu Gebäude bezeichnen wir als Gebäudekragen und betrachten Treppen in eben diesem Gebäudekragen. Einfache und komplexe Treppen werden durch UML-Klassendiagramme repräsentiert, wobei semantische und funktionale Aspekte in OCL ausgedrückt werden. Es wird die systematische Ableitung einer attributierten Grammatik mit ihren Produktionsregeln und semantischen Regeln aus UML/OCL vorgestellt. Schließlich wird gezeigt, wie aus Beobachtungen Hypothesen mit umfassenden Prädiktionen unter Verwendung von Methoden des Mixed Integer/Real Programming und gesteuert durch Grammatikregeln konstruiert werden.

1 Introduction

3D building models have been developed since more than a decade, now being readily available for everybody through the Internet. Google Earth, Microsoft's Virtual Earth and NASA World Wind provide high visibility. As a rule, however, existing building models have a low level of detail. Buildings are represented as boxes without roofs and substructures. File formats, such as KML (Key-hole Mark-up Language) used by Google Earth, are poor. Neither topology nor semantics are available. Advanced applications such as planning, disaster management, escape route findings and integrity preserving transactions require a higher level of detail and an explicit representation of semantics in a formal language.

Our research focuses on the design of models, methods and tools for the semi-automatic, interactive refinement of 3D city models. Starting from models consisting of simple blocks and roof structures, the aim

1432-8364/07/2007/0415 $ 3.25
© 2007 DGPF / E. Schweizerbart'sche Verlagsbuchhandlung, D-70176 Stuttgart

is to identify and reconstruct stairs, balconies, windows, doors, and arcades from terrestrial images or laser scans. It is rather obvious to start with entrance stairs. Due to their regular, recursive structure, high-resolution and detailed geometry they may be specified by a handful of parameters. Their semantic relevance is high. An entrance stair connects a route with the entrance door, a building with the outside world. Its detection and reconstruction helps to identify the entrance door and the ground floor level of a building which is not directly observable.

Building models and digital terrain models (DTM) often stem from different sources. This leads to problems of interoperability and homogeneity. A *ground edge*, if given, helps to glue DTM and building together. If there is no exact match, the fiction of a ground edge helps to find an optimal graduation. On a low level of detail a ground edge may be presupposed, although, due to occlusion, it may not be visible. On a higher level of detail a ground edge may never exist. Instead, there could be a transition zone with intermediate objects such as arcades, light wells etc. We call this transition area the *building collar* (Schmittwilken et al. 2006).

The aim of this paper is to provide a semantic model of stairs as part of a building collar especially exposing the path from the semantic model to a model usable for data interpretation. It should enable strong hypotheses and predictions based on a small number of noisy and incomplete observations (3D blobs from images or laser range data).

After introducing the main concepts in Section 2 our starting point is the ontology of stairs and a formal specification of object classes, generalization, aggregation and constraint rules defined by the Unified Modelling Language UML and the Object Constraint Language OCL (cf. Section 3). The UML-OCL model allows deciding whether or not a given object is a valid stair but it can't be used to generate a valid stair. Just for interpreting image data using hypotheses-test mode a generative model is needed. For this an attributed context-free grammar is introduced in Section 4. By translating aggregation into (right linear) recursion and OCL rules into semantic rules we show how the UML-OCL model is mapped to an attributed grammar. Whereas a grammar is able to generate all valid sentences of a language, it also may identify the syntactical and – if attributed – semantical structure of a given sentence. The latter is called parsing. Here the problem, however, is not to parse a complete sentence with fixed symbols but to generate reasonable hypotheses from incomplete, noisy observations. In Section 5 we illustrate how hypotheses with comprehensive predictions are derived from observations using thresholds as a bound on distributions from a database of real stairs and mixed integer/real programming. In Section 6 the paper finishes with a discussion on the genericity of the model and future work.

2 Ontology and Grammars

Semantics deals with the meaning of words. Ontology deals with the essence of things. Having its roots in philosophy and metaphysics, Plato and Aristotle, Hegel and Kant, Husserl and Heidegger ('Sein und Zeit', 1967) are some of the most prominent thinkers having meditated on this topic. Philosophers put questions like: 'What is the essence of the world?' and 'What characterizes an object?' and studying their approach helps us to make the semantics of objects explicit.

Computer science deals with *ontologies* (in plural). Ontologies make semantics explicit in formal symbolic representations. In other words, an ontology is 'a description of the concepts and relationships that can exist for an agent or a community of agents', generally written 'as a set of definitions of formal vocabulary' (Gruber 1995). They help to design software providing reasonable answers to sensible questions e. g. they are part of the *Semantic Web*. Ontologies can be used to identify meaningful patterns in noisy observations. Ontologies have roots in Artificial Intelligence (knowledge representation, logic, deduction and automated reasoning), database design (semantic data

models) and software engineering (object oriented modelling). Nowadays, there are many ontologies (Staab 2004), description languages (OWL, Antoniou 2004), and platforms (Protégé, Gennari et al. 2003). We use UML and OCL, the *Unified Modelling Language* together with its *Object Constraint Language* (Booch et al. 2005, OMG 2007). The reason is pragmatic: We have developed a large database for approx. 5.000.000 buildings of the German capital city Berlin and the German state North Rhine-Westphalia which is based on the semantic data model of CityGML (Groeger et al. 2006). CityGML is specified in UML and our model of stairs uses and extends CityGML.

UML is a formal specification and graphical notation which supports the definition of classes, attributes and associations (relations), namely aggregation (building of composites from parts) and inheritance (defining subclasses which inherit attributes and methods from the super class refining the latter by additional attributes and constraints). Generalization/specialization allows defining that the general concept of a stair includes linear stairs having a fixed number of steps, composite stairs with one or more landing in between, branching stairs and many other variants. Aggregation describes the relation between a stair and its steps. In order to claim that all steps of a stair have the same shape OCL has to be used. OCL is a subset of first-order predicate logic including predicates, logical operators such as *AND*, *OR*, *NOT*, *IMPLIES* and existential and universal (∃ and ∀) quantifiers (OMG 2006).

Generative grammars have been introduced by Chomsky (1956, 1959) in order to reconstruct the syntax of sentences in formal or natural languages. *Context-free* (type-2) languages have a special importance. A context-free language is given by a start symbol S, a set of non-terminals denoted by uppercase letters, a set of terminals denoted by lowercase letters and a set of production rules. Production rules have the form $A \rightarrow a$ where A is a non-terminal and a a sequence of terminals and/or non-terminals. It reads: Each occurrence of the symbol A may be substituted by the string a.

Stairs are taken as an example and denoted as follows: stair by S, riser by R and tread by T. The start symbol S, the set of non-terminals $N = \{S,R,T\}$, the set of terminals $T = \{r, t\}$ and the following set of production rules are given:

$$P = \{\ S \rightarrow R\ T \quad S \rightarrow R\ T\ S \quad R \rightarrow r \quad T \rightarrow t\}$$

A typical sentence of this language would be '$rt\ rt\ rt\ rt\ rt = (rt)^5$' denoting a stair with five steps ('rt's).

Context-free grammars define context-free languages. They are expressive enough to define regular, recursive structures such as $\{a^n b^n | n > 0\}$. Where a^n stands for n occurrences of a. It is generated by the rules $A \rightarrow ab$ and $A \rightarrow aAb$. The string $a^n b^n c^n$, however cannot be generated by a context-free grammar. An informal reason is 'the number of occurrences of a and c is the same' is context-sensitive and not context-free. This deals with remote substrings. A technical argument using *pumping lemma* is given by Hopcroft et al. (2001). An important consequence for us is the statement 'all steps of the same stair have the same size' is not context-free.

Knuth (1968, 1971) proposed the concept of *attribute grammars* which extend context-free grammars to remedy these deficiencies. Terminals and non-terminals are augmented by *attributes* (variables) and production rules by *semantic rules* which specify constraints between attributes. In our example semantic attributes could be `height` and `depth`. Semantic rules for the production $S \rightarrow RTS'$ could be `S.depth=R.depth` and `S.depth=S'.depth`. Note that an apostrophe is used to differentiate between occurrences of the same symbol. A more elaborated attribute grammar for stairs will be given in Section 4.

Attribute grammars are applied in compiler design (Aho et al. 1985). Attributes are used to link remote occurrences of the same symbol, for instance the declaration and instantiation of a variable. Semantic rules are

used to generate look-up tables for variable declaration and the generation of fragments of the target code language. Attribute grammars differentiate between *synthesized* and *inherited* attributes in order to specify the flow of information (from right to left or from left to right in a production rule) and the evaluation strategy in a derivation tree. As can be seen later, in our case the flow of information cannot and should not be fixed in advance. On the one hand observation resp. estimation of parameters leads to parameters of a stair, but on the other hand these stair parameters give prediction of unobserved steps and their parameters. Parsing of sentences of a formal language and classification and object reconstruction in noisy images are as a rule rather different. Parsing recognizes the syntactic structure of complete sentences with safely identified non-terminals. Our aim is to deduce reasonable, comprehensive hypotheses and predictions from incomplete observation in noisy images. Attribute grammars and their variants have been used in computer vision and pattern recognition since more than two decades (FU 1982). More recent work on their application in building reconstruction is reported in BRAUN et al. (1995) and STEINHAGE (1999).

Probabilistic (or stochastic) grammars are used to control the generation of words by known a priori probabilities. A stochastic grammar is a grammar where the production rules are associated with probabilities. The probability of a whole derivation equals the product of the probabilities of the single productions of that derivation. The probabilities of all productions $A \rightarrow a$ starting with the same non-terminal A sum up to 1 (FU 1982).

In shape grammars (STINY & GIPS 1972) terminals and non-terminals are geometrical patterns. They have been used to create designs with symmetric, recursive patterns. Shape grammars are appealing because shaped production rules address the intuition of the designer.

For our purpose the main disadvantage of shape grammars is that regularities just have geometrical but no formal symbolic representations. It is more a sketch than a specification. Dependencies and constraints remain implicit. Therefore we prefer attribute (probabilistic) grammars which are as expressive as shape grammars but represent constraints explicitly. In our grammar shapes are represented by attributes and constraints on them. This will be shown in Section 4. Attribute grammars nicely fit to sophisticated constraint solvers (HOOKER 2006).

Research on object modelling using stochastic attribute grammars has been done by MÜLLER et al. (2006) and PARISH & MÜLLER (2001). They use L-Systems, set grammars, and split grammars to generate virtual scenes. However they do not deal with reconstruction from images or 3D point clouds. WONKA et al. (2003) also generate virtual models using split grammars. BRENNER & RIPPERDA (2006) and RIPPERDA & BRENNER (2006) use rjMCMC techniques to control the generation of the derivation tree. MAYER & HUANG (in this issue) use L-Systems and MCMC techniques to generate 3D hypotheses for trees.

3 Ontology of Stairs

Stairs are designed to allow pedestrians to surmount the altitude difference of several levels. Contrary to ladders on one hand and ramps on the other hand stairs combine the comfort of walking with saving of space. Stairs are regularly formed from steps having the same rise and depth. Sometimes they are discontinued by one or more landings. The rise of a riser and depth of a tread are adapted to the step length of adult humans as BLONDEL already assumed in 1698. The vocabulary of stairs and parts of the taxonomy are defined in ISO 3881 (ISO 1977) to which the following specification adheres.

It is a general observation in ontological engineering that there are several ontologies for the same class of objects. An ontology refers to a domain of discourse, a cognitive interest and a greater context. It is difficult in general to match different ontologies for the same entities, but this discussion is beyond the scope of this paper. FRANK (1996,

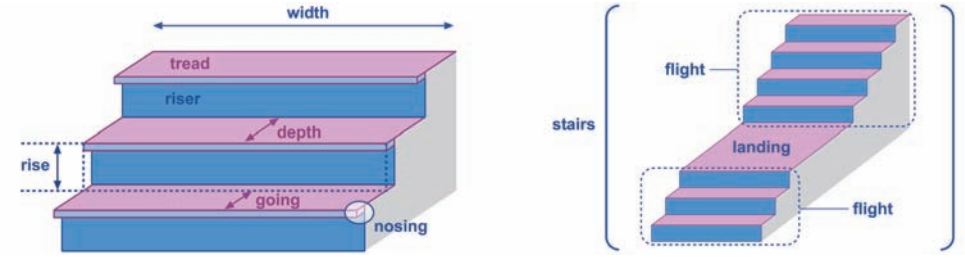

Fig. 1: Basic terms of stairs defined by ISO 3881.

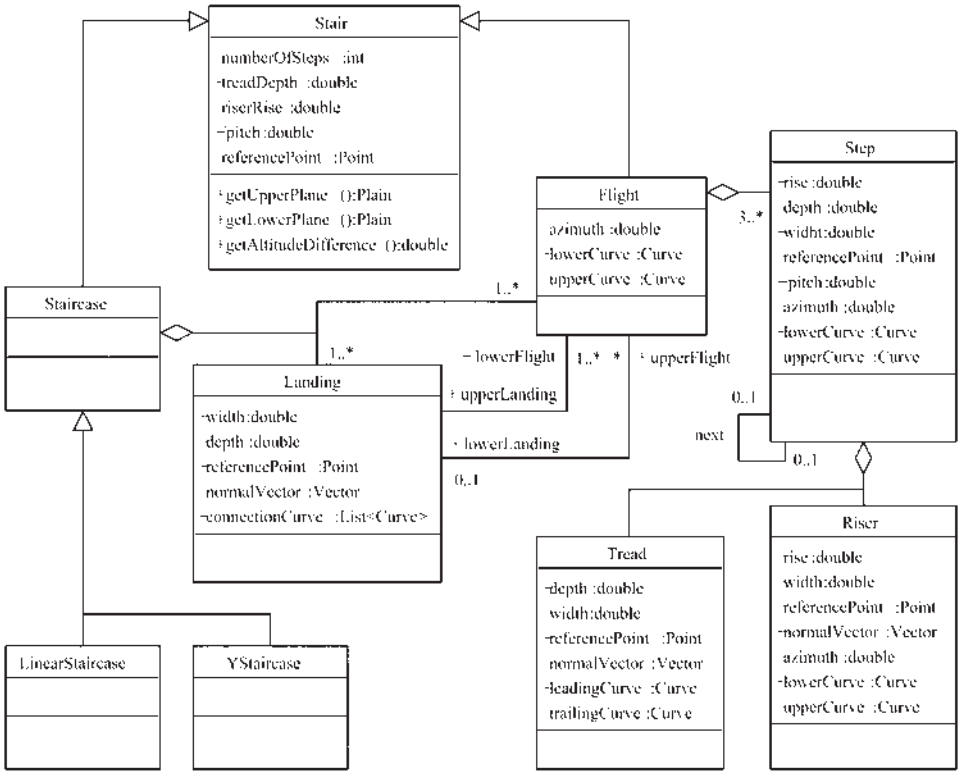

Fig. 2: UML diagram of stairs as one part of an ontology.

2003) and KUHN (2005) advice to identify the actions to which the ontology refers. For us two perspectives and two actions are relevant: The pedestrian has to cover an altitude difference. To this end he uses a stair. His action is stepping. A completely different perspective starts with terrestrial images or laser scans and ends up with the reconstruction of a stair. The action is classification and object reconstruction. Both kinds of action include that the visible surface plays an important role – walking on the surface on one hand and its observation and reconstruction on the other hand. The visible surface is associated with meaningful objects. Stairs are specified in ISO 3881 (cf. Fig. 1). In this norm the basic terms of stairs are standardized, their meaning is defined, and their translation into German and French is fixed. ISO 3881 provides a thesaurus.

The UML diagram shown in Fig. 2 provides a semantic model of stairs. It gives a general overview of stairs, its parts and their dependencies. On the top level there is the abstract class `Stair (S)` which is specialised into the classes `Staircase` and `Flight (F)`. Since attributes should be observable `pitch` is marked as a *derived* attribute. The class `Step` is aggregated by one `Tread` and one `Riser`. As mentioned above tread and riser are the basic components of a stair. All `reference-Point` attributes are dependent on each other due to the recursive and regular structure of stairs. The step class also has an association `next` with itself. So each step points to its successor (in upward direction). Three or more steps aggregate to a stair. Here we follow the German norm DIN 18065 (DIN 2000). Single and double steps have rather different patterns. A flight is exclusively made of steps. A staircase is an aggregation of flights and landings. Two associations express which flights are upstairs from the landing (optional) and which are downstairs (at least one flight). A landing has to have at least one lower flight and optional upper flights. Furthermore `LinearStaircase (L)` and `YStaircase (Y)` are special staircases and hence special stairs. A linear staircase is introduced as a staircase with one lower flight and one upper flight per landing and a Y-staircase as an aggregation of a landing with two lower flights and one upper flight or vice versa. In the following it is assumed that each flight has a constant width and a straight ground plan.

Whereas classes, attributes association, aggregation and inheritance structures are represented in the UML diagram as described above, essential regularities which characterize the essence of stairs are not given yet. They are specified by object constraint rules. Some basic assumptions and observations on stairs in general and thus on entrance stairs are the following ones, given in natural language. The corresponding lines of the transcribed OCL expressions shown in Tab. 1 are given in brackets.

1. All steps of a stair have the same rise and tread. This holds both for linear staircases with or without landing and for Y-staircases (lines 4–5).
2. Subsequent steps are connected seamlessly, i. e. the succeeding step follows the preceding step immediately (line 8).
3. The altitude difference between the two ends of a stair is given by the number of steps times the rise (line 9).
4. Tread and rise specify a vertical angle called pitch (line 3). There is another horizontal angle, the azimuth (line 3), which is the same for the steps of a flight (line 6), which may or may not be different for flights connected by a landing and which is different for y-staircases (lines 10–12). In all cases, there is a defined relationship between the azimuth and the orientation of the front façade of the corresponding house.
5. Since stairs are for human adults, depth and rise both have a given range. On the basis of a database containing parameters

Tab. 1: OCL rules expand the UML diagram with constraints.

```
01 context Riser inv: rise ~ N(17.0, (1.2)^2)
02 context Tread inv: depth ~ N(30.8, (2.7)^2)
03 context Step inv:  pitch = riser.rise / tread.depth & azimuth = riser.normalVector
04 context Flight inv: step → ∀ s1, s2 : s1.tread.depth = s2.tread.depth &
05                     step → ∀ s1, s2 : s1.riser.rise = s2.riser.rise &
06                     step → ∀ s1, s2 : s1.azimuth = s2.azimuth &
07                     step → ∀ s : (2 * s.rise + s.tread) ~ N(62.6, (3.0)^2) &
08                     step → ∀ s : s.upperCurve = s.next.lowerCurve
09 context Staircase inv: getAltitudeDifference() = sum(flight.riser.rise) * flight.riser->size() &
10                     flight → ∀ f1, f2 : abs(f1.azimuth - f2.azimuth) = 0 or
11                                         abs(f1.azimuth - f2.azimuth) = π/2 or
12                                         abs(f1.azimuth - f2.azimuth) = π
```

Tab. 2: Parameter distribution on staircases.

	riser [cm]	**tread** [cm]	**r/t** []	**2r + t** [cm]
μ	17.0	30.8	0.6	62.6
σ	1.2	2.7	0.1	3.0
μ/σ []	0.07	0.09	0.13	0.05

of 120 entrance stairs, one may very well (a) assume the rise, the tread depth, and the footstep of different stairs to be normally distributed (cf. lines 1, 2, and 9) and (b) observe the footstep formula 2r + d as the most stable parameter for stairs (line 7). Average values and standard deviations for the observed parameters are given in Tab. 2.

The OCL rules are given in Tab. 1. For the sake of readability the notation is slightly modified for this example. Each rule defines an invariant in the **context** of a specific class.

4 From Ontology to Grammar

Tab. 3 gives the production rules corresponding to the UML diagram shown in Fig. 2. The production rules are derived from the UML diagram by a method which consists of the following rules plus some technical details.

Each class name is associated with a non-terminal symbol. Non-terminals are given by the set {`S`, `L`, `Y`, `F`, `LANDING`, `STEP`, `RISER`, `TREAD`}. For the sake of brevity the non-terminals for stair (`S`), linear staircase (`L`), Y-staircase (`Y`), and flight (`F`) are abbreviated. The rest of them will drop out as can be seen in the following. All non-terminals have a one-to-one correspondence with the classes of the UML diagram.

The terminals are in the set {`riser`, `tread`, `landing`}. They correspond to – but are not identical with – features which are directly observable. The main point here is that the corresponding classes of the terminals are primitive in a specific sense: delete from the UML 'graph' all edges which are neither aggregation nor generalization. Give the generalization edges a direction from super class to subclass and the aggregation edges a direction from whole to part. If a super class defines an aggregated object, add the corresponding edges to their subclasses, too. The primitive classes are just those nodes of the resulting directed graph which have no off-going edges. Thus one gets the primitive classes `Landing`, `Tread` and `Riser`. These primitive classes form the set of terminal symbols. Note that `L` and `Y` are no leaves since they participate in the aggregation of `Staircase`.

As the three rules for stair illustrate, generalization is easy to handle: Translate generalization to a set of rules, where the super class non-terminal is on the left hand side and each subclass non-terminal gives a right hand side.

Aggregation is more difficult to handle, but there are two important patterns. The first case is where the whole (step) is composed from two or more different parts (tread and riser) the number of which is given in advance. Then the part non-terminals form the right hand side and the aggregate forms the left hand side of the corresponding rule. Second case: if there is a non-fixed number of parts, as it is the case with flight and steps, recursion is applied. As the case of flight shows, one linear *right recursive* rule

`(F → STEP F)`

and one non-recursive rule

`(F → STEP)`

suffice. A recursive production rule is called *linear recursive*, if there is only one occurrence of the left hand symbol on the right hand side of the production. It is called *right recursive*, if the recursive symbol is rightmost. This is an easy case of recursion since it forms the special case of a regular (Chomsky 3) grammar. The case of landing is similar.

The derived grammar makes the immediate correspondence to the UML diagram ex-

plicit but it has some redundancies which can easily be removed by rewriting. The three production rules for landing, tread, and riser just have one non-terminal symbol on the left and one terminal on the right hand side. When all occurrences of these non-terminal symbols in the remaining productions are replaced by the terminals, both the three production rules and the non-terminals can be cancelled.

There are three additional terminal symbols which correspond to no class, namely ';', '(' and ')'. These are just meta-symbols which are used in the case of Y-staircases in order to make a non-linear tree-like structure explicit. They support readability, but do not provide additional information.

Tab. 3: Excerpt of the grammar derived from UML diagram.

p1: **S** → **L** | **Y** | **F**
p2: **L** → **F** landing **L** | **F**
p3: **Y** → **F** landing (**F** ; **F**) | (**F** ; **F**) landing **F**
p4: **F** → riser tread **F** | riser tread

So far only classes, aggregation and generalization have been reflected. Other associations may be mapped as well provided that their multiplicity is given and their precise semantics are specified by constraint rules. This, however, is beyond the scope of this paper. What about the *attributes*? They become attributes of the respective terminals and non-terminals. Tab. 4 shows some examples of semantic rules. (Apostrophe is used to differentiate between multiple occurrences of the same symbol in one production. Non-scalar attributes are underlined.) Following this pattern the other semantic rules may easily be derived from the OCL rules. A comprehensive list of OCL rules (as well as detailed production rules, attributes and semantic rules) can be found at http://www.ikg.uni-bonn.de/fileadmin/data/schmittwilken_07_modeling_appendix.pdf. Although the correspondence between semantic rules and OCL rules is rather obvious, the translation is a bit tricky. Firstly, note that aggregation with identities on a non-fixed number of parts (flight and step) affords a quantifier on all steps `s1` and `s2` in the OCL rule. In the production rule, aggregation is represented by (right linear) recursion, and in the semantic rule, there are dependencies between the attributes of the two recursive occurrences of flight and the non-recursive non-terminal riser and tread. Secondly, inaccuracy is represented differently. OCL rules define constraints for the strict ontological model. For example they ensure that all steps have the same attributes and they provide parameters (mean μ and standard deviation σ) of the respective distributions. On the other hand there are the semantic rules of the attribute grammar. They are used in an operational scenario in order to guarantee a given functional context for each attribute e. g. the attributes of a specific flight equal the weighted mean of the respective attributes of its steps. The exact stochastic definition of the functions *weightedMean( )* and *f( )* is not precisely specified here.

Tab. 4: Excerpt of the semantic rules for the unknown attributes.

F → riser tread F'
F.numberOfSteps = F'.numberOfSteps + 1
F.rise = weightedMean(riser.rise, F'.rise)
F.azimuth = f(riser.normalVector, F'.azimuth)
tread.referencePoint = sum(riser.referencePoint, dx, dy)
... (*e. g. topological connection of riser and tread)*

Thresholds may be derived from the deviation (f.i. $\Delta = 3\sigma$) and used to define boolean-valued semantic rules called *guards*. Guards control the search and allow to prune the search space for candidate models early. Examples are given in Tab. 5 referring to a (simplified version of) production rule p2 in Tab. 3. *Guarded semantic rules* mimic *Guarded Horn Clauses* described in Ueda (1985). The *guard* of a production rule is the conjunction of all guarded semantic rules. An application of a production is *valid* if its guard evaluates true. A *derivation* is *valid* if all applications of production rules are valid, and a sentence is valid if it has a valid derivation. Formally, the valid sentences form a subset of the syntactically correct

Tab. 5: Guard rules for the unknown attributes.

L → F landing F'
| F.treadDepth – F'.treadDepth | < 8.02
| F.rise – F'.rise | < 3.57
| F.pitch – F'.pitch | < 0.22
...

sentences specified by the corresponding non-attribute grammar.

5 Hypotheses by Solving Mixed Linear Constraints

In this paragraph it is illustrated how the semantic rules and the guard rules may be used to derive hypotheses being consistent with given observations. To simplify illustrations we study a 2D projection of stairs. Generalization to 3D is obvious.

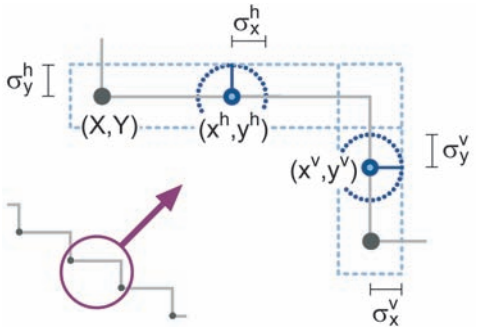

Fig. 3: Variables used for the reasoning example.

Assume that for the surfaces of stairs we have observed 2D points and their accuracies distinguished from points on horizontal and vertical faces (c.f. Fig. 3). Each point 'knows' his direction i. e. the normal vector of 'his' face, so we call this points *2D needles*.

$$x_i^h, y_i^h, \sigma_x^h, \sigma_y^h, x_i^v, y_i^v, \sigma_x^v, \sigma_y^v \tag{1}$$

The major construction principle of stairs is the constancy of rise and tread depth within a stair. So we claim

$$r = const \wedge t = const \tag{2}$$

Thresholds for these values can be derived from the stair database mentioned in section 3:

$$14 \leq r \leq 20 \wedge 18 \leq t \leq 38 \tag{3}$$

The number of steps within a stair is an integer number, i. e. the number of treads (n) and the number of risers (m):

$$n \in \mathbb{N} \wedge m \in \mathbb{N} \tag{4}$$

The structure of a 2D stair with origin (0,0) is defined by its rise and tread depth apart from its number of steps. So *reference points* can be introduced with coordinates resulting from a multiple of rise or tread depth:

$$Y_n = Y_0 + m*r \wedge X_n = X_0 + n*t \tag{5}$$

The coordinates of observed points supporting a given set of stair parameters can be specified by the following constraints. Inequalities of 'vertical points' can be derived analogously.

$$\begin{aligned} X_n - \sigma_x^h \leq x_i^h \leq X_n + \sigma_x^h + t \\ Y_n - \sigma_y^h \leq y_i^h \leq Y_n + \sigma_y^h \end{aligned} \tag{6}$$

We give the chain of reasoning which is used to determine the parameters of a stair from given observations. We assume that the observations are given within the stair coordinate system. Without loss of generality, the origin of the stair is claimed to be $(X_0, Y_0) = (0,0)$ and all accuracies are assumed to be 1.

Which kind of solver is necessary in order to support and implement this kind of reasoning?

Based on the ontology of stairs, constraints defined in OCL and the semantic rules of the attribute grammar turn out to become linear equalities and inequalities. The products $m*r$ and $n*t$ will disappear by constraint propagation and/or tentative instantiations of n and m. Thus a solver for linear inequalities (where equalities are a special case) is needed. The method of choice would be Simplex, but it does not suffice in our case. Simplex assumes real valued variables but there are integer variables as well (number of steps). Linear optimization on integers is computationally more demanding than optimization on real numbers. Optimization here becomes combinatorial and in fact NP-complete. More specifically, the given problem is mixed-integer.

Fortunately, there are algorithms and solvers for this kind of constraints (Hooker 2006).

We use the constraints solver ECLiPSe (Apt & Wallace, 2007) to estimate integer values or real intervals for the interdependent parameters. All constraints have been implemented. Tab. 6 shows an excerpt of the program: two clauses for estimating the reference point from an observed point lying on a horizontal face.

Tab. 6: Excerpt of the constraint solver program.

```
estimate_RefX_from_hPoint( ) :-
  x $> = RefX − Sigma_x_horizontal,
  x $= < RefX + Sigma_x_horizontal + Tread,
  RefX $= X0 + N * Tread.

estimate_RefY_from_hPoint( ) :-
  y $> = RefY − Sigma_y_vertical,
  y $= < RefY + Sigma_y_vertical,
  RefY $= Y0 + M * Riser.
```

Starting with the observation $p_1 = (22,0)$ we cannot reduce the intervals for riser and tread depth: $tread \in [18,38]$ and $riser \in [14,20]$. Obviously we can fix the coordinates of the origin as $(0,23)$. Insertion of the observation values into (7) we get

$$22 - 1 - [18..38] \leq X_1 \leq 22 + 1$$
$$0 - 1 \leq Y_1 \leq 0 + 1$$

and due to (2-5) we get $X_1 \in [0,23]$, and $Y_1 \in [0,1]$. This leads to the number of risers and treads (from origin to reference point): $n_1 = 0$ and $m_1 = 0$.

The second observation $p_2 = (140,75)$ restricts the domain but the parameters are still ambiguous: $tread \in [20.19,35.25]$ and $riser \in [18.25,19.0]$. So we estimate for the second reference point $X_2 \in [103.75,141.0]$, and $Y_2 \in [74.0,76.0]$ with $n_2 \in [4,5]$ and $m_2 = 4$. The intervals of the coordinates of the origin and the first reference point remain the same.

The third observation $p_3 = (71,26)$ will resolve the ambiguity: $tread \in [34.5,35.25]$ and $riser \in [18.25,19.0]$ with the following numbers of steps: $n_1 = 0$, $m_1 = 0$, $n_2 = 4$, $m_2 = 4$, $n_3 = 2$, and $m_3 = 2$.

If we get a fourth observation such as $p_4 = (291,115)$ we can even detect landings within the stair. The estimation of the fourth reference point and its numbers of steps from the origin leads to an inconsistency: $n_4 = 8$ and $m_4 = 6$. The difference between n_4 and m_4 can only be explained by a landing between p_4 and p_3.

6 Conclusion and Final Remarks

We have specified an ontology of stairs in the building collar based on 'common sense' and ISO 3881. The UML/OCL model presented in Section 3 and the attribute grammar of Section 4 specify a subset of the variety of stairs which is observed in reality. How to derive the UML diagram and the constraint rules from examples automatically remains an open question. We have sketched a procedure which maps the UML diagram and the constraint rules to an attribute grammar. With regard to associations the focus was on aggregation and generalization. The mapping of other associations depends on their multiplicity and their semantics specified by OCL rules.

One main contribution of this paper is the derivation of logical formulas (conjunctions of constraints) from the ontology of stairs linked by the (stochastic) attribute grammars. We provide a method to solve these constraints and generate candidate models. This method has been implemented as a prototype in ECLiPSe and is available on the internet at http://www.ikg.uni-bonn.de/fileadmin/data/schmittwilken_07_modeling_appendix.pdf. The candidate models fix numbers of integral domains and specify precise intervals for the real domains. The stochastic estimation of most probable values within these intervals with Bayes statistic (Russel & Norvig 2003, Bishop 2006, Pearl 2000) is left as a topic for future research.

The constraints solver generates candidate models from consistent observations. A set of observations is consistent if there exists an instantiation of the model parameters satisfying the constraints. These

parameters constitute a model explaining the observations. Under certain conditions three points (3D needles) suffice to specify a single candidate model. Thus the constraint solver is able to handle incomplete observations. In realistic scenarios, however there will be a large number of 3D points with substantial percentage of outliers. The reconstruction of stairs from noisy observations requires the differentiation between inliers and outliers. From a logical point of view inliers constitute the largest consistent subset of the given observations. Exact algorithms for the identification of this subset are not feasible. An obvious probabilistic approach to identify this subset is the RANSAC algorithm (Fischler & Bolles 1981) which nicely fits to our constraint solver and will b implemented in a next step. For a randomly chosen minimal set of observations the constraint solver derives one or more candidate models if this set of observations is consistent. In a next step RANSAC calculates the 'consensus set' of consistent observations which fits into the candidate model. The RANSAC algorithm iterates its steps and terminates when a sufficiently large set has been found.

Although we focus on entrance stairs of 'ordinary' buildings which are less complex than those to be seen in pleasure grounds, natural parks and in front of castles, some restrictions should be mentioned. We assume that stairs have constant width. In some cases the width decreases (e. g. parallelogram layout), and it is a linear function of the index of the step. We also assume that stairs are straight, branching (Y-stairs) or interrupted by landings. Some stairs have more sophisticated, curved layouts. Another assumption is that steps are compact solids. However some steps have nosing, others rest on piles (open stairs). Reconstruction and classification of the mentioned details is more demanding and requires an extension of the ontology and thus additional model parameters.

An important subclass of stairs is spiral stairs. They have limited relevance for building collars, fire escapes and indoor staircases, however, sometimes have this shape. Whereas riser and number of steps can be estimated by the constraints given in Section 5 the tread depth and the width of the stair are more sophisticated and require trigonometric functions in a Cartesian coordinate system. Linearity can be achieved in a polar coordinate system. But this affords that the centre point(s) of the spiral stair can be derived Curved transitions between two flights have a similar pattern. An effective system of constraints for this case is topic for future research.

The ontology and the attribute grammar are generic. Extensions for the cases mentioned above can be handled in this framework. They require additional parameters and more sophisticated constraints. Above, they require non linear constraint solvers. However there are algorithms (Boyd & Vandenberghe 2006) and systems like CPLEX (ILOG 2006), ECLiPSe (Apt & Wallace 2007), or Frontline solver (Frontline Systems, Inc., 2007) which are able to handle large numbers of non-linear constraints if they are convex. Even non-convex collections of constraints can be solved if they meet the requirement that the modelling of the independencies of the constraints supports constraint propagation. In Section 5 we have shown that the non-linearity of $n*tread$ disappears when the domain of n becomes a singleton due to constraint propagation. Thus we claim that the integration of strong constraint solvers with graphical models/Bayes networks will help to bridge the gap between ontological engineering and 3D object reconstruction.

References

Aho, A.V., Sethi, R. & Ullmann, J.D., 1985: Compilers. Principles, Techniques and Tools. – Addison-Wesley Longman, Boston, USA.

Antoniou, G. & van Harmelen, F., 2004: Web Ontology Language OWL. Handbook on Ontologies. – Springer, New York, USA: 68–92.

Apt, K.R. & Wallace, M., 2007: Constraint Logic Programming Using ECLiPSe. – Cambridge University Press, Cambridge, UK.

Bishop, C.M., 2006: Pattern Recognition and Machine Learning. – Springer, New York, USA.

Blondel, F., 1698: Cours D'Architecture. – L'Academie Royale D'Architecture, Paris, France.

Booch, G., Rumbaugh, J. & Jacobson, I., 2005: Unified Modeling Language User Guide. – Addison-Wesley Longman, Boston, USA.

Boyd, S. & Vandenberghe, L., 2006: Convex Optimization. – Cambridge University Press, Cambridge, UK.

Braun, C., Kolbe, T.H., Lang, F., Schickler, W., Steinhage, V., Cremers, A.B., Förstner, W. & Plümer, L., 1995: Models for Photogrammetric Building Reconstruction – Computers & Graphics **19** (1): 109–118.

Brenner, C. & Ripperda, N., 2006: Extraction of Facades using rjMCMC and Constraint Equations – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (3): 155–160.

Chomsky, N., 1956: Three Models for the Description of Language. – IEEE Transactions on Information Theory **2** (3): 113–124.

Chomsky, N., 1959: On Certain Formal Properties of Grammars. – Information and Control **2:** 137–167.

DIN (Deutsches Institut für Normung), 2000: DIN 18065 Gebäudetreppen – Definitionen, Meßregeln, Hauptmaße – Berlin, Germany.

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

Frank, A.U., 1996: Ontology: A Consumer's Point of View, Spatial and Temporal Reasoning. – Kluwer, Dordrecht, Netherlands.

Frank, A.U., 2003: Ontology for Spatio-temporal Databases. – In: Koubarakis, M., Sellis T.K., Frank, A.U. et al. (eds.): Spatio-Temporal Databases: The CHOROCHRONOS Approach. Lecture Notes in Computer Science 2520, Springer, Berlin, Germany: 9–78.

Frontline Systems, Inc., 2007: website. http://www.solver.com (last visited: 2007-06-20).

Fu, K.S., 1982: Syntactic Pattern Recognition and Application. – Prentice Hall, Englewood Cliffs, USA.

Gennari, J., Musen, M.A., Fergusson, R.W., Grosso, W.E., Crubezy, M., Eriksson, H., Noy, N.F. & Tu, S.W., 2003: The Evolution of Protégé: An Environment for Knowledge-Based Systems Development – International Journal of Human-Computer Studies, **58** (1), 89–123.

Gröger, G., Kolbe, T.H. & Czerwinski, A., 2006: Candidate OpenGIS CityGML Implementation Specification (City Geography Markup Language). – OGC Doc. No. 06–057r1.

Gruber, T., 1995: Toward Principles for the Design of Ontologies Used for Knowledge Sharing. – International Journal of Human-Computer Studies **43** (4–5): 907–928.

Heidegger, M., 1967: Sein und Zeit. – Niemeyer, Tübingen, Germany.

Hooker, J.N., 2006: Operations Research Methods in Constraint Programming. – In: Rossi, F., van Beek, P. & Walsh, T. (eds.): Handbook of Constraint Programming. Elsevier, Amsterdam, Netherlands: 527–570.

Hopcroft, J.E., Motwani, R. & Ullman, J.D., 2001: Introduction to Automata Theory, Languages, and Computation. – Addison-Wesley Longman, Amsterdam, Netherlands.

Huang, H. & Mayer, H., 2007: Extraction of the 3D Branching Structure of Unfoliaged Deciduous Trees from Image Sequences. – this issue.

ILOG, 2006: Efficient modelling in ILOG OPL-CPLEX Development System. – White Paper, http://www.ilog.com/products/oplstudio/whitepapers/index.cfm (last visited 2007-05-15).

ISO (International Organization for Standardization), 1977: Building construction – Stairs – Vocabulary – Part I. – ISO 3880/I-1977, First Edition, Geneva, Switzerland.

Knuth, D.E., 1968: Semantics of Context-Free Languages. – Theory of Computing Systems **2** (2): 127–145.

Knuth, D.E., 1971: Top-Down Syntax Analysis. – Acta Informatica **1** (2): 79–110.

Kuhn, W., 2005: Why, of What, and How? – In: Spaccapietra, S. & Zimányi, E. (eds.): Semantic-based Geographical Information Systems. Journal on Data Semantics **III**. Lecture Notes in Computer Science 3534, Springer, Berlin, Germany: 1–24.

Müller, P., Wonka, P., Haegler, S., Ulmer, A., & Gool, L. van, 2006: Procedural Modeling of Buildings – In: Proceedings of ACM SIGGRAPH 2006, New York, USA / ACM Transactions on Graphics **25**, ACM Press: 614–623.

OMG (Object Management Group), 2006: Object Constraint language (OCL) 2.0.

OMG (Object Management Group), 2007: Unified Modeling Language (UML) 2.1.1.

Parish, Y.I.H., & Müller, P., 2001: Procedural Modeling of Cities – In: Fiume, E. (ed.), Proceedings of ACM SIGGRAPH 2001, New York, USA, ACM Press: 301–308.

Pearl, J., 2000: Causality. Models, Reasoning and Inference. – Cambridge University Press, Cambridge, UK.

Russell, S.J., Norvig, P., 2003: Artificial Intelligence: A Modern Approach. – Prentice Hall, Pearson Education, Upper Saddle River, USA.

Ripperda, N., & Brenner, C., 2006: Reconstruction of facade structures using a formal grammar and rjMCMC – Lecture Notes in Computer Science **4174**: 750–759.

Schmittwilken, J., Kolbe, T.H. & Plümer, L., 2006: Der Gebäudekragen – Eine detaillierte Betrachtung des Übergangs von Gebäude und Gelände. – Proceedings of 26. Wissenschaftlich-Technische Jahrestagung der DGPF, September 2006, Berlin, Germany: 127–135.

Staab, S. & Studer, R., 2004: Handbook on Ontologies. – Springer, Berlin, Germany.

Stadler, A. & Kolbe, T.H., 2007: Spatio-Semantic Coherence in the Integration of 3D City Models. – In: Proceedings of the 5th International Symposium on Spatial Data Quality ISSDQ 2007, June 2007, Enschede, Netherlands, ISPRS Archives (in print).

Steinhage, V., 1999: Zur automatischen Gebäuderekonstruktion aus Luftbildern. – Habilitationsschrift an der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Bonn.

Stiny, G. & Gips, J., 1972: Shape Grammars and the Generative Specification of Painting and Sculpture. – Information Processing **71**. North-Holland, Amsterdam, Netherlands: 1460–1465.

Ueda, K., 1985: Guarded Horn Clauses. – In: Wada, E. (ed.): Logic Programming '85: Proceedings of the 4th Conference. Lecture Notes in Computer Science **221**, Springer, New York, USA: 168–179.

Wonka, P., Wimmer, M., Sillion, F., & Ribarsky, W., 2003: Instant Architecture – ACM Transactions on Graphics, **22**(4), 669–677.

Addresses of the Authors:

Dipl.-Ing. Jörg Schmittwilken, Prof. Dr. Lutz Plümer, Department of Geoinformation, Institute of Geodesy and Geoinformation, University of Bonn, Meckenheimer Allee 172, 53115 Bonn, Tel.: +49 228/73-{1756, 1751}, E-Mail: {schmittwilken, pluemer}@ikg.uni-bonn.de

Dipl.-Ing. Jens Saatkamp, Prof. Dr.-Ing. Wolfgang Förstner, Department of Photogrammetry, Institute of Geodesy and Geoinformation, University of Bonn, Nussallee 15, 53115 Bonn, Tel.: +49 228/73-{2711, 2713}, E-Mail: saatkamp@uni-bonn.de, wf@ipb.uni-bonn.de

Prof. Dr. Thomas H. Kolbe, Institute for Geodesy and Geoinformation Science, Technical University Berlin, Straße des 17. Juni 135, 10623 Berlin, Tel.: +49 30/314-23274, E-Mail: kolbe@igg.tu-berlin.de

Manuskript eingereicht: Mai 2007
Angenommen: Juni 2007

Photogrammetrie • Fernerkundung • Geoinformation 6/2007, S. 429–436, 5 figs.

Extraction of the 3D Branching Structure of Unfoliaged Deciduous Trees from Image Sequences

Hai Huang & Helmut Mayer, Neubiberg

Keywords: 3D Tree Extraction, Branching Structure of Trees, Statistical Generative Models, L-Systems, Markov Chain Monte Carlo – MCMC

Summary: In this paper we propose an approach for the three-dimensional (3D) extraction of the branching structure of unfoliaged deciduous trees from urban wide-baseline image sequences. The trees are generatively modeled in 3D by means of L-systems. A statistical approach, namely Markov Chain Monte Carlo (MCMC) is employed together with cross correlation for the extraction of the branches. With this generative statistical approach we avoid the complexity and uncertainty of extracting and matching branches in several images due to weak contrast, background clutter, and particularly the varying order of branches when projected into different images. First results show the potential of the approach.

Zusammenfassung: *Extraktion der 3D Verzweigungsstruktur unbelaubter Laubbäume aus Bildsequenzen.* Dieses Papier stellt einen Ansatz für die Extraktion der drei-dimensionalen (3D) Verzweigungsstruktur unbelaubter Laubbäume aus städtischen Bildsequenzen mit langer Basis vor. Die Bäume werden in 3D mittels L-Systemen modelliert. Markoff Ketten Monte Carlo (MCMC) wird zusammen mit Kreuzkorrelation für die Extraktion der Äste genutzt. Mit diesem generativen statistischen Ansatz wird die Komplexität und Unsicherheit der Extraktion und Zuordnung von Ästen in mehreren Bildern wegen schwachem Kontrast, Störobjekten im Hintergrund und insbesondere der z. T. unterschiedlichen Ordnung der Äste nach Projektion in verschiedene Bilder vermieden. Erste Ergebnisse zeigen das Potential des Ansatzes.

1 Introduction

In our environment trees play an essential role. This is particularly true in urban areas, where they are too often the only prominent representatives of nature. Because of their complex structure their acquisition is costly. Thus, they are often neglected, or at least only acquired in a very simplified form for geoinformation systems (GIS), especially for three-dimensional (3D) city models. The distinctive shape and texture of some of the trees that influence the appearance of their whole environment is only represented for very locally limited architectural models.

In this paper we aim at extracting the 3D branching structure of individual unfoliaged deciduous trees from wide-baseline image sequences. Deciduous trees are popular in cities worldwide as they provide shadow in summer and yet let through most of the light in winter. Thus, they often form the majority of trees in urban areas. From a practical point of view images for data acquisition in cities will often be taken when the trees are unfoliaged, as facades etc. are then more readily visible.

From a scientific point of view extracting the 3D branching structure of unfoliaged trees by matching in multiple images is a difficult problem nobody to our knowledge has tried to solve yet. Extraction and matching of branches is difficult because of bad contrast, clutter by background objects, and because the order of the branches even in neighboring images can vary considerably

1432-8364/07/2007/0429 $ 2.00
© 2007 DGPF / E. Schweizerbart'sche Verlagsbuchhandlung, D-70176 Stuttgart

due to the pronounced 3D structure of trees.

Former work has mostly dealt with tree extraction in aerial images and especially recently laser scanner data. Much work focuses on forests. Work for tree extraction from terrestrial urban images is scarcer. Haering et al. (1997) segment groups of foliaged deciduous trees in color images based on texture without any 3D interpretation. Also Forsyth et al. (1996) focus only on a two-dimensional (2D) interpretation, yet for individual trees. They particularly model the symmetries of coniferous trees. (Sakaguchi & Ohya 1999) is mostly dealing with the animation of trees, but is one of the few papers actually concerned with the 3D extraction of trees. A volume is carved out by intersecting the view cones generated from the tree silhouettes in multiple images. The voxels of the volume are colored with the average brightness of the rays from the different images. A branching process is started at the ground extending into dark areas assumed to correspond to the trunk or branches. The given results are plausible, but there is much human intervention involved. The most sophisticated automatic approach today is arguably (Shlyakhter et al. 2001). 3D volumes are generated as in (Sakaguchi & Ohya 1999). From the volumes 3D medial axes are constructed. The medial axes are constrained to the ''botanical fidelity of the branching pattern and the leaf distribution'' (Shlyakhter et al. 2001) via an open Lindenmayer-, or in short L-system (Měch & Prusinkiewicz 1996). Again, manual interaction is employed to generate results which are good in terms of visualization.

We show how generative statistical modeling based on L-systems and Markov Chain Monte Carlo – MCMC makes it feasible, to match branches in wide-baseline image sequences taken unconstrained with a standard consumer camera in spite of the problems with clutter and occlusions stated above. Our basis is a procedure (Mayer 2005) for the highly precise automatic determination of the orientation of images making use of calibration via the five-point-algorithm (Nistér 2004). Corresponding points are obtained with high precision by least-squares matching and bundle adjustment is used after every step.

In Section 2 the basic idea of generative statistical extraction employing L-systems and MCMC is described. We use statistical sampling in the form of MCMC to generate the parameters of an L-system modeling the 3D characteristics of trees which comply with the data evaluated in terms of likelihood. The generation of 3D hypotheses, their 2D projection, and evaluation are described in Sections 3 and 4. Hypotheses for trunks are generated from approximately vertical lines matched in several images. For the branches suitable prior distributions for the parameters are discussed particularly focusing on the branching angles. The evaluation of new hypotheses is conducted using the (normalized) cross correlation coefficient (*CCC*) as a substitute for likelihood. After presenting first results demonstrating the potential of the approach in Section 5, the paper ends up with conclusions.

2 Generative Statistical Extraction Using L-systems and MCMC

The branching structure of trees is difficult to extract from terrestrial wide-baseline urban image sequences because of possibly weak contrast and background clutter from other objects, e. g., facades or other trees. To construct 3D models of trees, we need to match the branches. Often, the ordering constraint, i. e., a point left of another point on an epipolar line in one image is also left of the corresponding point on the epipolar line in the other image, is employed to guide matching. Yet, because of the complex 3D structure of trees, the ordering constraint is often not valid even for images taken close to each other. All this means that the bottom-up / data-driven extraction of branches and matching them in 3D does not seem promising and suitable constraints describing the structure of trees are essential for their 3D reconstruction.

We describe the structure of trees in terms of their growth, or more particularly branch-

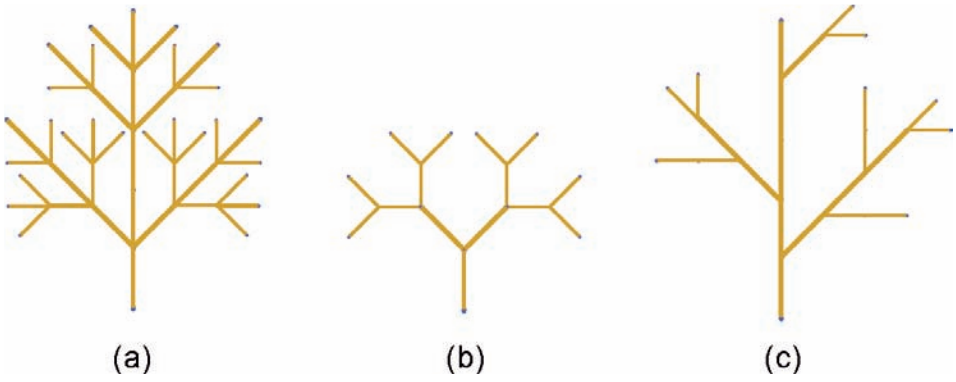

Fig. 1: Types of branching structures: (a) monopodial (b) sympodial, dichasium (c) sympodial, monochasium.

ing, by a Lindenmayer-, or in short L-system (Měch & Prusinkiewicz 1996). It is a parallel string rewriting system representing branching structures in terms of bracketed strings of symbols with associated numerical parameters, called modules.

The simulation of branching starts with an initial string (axiom). By means of productions all modules in the predecessor string are substituted by successor modules. Whether a production is applicable can depend on a predecessor's context, values of parameters, and like in our case on random factors (also termed stochastic L-systems). By means of context-sensitive L-systems interactions between plant parts can be represented. We do not use this for our first proof-of-concept implementation described in this paper, although it would certainly be helpful. By recursively using the same productions, L-systems represent self-similarity, an important biological characteristic of plants.

Basically, branching structures of trees can be divided into two main groups for which different production rules have to be used: monopodial and sympodial (Deussen & Lintermann 2005). The monopodial branching system (cf. Figure 1 (a)) has a prominent main axis, which is stronger and longer than the side branches. The side branches are again stronger and longer than the side branches of the second order, etc. Because of the dominant axes monopodial branching structures have a radially symmetric crown.

Fig. 1 (b) and (c) show the two main types of sympodial branching. Sympodial, dichasium branching means that two buds of a branch sprout and grow synchronously. For this kind of tree trunk and crown are clearly separated. The most common branching structure for trees is sympodial, mono-chasium branching, where one of the secondary branches has approximately the same direction as the original branch. Sympodial, monochasium branching results into only partially symmetric branching structures, which will still often appear very similar to monopodial branching.

For the sake for flexibility, we employ at the moment a mixture of all three sorts of branching: We let the branches sprout at the end of the trunk or a branch in all possible directions, yet preferring inclinations around 45° via a prior function (cf. below).

Modeling with L-systems enforces tree-like branching structures. Yet, L-systems alone only give means to generate and also visualize trees. For their extraction from images they need to be linked to a means for extraction. We decided to employ a generative statistical approach based on MCMC, where likely candidates of branches are generated by stochastical sampling and are verified by comparing simulated and real images.

In Fig. 2 the basic idea of our approach is presented. After extracting the trunk as described below, branches are grown ran-

domly guided by appropriate prior distributions and are then projected into the images via the given highly precisely known orientation parameters. The hereby generated simulated images are matched to the given images. As model for the background clutter we use Gaussian noise.

By linking stochastic sampling, L-systems, and likelihood from the images we find a tree structure very similar to that of the real tree. While L-system and MCMC can produce a typical tree, e. g., a beech, the link with the likelihood generated by matching with the images results into a beech with the particular characteristics that can be seen in the given images.

3 Generation of 3D Hypotheses

While we focus on the branching structure, a basic part of many trees we are interested in is the trunk. For it we extract straight lines, assuming that trunks correspond to thick, mostly vertical lines. The vertical direction is presumed to be known approximately by basically taking images horizontally. It can often be improved by computing the vertical vanishing point from the vertical edges of trunks or on facades as we focus on urban scenes. Vertical lines, i. e., hypotheses for trunks, are verified by matching in several images. We use the trifocal tensor (Hartley & Zisserman 2003) derived from the known orientation parameters to predict from lines in two images hypotheses for lines in other images. For the remainder of the paper we assume that the position of the tree is determined by the trunk.

This proof-of-concept paper is limited to the first two levels of branches. A branch in 3D object-space is modeled as a cylinder with known begin. As parameters azimuth (angle with x-axis of branch projected into horizontal plane), inclination (angle between branch and horizontal plane), length, and diameter are used.

We assume that the vertical direction is approximately known (cf. above). The x- and y-axis are taken from the local coordinate system of the first camera after aligning it with the vertical direction. The azimuth is sampled by MCMC with a uniform distribution between 0° and 360°. Because of the resulting symmetry, only a half circle is needed for the inclination (cf. Fig. 3). For most types of trees the majority of branches points upwards. We thus have empirically devised a prior distribution for the inclination with highest probability around 45°.

For length and diameter normal distributions are considered. Our first experiments were conducted with means between 0.7 m and 1.5 m for the length for the first level of branches. The diameter is set to a fixed value.

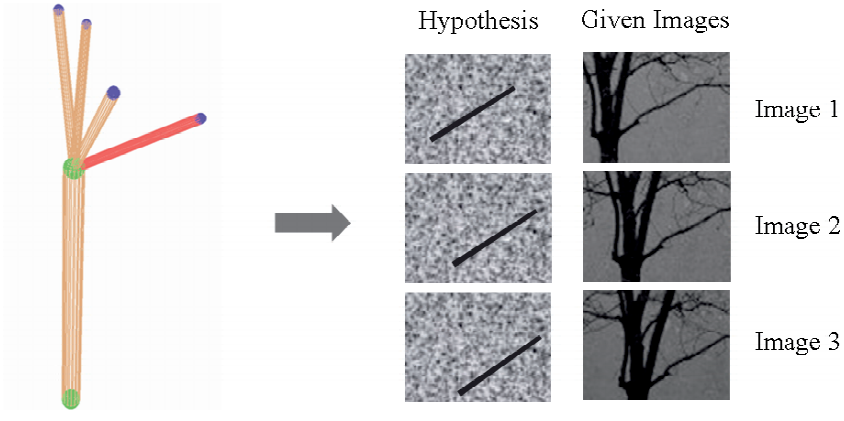

Fig. 2: Stochastical sampling based on an L-system results in a 3D tree hypothesis (left). Projection of a new branch (red) into three empty images with randomly textured background (center) and given image data (right). For the sake of clarity only the projection of the new branch is shown.

4 Projection and Evaluation

The hypotheses generated above are projected into 2D resulting into simulated images. They are evaluated by comparing the simulated images with the given images. The projection of 3D cylinders entails a larger computational effort. For MCMC many of these projections are needed. We thus decided for the proof-of-concept prototype, where we did not want to use a graphics processing unit (GPU) due to missing experience with its programming, to use a simple and efficient 2D representation derived from the 3D representation. Another reason for doing so is that the projection of the branches results into patches of nearly constant brightness anyhow. The chosen 2D representation consists of trapezoids. The color is taken as average of the trunk color. A trapezoid is described by the parameters direction (angle with x-axis), length, width of begin, and width of end.

We determine the parameters of the trapezoid as follows: The centers of the begin and the end are obtained by projecting the centers of the circles, i. e., the end points of the axis delimiting the cylinder on both sides, into the image via

$$\mathbf{x}' = \mathbf{P}\mathbf{X}$$

with (homogeneous) 3D points $\mathbf{X}$, image points $\mathbf{x}'$, and the projection matrix $\mathbf{P}$ (Hartley & Zisserman 2003). To compute reasonable approximations for the widths, we connect each end point of the axis of the cylinder with the camera center and determine a normal to this vector. The distance between the projections of the end point of the axis and of a point on the normal with distance radius of the cylinder from the axis equals half the width in the image.

The projected hypothesis is compared with the corresponding original image i by means of the cross correlation coefficient CCC_i for the intensities computed by HSI color transformation. To be able to compare different hypotheses, the matching is done against the projections of the convex 3D hull of all hypotheses. As MCMC sampling usually entails a larger number of iterations, the comparison has to be efficient. This is done by an incremental update of only those parts of the 2D projection and the corresponding variances and covariances, which have been changed.

CCC_i values for the n individual images are combined via multiplication into a global CCC value.

$$CCC = \prod_{i=1}^{n} CCC_i.$$

We use multiplication because we interpret the CCC_i values as likelihoods and we assume independence of the images given the 3D model. Moreover, we found empirically, that this conservative combination helps to sort out wrong hypotheses early. We are aware that the actual size of the CCC_i values can be far from correct likelihoods. Yet, our experiments give evidence to assume that they give a reasonable approximation to correct likelihoods. A function linking raw

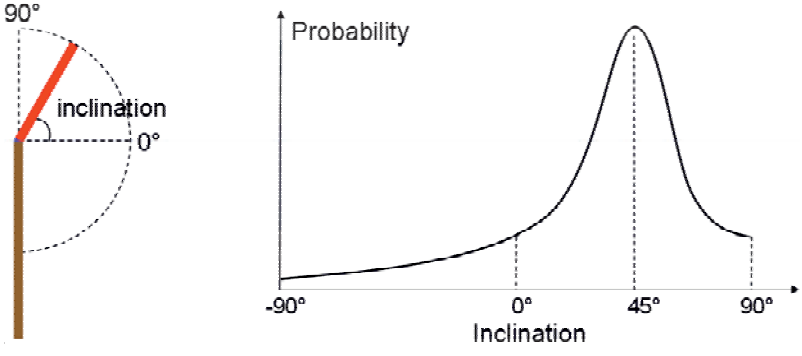

Fig. 3: left: Inclination of branch (red) – right: Empirically determined prior distribution.

CCC_i values and likelihoods could be obtained by determining the statistics of CCC_i values for a larger number of known correct and incorrect hypotheses for branches at a certain level.

By means of experiments we found that it is not useful to sample all parameters of a branch at the same time. Thus, MCMC sampling of the parameters is conducted sequentially. First, only azimuth and inclination are jointly varied over 1000 iterations while the length is kept fixed. The length is optimized only afterwards with 500 iterations. For future research we plan to relax the sequential sampling via conditional probabilities controling which parameter to sample next.

5 Results

Figs. 4 and 5 present first results. The input data consists of an image quadruple for the first and an image triplet for the second, both taken unconstrained with a hand-held 5 Megapixel camera. As output we obtain a VRML (virtual reality modeling language) model describing the trunk and the first two levels of the main branching system of the trees.

The scenes, taken in China and Germany under very different lighting conditions demonstrate, that we can basically determine the branching structure on the first two levels even though branches are partially occluded as, e. g., one of the left branches of the Chinese quadruple. Yet, we note that our proof-of-concept implementation still mis-

Fig. 4: Result for a Chinese image quadruple limited to the trunk and the first two levels of branches – Images (top) as well as the 3D tree, the cameras as (green) pyramids, and points used by the orientation procedure (bottom).

ses branches and reaches only a limited accuracy.

6 Conclusions and Discussion

We have proposed an approach for the extraction of the branching system of unfoliaged trees from wide-baseline image sequences. It combines the descriptive power for trees of L-systems with statistical sampling by means of MCMC and cross correlation into a statistical generative approach. Using the approach we are able to extract branches even when they are partly occluded as demonstrated by our first results. The envisaged final result of our approach is the basic branching structure of a particular tree. It will allow very realistic visualizations, e. g., for movies, and one could add leaves with different colors for different seasons. The branches could even be animated by simulating the forces of wind on them. When analyzing many trees, the resulting statistics for the parameters could be used for ecological applications or for simulating the interaction with radio waves for synthetic aperture radar (SAR).

Concerning future research, we first want to generalize the implemented L-system in the direction of open L-systems (Měch & Prusinkiewicz 1996) and distinguish between different types of trees (monopodial, sympodial; cf. above). We might need to change the parameterization away from the vertically centered azimuth and inclination angles to a more local, context-based representation using branching angles.

We also note that generative statistical modeling is not confined to L-systems. We basically just need a means to construct realistic trees that can be efficiently controlled. For this, e. g., also (Lintermann & Deussen 1999) could be a basis. We right now assume, that the upper stages of branches with very thin twigs might be grown stochastically to just match the image density, but it has to

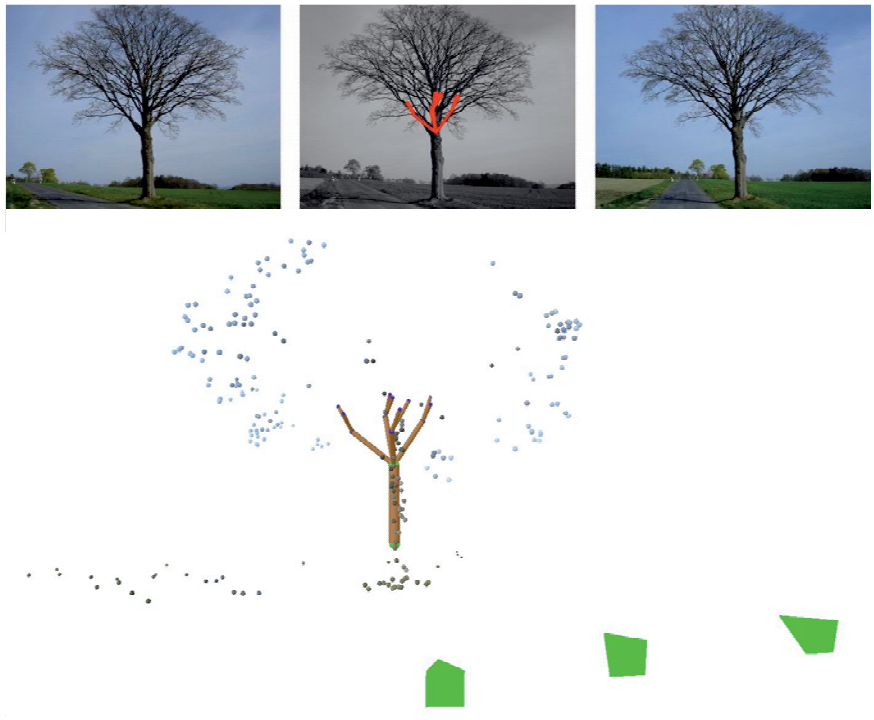

Fig. 5: Result for an image triplet limited to the trunk and the first two levels of branches – branches projected into images (top) as well as the 3D tree, the cameras as (green) pyramids, and points used by the orientation procedure (bottom).

be seen if and on which level of branching this is a valid assumption.

It should be possible to learn parameters such as contraction rates for lengths and diameters or branching angles by extracting a larger number of trees leading to priors probably conditional to the branching level. As noted above, by correlating against trees and representative samples of the background, a function to upgrade correlation coefficients to likelihoods could also be learned.

One question which arises is, how many branches are to be formed on a level and how many levels are appropriate for the tree, i. e., to control the complexity. If there are other trees or facades with strong linear textures in the background, there will be a strong tendency, that too many and thus too dense branches will be estimated and that they also extend beyond the perimeter of the tree. We want to tackle this issue by means of model selection. The idea is to balance the complexity of a hypothesis, i. e., the size of the tree or more particularly the number of parameters, against its likelihood according to the data. For this, the theory developed for compositional systems (Geman et al. 2002) might prove helpful, possibly also in conjunction with reversible jump (RJ) MCMC (Green 1995), to dynamically add better and delete worse hypotheses, the latter, if better solutions evolve.

Finally, we note that our modeling should be useful to find trees in much more explicit laser-scanner data, though the latter is linked to more effort for data acquisition.

Acknowledgements

We thank Deutsche Forschungsgemeinschaft for supporting Hai Huang under grant Ma 1651/11-1 and the anonymous reviewers for their helpful comments.

References

Deussen, O. & Lintermann, B., 2005: Digital Design of Nature. – Springer. Berlin, Germany.

Forsyth, D., Malik, J., Fleck, M., Greenspan, H., Leung, T., Belongie, S., Carson, C. & Bregler, C., 1996: Finding Pictures of Objects in Large Collections of Images. – Object Representation in Computer Vision II. Springer-Verlag, Berlin, Germany, 335–360.

Geman, S., Potter, D. & Chi, Z., 2002: Composition Systems. – Quarterly of Applied Mathematics **LX**: 707–736.

Green, P., 1995: Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. – Biometrika **82**: 711–732.

Haering, N., Myles, Z. & Vitoria, N., 1997: Locating Deciduous Trees. – IEEE Workshop on Content Based Access of Image and Video Libraries, 18–25.

Hartley, R. & Zisserman, A., 2003: Multiple View Geometry in Computer Vision – Second Edition. – Cambridge University Press. Cambridge, UK.

Lintermann, B. & Deussen, O., 1999: Interactive Modeling of Plants. – IEEE Computer Graphics and Applications **19** (1): 2–11.

Mayer, H., 2005: Robust Least-Squares Adjustment Based Orientation and Auto-Calibration of Wide-Baseline Image Sequences. – ISPRS Workshop in conjunction with ICCV 2005 "Towards Benchmarking Automated Calibration, Orientation and Surface Reconstruction from Images" (BenCos), Beijing, China, 1–6.

Měch & Prusinkiewicz, P., 1996: Visual Models of Plants Interacting with Their Environment. – SIGGRAPH '96, 397–410.

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.

Sakaguchi, T. & Ohya, J., 1999: Modeling and Animation of Botanical Trees for Interactive Environments. – Symposium on Virtual Reality Software and Technology, 139–146.

Shlyakhter, I., Rozenoer, M., Dorsey, J. & Teller, S., 2001: Reconstructing 3D Tree Models from Instrumented Photographs. – IEEE Computer Graphics and Applications **21** (3): 53–61.

Address of the Authors:

Dipl.-Ing. Hai Huang, Prof. Dr.-Ing. Helmut Mayer, Bundeswehr University Munich, Institute of Photogrammetry and Cartography, D-85577 Neubiberg, Tel.: +49-89-6004-3429, Fax: +49-89-6004-4090, e-mail: Hai.Huang@unibw.de, Helmut.Mayer@unibw.de

Manuskript eingereicht: Mai 2007
Angenommen: Juni 2007

Photogrammetrie • Fernerkundung • Geoinformation 6/2007, S. 437–445, 6 figs.

Road Part Extraction for the Verification of Suburban Road Databases[1]

Anne Grote, Matthias Butenuth & Christian Heipke, Hannover

Keywords: Road Database, Suburban Areas, Segmentation, Verification, Normalized Cuts

Summary: In this paper, a new strategy for the verification of road databases in suburban areas using aerial orthoimages is suggested. It consists of two stages: the extraction of road parts and the comparison with the database roads. High resolution aerial images are used for the road part extraction which comprises the steps segmentation, grouping and evaluation. The segmentation is designed to yield a good division between road areas and the surroundings. We use the Normalized Cuts algorithm, which is a graph-based approach that divides the image on the basis of pixel similarities. The initial segments have to be grouped due to an enforced oversegmentation. The criteria for both segmentation and grouping are based on colour and edge properties. The grouped segments are then evaluated in order to extract road parts, based mainly on shape criteria. In the second stage, the extracted road parts are compared to the database roads, leading to the acceptance or rejection of each database road. This paper focuses on road part extraction in suburban areas; the verification stage itself is beyond the scope of the current article. Nevertheless, the general approach is briefly discussed as part of the overall strategy. Preliminary results for the road extraction stage show that reliable road parts can be extracted with the presented approach; ideas for the procedure in the verification stage are presented as outlook.

Zusammenfassung: *Extraktion von Straßenstücken zur Verifikation von Straßendatenbanken in Vorstadtgebieten*. In dieser Publikation wird eine neue Strategie zur Verifikation von Straßendatenbanken in Vorortgebieten mit Hilfe von Luftbildern vorgeschlagen. Sie besteht aus zwei Stufen: der Extraktion von Straßenstücken und dem Vergleich mit den Straßen in der Datenbank. Die Straßenextraktion, für die hoch aufgelöste Luftbilder verwendet werden, wird in drei Schritten durchgeführt: Segmentierung, Gruppierung und Evaluierung. Die Segmentierung ist so ausgelegt, dass sie eine gute Unterscheidung zwischen Straßenflächen und der Umgebung liefert. Wir nutzen den Normalized-Cuts-Algorithmus, eine graphbasierte Methode, die das Bild mit Hilfe von Pixelähnlichkeiten segmentiert. Die vorläufigen Segmente müssen dann aufgrund einer erzwungenen Übersegmentierung gruppiert werden. Die Kriterien für die Segmentierung und die Gruppierung beruhen auf Farb- und Kanteneigenschaften. Die gruppierten Segmente werden anschließend evaluiert, hauptsächlich anhand von Formeigenschaften, um Straßenstücke zu extrahieren. In der zweiten Stufe werden die extrahierten Straßenstücke mit den Straßen in der Datenbank verglichen, was zur Annahme oder Ablehnung der Datenbankstraße führt. In dieser Publikation liegt der Schwerpunkt auf der Extraktion von Straßenstücken in Vorstadtgebieten, die Verifikation selbst geht über den Rahmen dieses Artikels hinaus, die allgemeine Vorgehensweise wird aber als Teil der Gesamtstrategie kurz beschrieben. Für die Straßenextraktion werden erste Ergebnisse vorgestellt, die zeigen, dass mit dem vorgestellten Ansatz zuverlässige Straßenstücke extrahiert werden können; für die Vorgehensweise bei der Verifizierung wird ein Ansatz im Ausblick vorgestellt.

[1] Enhanced version of a paper published in the International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (3/W49A).

1432-8364/07/2007/0437 $ 2.25
© 2007 DGPF / E. Schweizerbart'sche Verlagsbuchhandlung, D-70176 Stuttgart

1 Introduction

Today, GIS data are widely used and incorporated in decision finding procedures, navigation systems and many other applications. As users rely on the correctness of GIS data, it is necessary to check databases frequently to eliminate errors and to add new objects. Manual database assessment is very time-consuming, which is why many approaches have been developed over the past years to automate this process. In many of these approaches, up-to-date aerial or satellite images are used to automatically extract objects and to compare them to objects contained in the database (Baltsavias 2004).

One prominent class of objects in geospatial databases are roads. Road databases are very important for many applications, for example navigation systems or spatial planning. As they are also subject to frequent changes, road database assessment and update is a major topic. Accordingly, many approaches deal with road database assessment and update in rural areas, for example (Zhang 2004) or (Gerke 2006). In contrary, only few approaches work in urban or suburban areas due to the highly complex structure found in urban scenes, which complicates the task of automatic road extraction. Many approaches have been developed to extract roads, some of them are summarized in (Mayer et al. 2006), but only few of them are designed for urban areas. In (Price 1999, Youn & Bethel 2004) the road network is expected to be a more or less regular grid, but this constraint is not suitable for many European urban areas. Another approach uses a very sophisticated road and context model and is based on grouping small extracted entities to lanes, carriageways and road networks (Hinz 2004). It employs a large set of parameters that must be carefully adapted for different scenes. In recent work, colour properties are exploited, for example in (Zhang & Couloigner 2006): the authors perform a pixel-based multispectral classification and use shape descriptors to reduce the number of misclassifications, but the completeness and correctness rate is only about 50 %. In our opinion, a reason for this fact is that the multispectral classification does not take into account the spatial relations of the pixels and colour and shape properties are treated separately.

In this paper, a new approach for the extraction of road parts in suburban areas for the verification of road databases is presented. From the above mentioned approaches we can deduce that a proper segmentation algorithm is essential as a first step for road extraction in suburban areas and that it is important to combine several features in the segmentation step because of the complex surroundings. A simple line based road model, as used in many road extraction approaches for rural areas, is not applicable. Therefore, we use Normalized Cuts (Shi & Malik 2000) for the segmentation. One advantage of this method is the possibility to combine several different criteria in one step, which can be selected according to the application. In this way, it is possible to use modelled knowledge about roads already in the segmentation step. Another important advantage is the incorporation of both local and global characteristics of the objects present in the image. Afterwards, the initial segments are grouped to reverse oversegmentation. The resulting segments are then evaluated in order to find road parts. The verification stage follows the approach of (Gerke 2006). The verification itself is not part of this paper; nevertheless, the verification strategy is briefly explained.

In the next section the general strategy for road database verification in suburban areas is explained. The specific approach for road part extraction (segmentation, grouping and evaluation) is described in section 3, results are presented in section 4. Some conclusions and an outlook on further work are given in section 5, including an example demonstrating the concept for the verification strategy.

2 Verification Strategy

The aim of our research is the assessment of a road database. This process consists of two stages (cf. Fig. 1):

- Road part extraction
- Comparison to the database road.

As mentioned before, road extraction in suburban areas is more complicated than in rural areas due to the inhomogeneous background. We use an area-based road model and apply our strategy to high resolution CIR images.

The two stages are carried out for each database road individually. A region of interest is defined around the database road, whose length matches the length of the road to be assessed and whose width exceeds the expected road width in order to avoid segments to be forced into a roadlike shape by the region borders, thus distorting the evaluation. The following road extraction steps are carried out in each region of interest.

The first stage of the verification strategy, the road part extraction, consists of three steps: segmentation, grouping, and evaluation. The region of interest within the image is segmented using the Normalized Cuts method as described in section 3.1. The aim of the segmentation is a distinct separation between road segments and non-road segments. The result is an oversegmentation, which is necessary to obtain as many parts as possible of the road border belonging to a segment border, even if the image information is weak. In a second step the initial segments have to be grouped to larger segments, before they can be evaluated and divided into road segments and non-road segments. As will be shown later, it is not possible to evaluate the initial segments correctly based on their shape characteristics, because they are too small for deriving reliable shape attributes. The grouping is currently done using a simple iterative algorithm merging initial segments with similar mean colour, weak edges at shared borders and low overall colour variance. Details are described in section 3.2. Next, the grouped segments have to be evaluated in order to select the road segments. The evaluation is based on shape and colour criteria of roads like elongation, maximum width and homogeneity. The aim of this step is to extract only reliable road parts, in order to avoid false road parts in the comparison with the database road.

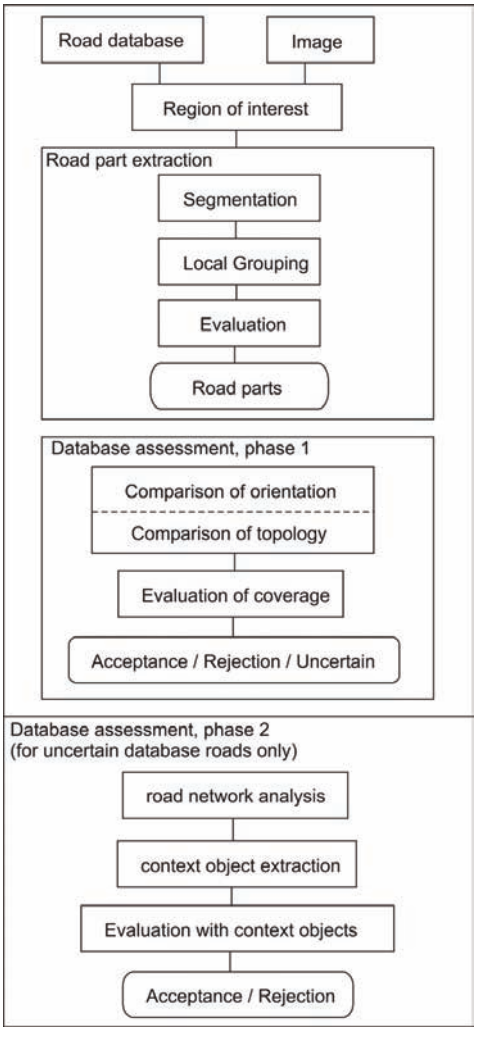

Fig. 1: General strategy.

In the second stage of the verification strategy, the database road is compared to the extracted road parts in order to assess the database road. The assessment is based on the road database assessment method described in (Gerke 2006). A brief description of the general concept is given here. If road parts could be extracted, their topology and orientation are compared to that of the database road. The orientation should be approximately the same, and the buffer

around the database road should contain the extracted road part. This is checked for each extracted road part. After the comparison we check to what extent the database road is covered with extracted road parts. Based on the results of this check, the database road is labelled accepted, rejected or uncertain.

The database roads labelled uncertain are re-examined. For this purpose, the road network properties and context objects are used. If the uncertain database road is important for the road network, context objects are extracted in the region of interest and it is tested whether they can explain the failure to extract reliable road parts in the first stage. The result is a final acceptance or rejection.

3 Approach for Road Part Extraction

3.1 *Segmentation*

The Normalized Cuts method is used for the initial segmentation. Normalized Cuts is a graph-based method using an undirected graph with weighted edges that is to be divided into segments with similar features (Shi & Malik 2000). Pixels are defined as nodes which are connected by weighted edges. The edge weights describe the similarities between the pixels. This graph is cut into segments to meet the following minimization criterion, dividing dissimilar pixels into different segments:

$$Ncut(A_1, \ldots, A_n) = \sum_{i=1}^{n} \frac{link(A_i, V \backslash A_i)}{link(A_i, V)} = \min \quad (1)$$

Using Equation 1, the graph is divided into n sets of nodes A_i. V is the set that contains all nodes in the whole graph. *Link* is the sum of all weights of the connecting edges between two sets:

$$link(P, Q) = \sum_{p \in P, q \in Q} w(p, q) \quad (2)$$

where $w(p, q)$ is the weight between two nodes p and q belonging to the two sets P and Q. The calculation is based on computing eigenvectors of a matrix containing all similarity weights, each segment is defined by one eigenvector. Details of the calculation can be found in (Shi & Malik 2000).

The aim of the segmentation is to separate the road parts in the image from the non-road parts. The similarity criteria are derived from the road model:

- Presence and strength of edges between two pixels
- Colour difference
- Hue difference
- Road colour derived from database information.

Roads are divided from their surroundings by edges, therefore edges are used as a criterion: if there is an edge between two pixels, the pixels are considered to be dissimilar. The similarity measure depends on the edge strength: a high edge strength leads to a low similarity.

The second criterion is colour because roads usually have homogeneous surfaces and the pixel colour stays approximately the same. A measure for the colour similarity of two pixels is the distance between these two pixels in colour space. If the distance is short, the pixels have a high similarity regarding colour.

As a third criterion hue is used because a significant hue difference almost certainly indicates a different object. On the other hand, in parts darkened by shadows the hue of an object remains the same if certain conditions are met (Perez & Koch 1994). Thus, if the hue difference is large, the pixels are considered dissimilar. There is certainly some correlation between the colour and the hue criterion; however, our experiments have shown that the use of both yields rather good results.

The database information is used to obtain colour information about the roads: assuming that the position of the road is approximately correct, we compute the average colour values from the position of the database road centerline. This information is used to increase the chance that segments are divided along a road border: if the colour

of one pixel is close to the average road colour, while the other is not, the pixels are dissimilar.

The similarity measures are combined to one similarity weight for each pixel pair. The Normalized Cuts method is performed with these weights. The number of segments, which must be specified before the calculation, must be large enough to prevent merging of road and non-road segments.

3.2 *Grouping*

As mentioned before, the image segmentation algorithm results in an oversegmentation. This is necessary to minimize the risk of losing road borders. Therefore, in the next step the initial small segments have to be grouped to larger, more meaningful segments before they can be evaluated. We follow the approach suggested by (Luo & Guo 2003). The authors aim at a general grouping algorithm as a bridge between image segmentation and high-level object extraction algorithms. The region properties they use include, among others:

- Colour mean difference between two regions
- Edge strength along shared border
- Colour variance of cross-border area
- Contour continuity between two regions

The first two criteria are particularly suitable for our approach, because the enforced oversegmentation often produces segment borders where the image information does not really justify a separation. The edge strength along a shared border and the mean colour difference between both segments are suitable criteria to group these initial segments to larger ones.

At present, we use a simple iterative approach for grouping the segments. For each pair of initial segments, several criteria are calculated. Based on the similarity criteria used by (Luo & Guo 2003) we use the difference of mean colour (separately for the three channels) and the edge strength of the intensity channel in the region around the shared border (border region). The border region is a small band along the shared border. Instead of the colour variance of the cross-border area we use the joint standard deviation of colour of the merged regions (separately for the three channels) because tests have shown this to be the better criterion for our application. For all seven criteria, the calculated values have to be below predefined thresholds for the segments to be considered for merging. In each iteration step, only the two segments with the best values for all criteria (the least colour differences, the least edge strength and the least colour standard deviations) are merged. The iteration continues until the values for every segment pair exceed at least one threshold.

3.3 *Evaluation*

After grouping, the segments can be evaluated in order to extract road parts. The evaluation is based on shape and spectral characteristics of roads. The following characteristics are used for evaluation:

- Elongation
- Rectangularity
- Width
- NDVI

The elongation indicates the difference of the object from a circle. It is given by the ratio of the squared perimeter and the area, which is high for elongated objects. A road part should have a high elongation. The rectangularity is a measure for the similarity of an object with a rectangle. It is calculated using the Discrepancy method described in (Rosin 1999). Road parts should be close to rectangular. The width of a road part should not be much larger than the average width of a road. The average NDVI (Normalized Difference Vegetation Index) is calculated for each segment. The NDVI is the difference between the infrared channel and the red channel normalized by the sum of both channels. For road parts the NDVI should be low. In our tests, thresholds are defined for each of these criteria and segments are extracted as road parts if they fulfil all of them.

4 Results

The first stage of the strategy, the approach for road part extraction, was tested on CIR aerial orthoimages with a resolution of 0.1 m. The images depict a suburban scene in Grangemouth, Scotland. Road database data were simulated by manually digitizing the visible roads in the images.

For segmentation with Normalized Cuts, the regions of interest are divided into subsets of approximately 200×200 pixels for computational reasons. Each subset is segmented by the Normalized Cuts algorithm yielding 20 segments, an empirical value that is suitable for this image size and scene complexity. The width of the region of interest is set to approximately three times the expected road width.

Fig. 2 shows an example of a suburban road with the database road centerline shown in green. In Fig. 3 the Normalized Cuts segmentation result obtained with the similarity criteria described in section 3.1 is displayed. Segment borders are indicated by yellow lines. The result demonstrates that the segmentation in general has succeeded: road and non-road areas are in most instances clearly divided by initial segment borders. Exceptions can be found in shadow areas or where the contrast between the road and the surroundings is low, for example in the right part of the image.

After segmentation, the initial segments are grouped as described in section 3.2. Fig. 4 shows the grouping result from the segments of Fig. 3. The following grouping parameters are used: the mean colour of two segments to be merged should be the same, the width of the border region is 7 pixels, the maximum for the mean edge strength is set to 50, and the maximum standard deviation for two merged segments is set to 40.

The grouping result shows one distinctive road segment in the left part of the image. To the right, there is one big road segment that is more problematic: it contains one part of a parking lot and some parts of the pavement. The pavement parts are parts of an initial road segment and are not critical because they do not affect the overall shape of the road to a great extent. The parking lot poses more difficulties for an evaluation because it has the same colour characteristics as the road and is not separated by a distinct border so it is merged with the road segment, distorting its shape.

The grouped segments are evaluated using the criteria described in section 3.3. The thresholds used for the evaluation are: elongation more than 50, rectangularity more

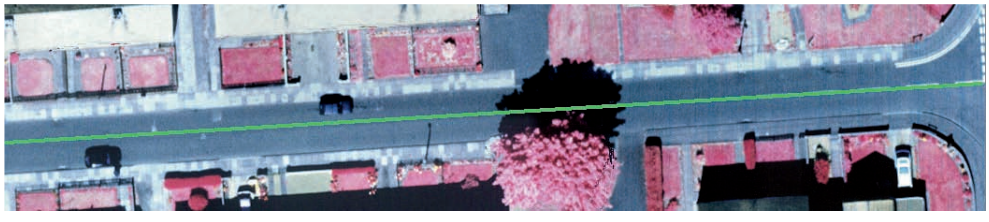

Fig. 2: Original image with database road.

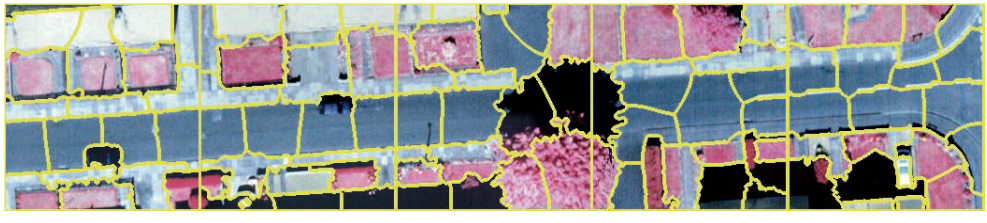

Fig. 3: Segmentation result.

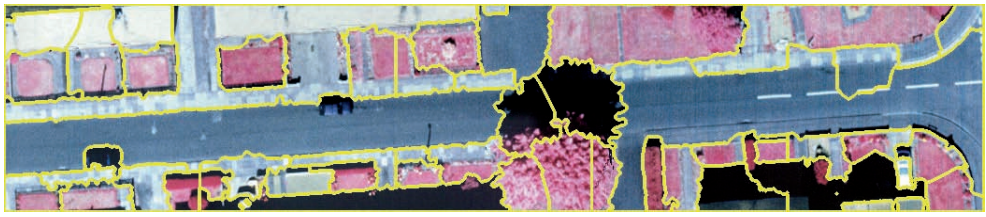

Fig. 4: Grouping result.

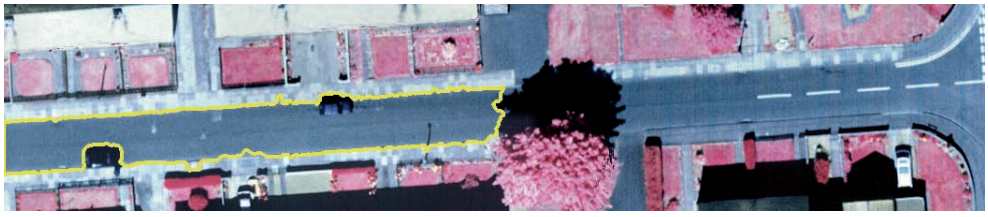

Fig. 5: Evaluation result.

than 0.7, width not more than 2 meters above assumed the road width, NDVI less than 0. The evaluation result for the grouped segments is displayed in Fig. 5. One segment is correctly extracted as road part. No false positives are extracted, which is important at this stage to prevent the verification from falsely accepting database roads. The other larger road segment seen in Fig. 4 on the right could not be extracted due to low rectangularity caused by the part of a junction on the right side and the inclusion of the parking lot on the left side.

5 Conclusions and Outlook

The results in this paper show the general usability of the approach for the detection of roads in aerial images of suburban areas. The Normalized Cuts method was shown to be suitable for the segmentation step. The combination of image features adapted to the appearance of roads yields a good division between road segments and non-road segments. By considering global aspects of the image as well as local ones, the algorithm is able to ignore noise, small surface changes and weak edges. The division of the image depends on the overall image content which allows the segments to be more coherent and perceptually meaningful than segments obtained by a local segmentation only.

One drawback of the Normalized Cuts algorithm is that the calculation is computationally expensive and the image has to be divided into smaller subsets to make the calculation possible. Another drawback is the fact that the number of segments has to be determined before starting the calculation. It is desirable to find a way to estimate the appropriate number of segments from the given data. One possibility is an iterative approach, repeating the Normalized Cuts algorithm with a varying number of segments and subsequently selecting the optimal segmentation.

The grouping result shows that it is possible to use the oversegmented results from the Normalized Cuts algorithm and group them to bigger segments whose shape can be assessed regarding their correspondence with the road model. The grouping works well for road parts without many disturbances by context objects. One problem are areas that are directly connected to the road and have the same colour as the road, like the parking lot in Fig. 5. Here, an additional grouping criterion, for example border continuation, could be helpful. Alternatively,

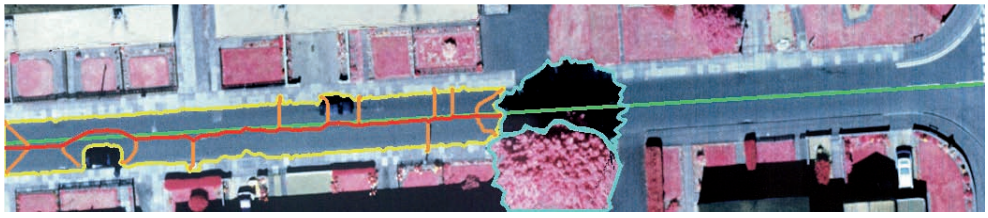

Fig. 6: Outlook on verification strategy.

the segments could be split such that protruding parts are removed.

Preliminary results for the evaluation show that the chosen extraction criteria seem to be suitable for the extraction of roads. However, the rectangularity criterion cannot be used for curved roads, which is why we plan to replace it with a criterion that combines elongation and constant width. The evaluation algorithm itself also has to be improved, at the moment we use hard thresholds without combining or weighing the criteria which is necessary to obtain a more reliable extraction in varying surroundings.

At the moment, we use thresholds for both the grouping and the evaluation. We plan to investigate how to reduce the number of these thresholds and how to select meaningful values for the remaining ones using image and object characteristics.

The next step will be the implementation of the verification strategy, which already has been shown to work in rural areas (Gerke 2006). We will need to adapt this approach for suburban areas because the road network properties and context objects play a different role: the connection function of the road network is less pronounced than in rural areas, dead ends are more frequent in suburban areas, for example. It has to be examined to what extent the network function can still be used. As for context objects, we plan to introduce a vehicle finder into our approach as vehicles are common context objects in suburban and urban areas. In addition, they may also be useful to deal with parking areas.

Using the example from section 4, the verification will roughly proceed as follows (cf. Fig. 6): in the region of interest around the database road (green) one road part was extracted in the first stage (yellow). The skeleton is then computed for this road part (orange) and the longest path between junctions in the skeleton is selected, which is then extended to the borders of the road part (red). This extracted road line is then compared to the database road: the orientation is about the same. The shape, which can for example be compared using moments as described in (Geke 2006), is again about the same, so the part of the database road that is covered by the extracted road can be confirmed. However, only about 50 % of the database road is confirmed in this way, which is not enough to accept it. Assuming the road is an important part of the road network, context objects are considered in a second step, in this case a tree and a shadow (manually extracted, light blue). As the existence of a road is compliant with a tree standing beside it, this sufficiently explains the failure to detect a road in this part, so more than 60 % of the database road can now be confirmed. This would still be not enough to accept the road, but after improving the road extraction stage as suggested above, more road parts will be extracted in the right part of the image, such that the database road can finally be accepted.

Acknowledgements

This project is funded by the DFG (German Research Foundation) under grant HE 1822/16-2. The calculations were made with a C++ program partly adapted from a MATLAB program written by Timothée Cour, Stella Yu and Jianbo Shi. Their pro-

gram can be found on http://www.seas.upenn.edu/~timothee/software_ncut/software.html (last checked July 2007). We also gratefully acknowledge the comments of the anonymous reviewers.

References

Baltsavias, E.P., 2004: Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. – ISPRS Journal of Photogrammetry and Remote Sensing **58** (3-4): 129–151.

Gerke, M., 2006: Automatic quality assessment of road databases using remotely sensed imagery. – Dissertation, Universität Hannover, DGK Reihe C, Nr. 599: Verlag der Bayerischen Akademie der Wissenschaften.

Heuwold, J., 2006: Verification of a methodology for the automatic scale-dependent adaptation of object models. – International Archives of Photogrammetry and Remote Sensing **36** (3): 173–178.

Hinz, S., 2004: Automatic road extraction in urban scenes – and beyond. – International Archives of Photogrammetry and Remote Sensing **35** (B3): 349–355.

Mayer, H., Hinz, S., Bacher, U. & Baltsavias, E., 2006: A test of automatic road extraction approaches. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (3): 209–214.

Luo, J. & Guo, C., 2003: Perceptual grouping of segmented regions in color images. – Pattern Recognition **36** (12): 2781–2792.

Perez, F. & Koch, C., 1994: Toward color image segmentation in analog VLSI: algorithm and hardware. – International Journal of Computer Vision **12** (1): 17–42.

Price, K., 1999: Road grid extraction and verification. – International Archives of Photogrammetry and Remote Sensing **32** (3-2W5): 101–106.

Rosin, P.L., 1999: Measuring rectangularity. – Machine Vision and Applications **11** (4): 191–196.

Shi, J. & Malik, J., 2000: Normalized cuts and image segmentation. – IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (8): 888–905.

Youn, J. & Bethel, J.S., 2004: Adaptive snakes for urban road extraction. – International Archives of Photogrammetry and Remote Sensing **35** (B3): 465–470.

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, Q. & Couloigner, I, 2006: Automated road network extraction from high resolution multi-spectral imagery. – Proceedings of ASPRS Annual Conference, Reno, Nevada, 10 p., on CD.

Address of the Authors:

Dipl.-Ing Anne Grote, Dipl.-Ing. Matthias Butenuth, Prof. Dr.-Ing. Christian Heipke, Leibniz Universität Hannover, Institute of Photogrammetry and GeoInformation, D-30167 Hannover, Tel.: +49-511-762-2599, -4922, -2482, Fax: +49-511-762-2483, e-mail: grote@ipi.uni-hannover.de, butenuth@ipi.uni-hannover.de, heipke@ipi.uni-hannover.de

Manuskript eingereicht: Mai 2007
Angenommen: Juni 2007

Photogrammetrie • Fernerkundung • Geoinformation 6/2007, S. 447–457, 8 figs., 1 tab.

Semantic Data Integration: Data of Similar and Different Scales

Birgit Kieler & Monika Sester, Hannover; Haiqing Wang & Jie Jiang, Beijing, China

Keywords: Data Integration, Data Matching, Ontology

Summary: The integration of data of different origin leads to a lot of benefits: firstly, the properties of the individual data sets can be exchanged and used for mutual enrichment and benefit, secondly, the information from both sources can be ''adjusted'', leading to a more precise and reliable information. Such integration presumes that semantic relations between the data are known: the semantic correspondences help to enrich similar object classes and objects. If such semantic relations are not explicitly known, our approach is to exploit geometric correspondences of individual object instances stemming from different sources. From these geometric correspondences inferences concerning their semantic relations can be drawn. For these analyses, differences with respect to geometric and semantic granularity have to be taken into account. In this paper, we describe approaches to solve these problems.

Zusammenfassung: *Semantische Datenintegration: Daten ähnlicher und verschiedener Maßstäbe*. Die Integration von Daten unterschiedlicher Herkunft ist von großem Nutzen: einerseits können die Vorteile wie auch die Reichhaltigkeit der individuellen Datensätze ausgetauscht werden, um damit gemeinsam eine Verbesserung zu erzielen. Andererseits können die Informationen beider Quellen einander angepasst werden, wodurch eine höhere Genauigkeit der verfügbaren Daten entsteht. Solch eine Integration setzt voraus, dass die Beziehungen zwischen den Daten in einem bestimmten Ausmaß bekannt sind. Die semantischen Korrespondenzen helfen dabei, ähnliche Objekte miteinander zu vergleichen. Sind diese Relationen nicht explizit bekannt, nutzt unser Ansatz die geometrischen Korrespondenzen der individuellen Objektinstanzen der verschiedenen Datenquellen. Von diesen geometrischen Korrespondenzen können Rückschlüsse bezüglich ihrer semantischen Beziehungen aufgezeigt werden. Für diese Analysen müssen unterschiedliche geometrische und semantische Auflösungen der Daten berücksichtigt werden. In diesem Beitrag werden Lösungsansätze für diese Probleme vorgestellt.

1 Introduction and Overview

The growing availability of data also via the internet allows a growing interoperability of geodata as well as information sharing and reuse. This, however, presumes that the content of the data is known in order to draw meaningful and correct conclusions. Thus, for a beneficial data integration, which is a prerequisite to interoperability of data and services in Spatial Data Infrastructures (SDI), both the semantic and the geometric correspondences of these data sets have to be known.

If the semantic relationships between different geo-ontologies, like equivalence-, disjunction- or inclusion-relations are known, a geometric integration can be accomplished, e. g. a fusion or alignment process of geometry in order to obtain one improved geometry. Also an attribute transfer between the data sets is possible to enrich the

1432-8364/07/2007/0447 $ 2.75
© 2007 DGPF / E. Schweizerbart'sche Verlagsbuchhandlung, D-70176 Stuttgart

existing knowledge about the objects. Although the semantics of the attributes still will remain unclear, the availability of this additional attributes may be helpful for new applications. The general case is that the semantic relationships between arbitrary data sources are unknown and the corresponding semantic object classes have to be determined.

In this paper we are discussing two cases. In case A, we use two data sets of similar scale where the semantic correspondences are unknown. Our approach is to use an instance-based determination of transformation rules between the ontologies. In case B we use two data sets that describe settlement objects stemming from different scales. Although the semantic relationships between the ontologies are known, a direct geometric integration is not possible due to the different granularity of the objects involved. In this case, an aggregation of the detailed data set has to be performed first in order to derive a comparable geometric object description.

The paper is organized as follows. In the next section the background of the research is sketched and references to existing work are given. Then, our methods for the two integration cases are presented, first the semantic integration of data of similar scale, then the geometric integration of data of different scales. A summary and an outlook conclude the paper.

2 Related Work

Interoperability and especially data integration faces different types of problems (Bishr 1997): it has to take structural, semantic and geometric differences in the data sets into account. Structural interoperability can be achieved using standardized data formats (e. g. ISO, OGC). The most difficult problem is semantic interoperability as it deals with the task of determining correspondences between object descriptions stemming from different user communities. In this way the corresponding semantics of the object category ''lake'' and ''See'' in an English and German topographic data set automatically have to be determined (which, of course, is not only a question of language). The general approach in the identification of semantic correspondences is to do a manual schema integration using expert knowledge together with given object catalogues or ontologies (Kokla 2006). Such a process is not adequate and not longer feasible if arbitrary data sets, e. g. downloaded from the internet, have to be integrated. Therefore, automatic techniques are needed. Rodriguez & Egenhofer (2003) use equality and similarity measures to determine relations between classes from different ontologies. Another method to automate the integration process, is a so called instance-based or extensional determination of schema transformation rules (Volz 2005, Duckham & Worboys 2005). The underlying idea of this approach is that, if two objects have an identical name and / or geometrically coincide, then they probably also have something in common on the semantic level. Duckham & Worboys (2005) use the lattice theory to determine possible class correspondences. This formally very elegant way does not take the relative frequencies of correspondences into account. This is done by Volz (2005), however, only with a manual evaluation. In our approach we want to link both approaches in order to be able to determine semantic relationsships with a corresponding probability and thus quality values.

By contrast Fonseca et al. (2006) presented a framework for measuring the degree of interoperability between geo-ontologies, which only compares the descriptions of the ontologies and not the data themselves. The drawback of this approach is, that in the general case, it can not be assumed, that the names or descriptions of objects in different data sets are the same - except objects with a unique given name like names of cities or roads. Thus using the geometric relations to infer a semantic relation seems to be more promising and open for automation.

Samal et al. (2004) take both the semantic, geometric and contextual similarity into account to derive corresponding objects in different data sets. There are also data matching approaches for data sets with different

granularity (Mustière 2006, Uitermark 2001). Dunkars (2003) matches data of different scales to automatically determine the links to build a multiple representation database. Therefore a number of different measures based on topology, geometry, semantics and inter-object relationships are used to compare similarity. Bel Hadj Ali (2001) also uses different measures for assessing shape and positional quality of polygons between two different data sets.

3 Data Sets, Tasks and Approaches

For our investigations we used two different test cases with two geo-data sets each, in vector format. For case *A* (described in detail in Section 4) two data sets describing topographic objects were used: GDF (data set *A*1) and ATKIS (data set *A*2) (see Fig. 1). Whereas the GDF data (Geographic Data Files) was specially developed for vehicle navigation purposes, the German ATKIS (Authoritative topographic cartographic information system) provides topographic base data. Each object is described geometrically and semantically using the object classes and attributes described in the respective object catalogues. Both data sets are of the same scale (approx. 1 : 25.000). In Tab. 1 the hierarchical organisation of the corresponding ontologies is given in a simplified form. These presentations are not complete, only the object classes used for the analysis are listed. The two data sets are modelled differently; whereas GDF uses a two level hierarchy, followed by a further distinction using special attributes, ATKIS distinguishes three hierarchical levels, where the third level is also distinguished with attributes. Looking at the data and the structure, it is clearly visible that ATKIS exhibits a higher granularity both with respect to the object classes and the number of individual object instances. Our task is to automatically determine the semantic correspondences between object classes of the two data sets. The identification of corresponding object classes is achieved by a geometric overlay and a comparison of the frequency of the occuring object relations.

For case *B* (described in detail in Section 5) we are using two data sets with different scales and contents from China (see Fig. 2). Data set *B*1 is of scale 1 : 50.000 and specially developed for GIS purposes with comprehensive attributes while data set *B*2 is developed mainly for cartographic visualization in scale of 1 : 10.000 and contains no attributes. We are concentrating on settlement structures, namely residential areas in the small scale data and buildings in the large scale data set. The task is to transfer the attributes of the small scale data set *B*1 to the individual buildings in *B*2. Whereas on a first glance the solution to this problem might look simple, as the semantic equivalence between the object classes is already known, the situation is more complicated, because the data sets have been acquired at different time instances: data set *B*2 is more current and contains a lot of new buildings, which are not reflected in the residential areas of data set *B*1. Therefore, a mere geometric overlay or containment check of buildings and residential areas is not suffi-

Fig. 1: Data sets for test case *A*: GDF – (data set *A*1) (*left*) and ATKIS – (data set *A*2) (*right*).

Tab. 1: The hierarchical organisation with semantically describing class names of the *A*1- (*top*) and A2-geo-ontology (*bottom*) – from superclasses (*Level 1*) to subclasses (*Level 2 or 3*) with available attributes (*right*). Σ is the total number of instances in the test area.

	Level 1	Level2	Level3	Attributes	Σ
*A*1-GDF	11 Administrative Areas	1119 Order 8			2
		...			
	43 Waterways	4310 Water Element		Dispclass 1, 2, 3, 4, 5	10
		...			
	71 Land cover and use	7120 Woodland			3
		7170 Park, Garden			1
		9715 Industrial Area			6
		...			
	...				
*A*2-ATKIS	2000 Residential Areas	2100 Built-on-areas	2101, 2111, 2112, 2113, 2114, 2126, 2132, 2134, ... e. g. Residential A., Industrial A.		754
		...			
	4000 Vegetation	4100 Vegetation Areas	4101, 4102, 4103, 4107, 4108, 4111, 4199, ... e. g. Forest, Moor, Farmland	FKT 2730, 9999 VEG 1000	226
		...			
	5000 Water Areas	5100 Water Areas	5101, 5102, 5112, ... e. g. Lake, River	BRG 9997, HYD 1000, OFL 11100, ...	66
		...			
	7000 Regions	7100 Administrative Areas	7101, ... e. g. Administrative Unit		2
		...			

Fig. 2: Data sets for test case *B*: 1 : 50.000 – (*B*1) (*left*) and 1 : 10.000 – (*B*2) (*right*).

cient. Thus, we propose to aggregate neighboring buildings first to derive a comparable geometric representation and calculate the geometric relations to the residential areas subsequently.

4 Derivation of Semantic Relations of Data with Similar Granularity

The set up of semantic correspondences of two different data sets can be done by means of expert knowledge based on the known meaning of the terminology used by the organizations which model and capture the

data sets. For example the expert could establish a semantic correspondence between *A*2-5100 (water areas) and *A*1-4310 (water element) by looking at the descriptions and definitions in the object catalogues. This inference, however, is at present only possible by a human analysis and detailed knowledge about the data. In the general case we cannot assume that expert knowledge is always available and human interaction is feasible. Therefore, the knowledge about semantic correspondences between object classes in different data sets has to be detected. Our method is to identify corresponding semantic object groups automatically by the analysis of geometric characteristics of the instances.

4.1 *Method of Geometric Overlay*

For identifying corresponding instances a geometric overlay of the data sets representing the same geometric extent is done. In the analysis we assume that the data sets are organized in layers each representing an object class. In our data sets the instances inside a layer are modelled in a tessellation, but instances of different layers within a data set can overlay. A typical example are administrative objects that often encompass larger areas and generally overlay all the other objects, so called ''container objects''. Accordingly a simple intersection of the data sets without consideration of further characteristics returns more than one matching candidate to an object, which could cause difficulties for the further analysis. But not only the layer structure, also the geometric discrepancies at the object boundaries themselves, occuring due to capturing by different organisations causes an increasing number of possible matching candidates, because adjacent objects may partially overlap.

For identifying corresponding objects all layers of data set *A*1 have to be intersected with all layers of data set *A*2 and vice versa. The result is a list of intersecting objects including also the relations to ''container objects''. That makes the procedure for large data sets with very different levels of detail or different amount of objects time consuming and inefficient. Therefore, in a preprocessing step all layers that do not overlay at all will be taken out of the analysis to reduce the data quantity for the actual overlay procedure. To this end, the instances on each layer are aggregated to one object, overlaid with each layer in the other data set and then the ratio R_c (similar Bel Hadj Ali 2001) between the intersection area and the union area is determined for every possible object class combination with

$$\mathsf{R}_c = \frac{|A1_i \cap A2_j|}{|A1_i \cup A2_j|} \tag{1}$$

in which *A*1 and *A*2 describe the different data sets and i, j = 1,2, ... are the occurring object classes on the lowest level of these data sets. The value of the ratio R_c ranges from 0 for disjunction to 1 for absolute equality. Using this ratio R_c a prediction about probable and improbable matching partners is not absolutely possible. Only matching candidates with a ratio $\mathsf{R}_c = 0$ are allowed to be dismissed, because even a low ratio does not necessarily mean that the matching is absolutely improbable. Due to different modelling, completeness and / or up-to-dateness of the data sets very different values may occur. By excluding all combinations, which fell below the value $\mathsf{R}_c \leq 0{,}01$ the results of case *A* show a reduction from 95 (5×19 matrix) object class combinations to 32. Among those, four complete *A*2-classes had no corresponding *A*1-classes and are no longer needed to be processed.

To further reduce the number of matching candidates an exclusion of neighbouring or minimal overlapping objects is done by introducing and considering the geometric criterion area, especially the object size and intersection area, in the analysis. In this process the overlay ratios of the size of the objects O and the intersection area I are calculated as follows

$$\mathsf{R}_{O,A1} = \frac{I \cdot 100\,\%}{\mathrm{O}_{A1}} \quad \text{and} \qquad \mathsf{R}_{O,A2} = \frac{I \cdot 100\,\%}{\mathrm{O}_{A2}}. \tag{2}$$

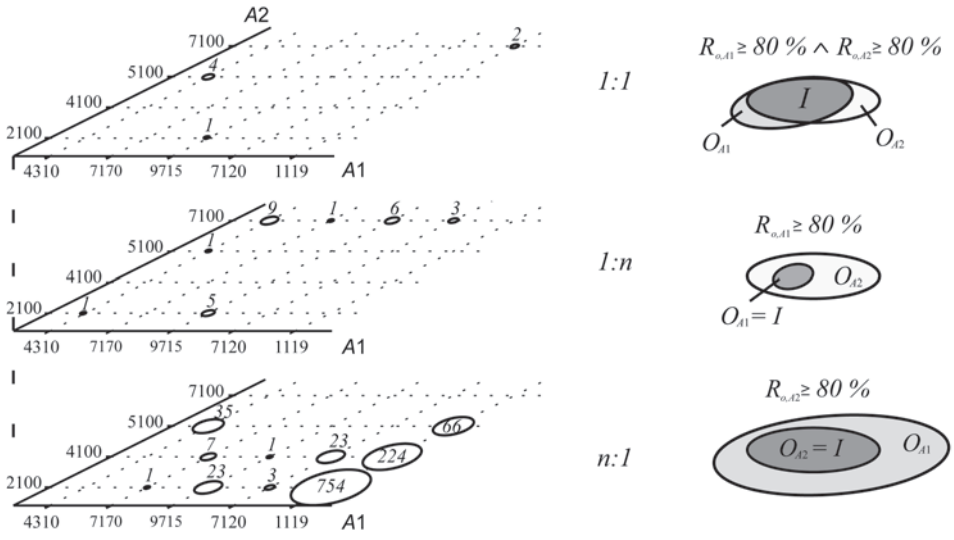

Fig. 3: Results of the overlay method displayed by a frequency matrix: level 1 – contains the matching candidates which fulfill the condition of 1 : 1 relation, level 2 and level 3 contain the candidates which fulfill the condition of 1 : n relation. The figures at the circles indicate the number of possible matching candidates.

Taking small geometric differences into account, we consider objects to match, when the ratio is 80 % or better. In the case of a 1 : 1-relationship following condition has to hold: $R_{o,A1} \geq 80\,\% \wedge R_{o,A2} \geq 80\,\%$.

But if the data sets are differently modelled, i. e. one data set contains more aggregated objects than the other data set, the search for 1 : 1 relations returns only few objects. Therefore the analysis is extended from 1 : 1 relations to 1 : n relations. In that case the condition $R_{o,A1} \geq 80\,\% \vee R_{o,A2} \geq 80\,\%$ has to be fulfilled. As a result a three-level frequency matrix can be prepared that displays the amount of the remaining matching candidates as illustrated in Fig. 3 (left). In order to ensure the readability the class levels 3 of the data set *A*2 are summarized to the respective level 2. On the right side of Fig. 3 the conditions are clarified with the schematic diagrams, that lead to the results presented on the left side.

In the next section the analysis of semantic correspondences from these individual results of the single object classes on the lowest levels is described, which is done up to now in a manual examination.

4.2 *Analysis of Relations for Semantic Integration*

The results of the geometric overlay method on instance level presented in section 4.1 have to reveal the correspondences on class level.

Firstly, clear relations can be derived. For example 1 : 0-disjunction relations ($A1_i \cap A2_j = \varnothing$) are detected between object classes *A*1-4310 and *A*2-4100 respectively *A*1-7120, *A*1-7170, *A*1-9715 and *A*2-5100. A 1 : 1-equivalence relation ($A1_i \equiv A2_j$) between the classes *A*2-7100 and *A*1-1119 is detected, because all (in this case two) objects of both object classes meet the condition $R_{o,A1} \geq 80\,\% \wedge R_{o,A2} \geq 80\,\%$. Additionally to the clear relations also a great amount of matching candidates is found in level 2 and 3 of Fig. 3 representing examples for the 1 : n-inclusion condition ($A1_i \subseteq A2_j$). But not all of these candidates represent true 1 : n relations, because all relations to "container objects" are also included. These large objects (administrative areas) contain nearly all objects of the other data set, i. e. objects of *A*2-7100 and *A*1-1119. These

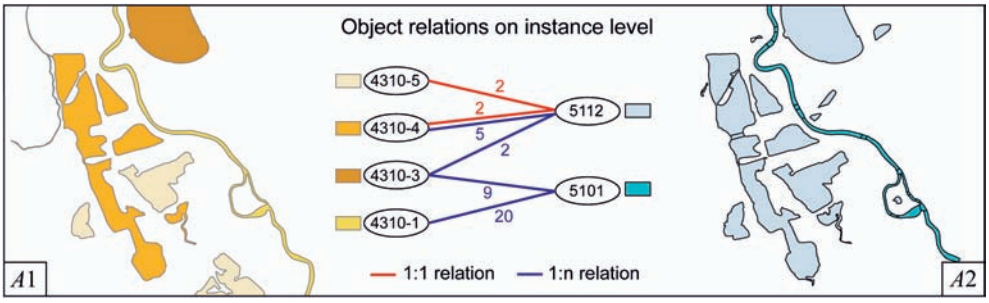

Fig. 4: Object relations on instance level of the object class *A*1–4310 with attributes (left) and *A*2–5101, –5112 (right). The red lines display 1 : 1 relations and the blue lines 1 : n relations with the number.

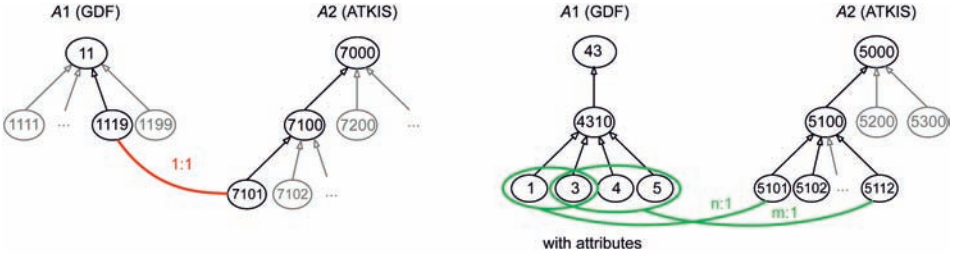

Fig. 5: Resulting semantic relations between object classes of both data sets derived from the instances-relations.

matching candidates have to be dismissed from the further analysis, because they are no real matching partners between corresponding classes. The remaining matching candidates are analysed in more detail. For example between object classes of *A*2-5100 and *A*1-4310 all kinds of instance relations exist, which is an indicator for semantic correspondences. If the hierarchy of the ontologies is regarded on the lowest class level a 1 : n relationships between object classes *A*1-4310 and *A*2-5101, *A*2-5112 exist. In order to be able to identify more detailed relations, in the further analysis additional attributes, i. e. attribute values like proper names or geometric properties will be included. This may lead to better distinctions, because in many cases the relevant semantic information is expressed by these attributes.

Including the available attributes of the object classes *A*1-4310 in the analysis the following relations between the instances can be detected: objects of object class *A*1-4310 with attribute values 1 (*A*1-4310-1) and 3 (*A*1-4310-3) match to objects of *A*2-5101 and *A*1-4310-3, -4 and -5 match to object class *A*2-5112. Whereas the former has only 1 : n relations the latter contains also 1 : 1 relations that have to be more important to the evaluation than 1 : n relations. Object class *A*2-5102 has no matching partners at all in this test area. The object class *A*1-4310-3 matches to *A*2-5101 and also to *A*2-5112 (see Fig. 4).

These derived semantic relations between the object classes of both data sets can be represented with the help of the ontology hierarchies. Therefore in Fig. 5 the results for the 1 : 1 equivalence relation between the lowest levels of *A*1-11 and *A*1-7000 and the relations between level 2 of *A*1-43 with attributes and level 3 of *A*2-5000 are presented; from the latter, also a higher level correspondence between *A*1-4310 and *A*2-5100 can be derived.

Beside the existing attributes also derived attribute values of the geometry can improve the analysis of the semantic corre-

spondences in the future. As geometric properties would be conceivable compactness, elongation, rectangularity, which are typical for certain objects. For example buildings in general have parallel borders and right angles whereas lakes or ponds have in most cases a round, organic outline. Furthermore the integration of context-dependent relationships will be considered. For example an object X of *A*1 corresponds to an object Y of *A*2 only when X is in context C1 and Y in context C2. In another context, X would correspond to another class.

5 Attribute Transfer of Data with Dissimilar Granularity and Up-to-dateness

5.1 *Aggregation and Matching*

For identifying corresponding objects in data with dissimilar granularity a simple geometric overlay of the data sets does not always lead to the desired result. Objects in large scale data sets have more geometrical details than objects in the small scale. Objects (in our case buildings) in border areas in the large scale data set may clearly belong to the group of buildings in their local environment although they might not fall within the corresponding settlement area in the smaller scale. This is even more true, as in our case due to temporal differences, new buildings are present in the detailed data set, that are not reflected by residential areas in the smaller scale data. In this test case the knowledge about the semantic corresponding object classes is presumed to be known, namely the aggregation relationship of buildings from *B2* and settlement areas in *B1*. As described above, the goal of the analysis is the identification of the individual corresponding objects in order to accomplish a transfer of semantic information (e. g. name attributes) from *B1* to *B2*. To establish the relationships between these objects a three-step-approach is proposed. With no additional information available (e. g. road network) only geometric reasoning can be used for the analysis.

The first step involves a distance-based aggregation of objects of data set *B*2. By means of this procedure a generalization is accomplished, in order to overcome the granularity differences. For this purpose a buffer is computed around the buildings. The determination of the buffer distance is crucial. The most suitable buffer distance can be determined depending on the building density. To this end clustering approaches can be used (e. g. Anders 2004). The aggregation of the building objects of *B*2 is illustrated in Fig. 6 (left). In this first experiment, we used a fixed buffer size of 10 m derived from experiments.

The second step involves the search for correspondences of the buffered building polygons (bps) and the residential areas (ras) of *B*1. We used a distance-based approach by determining Voronoi cells (Vcs) of the residential polygons of data set *B*1. These polygons, as illustrated in Fig. 6 (middle) divide the whole area into seamless cells and every cell is the best range corresponding to a 1 : 50.000 ra object. In this figure the red lines are Voronoi polygon boundaries and the pink blocks are residential objects in 1 : 50.000. In the third step a geometric overlay with the new bps of 1 : 10.000 and the Vcs of 1 : 50.000 is done (see Fig. 6, right).

To analyze the correspondences of bps and Vcs two methods can be used. One method is to calculate the centroid point of the bps. If the point is within one Vc, we decide that the bp belongs to this cell polygon, and thus to the corresponding ra. In Fig. 7 (left) the centroid points of bps and Vcs are displayed. In general case the mapping is clear, but at the border of the Vcs the result is not always satisfying. The bp in the red circle will be mismatched to the Vc above.

To improve the result, another method is to decide by means of the degree of overlap. If the larger part of the bp is within a Vc, the polygon is considered to belong to this cell. That means for the example displayed in Fig. 7 (right), that the polygon must be assigned to the lower cell, because the larger part displayed in dark grey is located there.

Fig. 6: Three-step-approach for establishing relationships between objects of data sets with dissimilar granularity; Creation of buffer polygons (bps) (orange) (*left*), Voronoi cells (Vcs) (red) of 1 : 50.000 (*middle*) and Result of overlay of bps and Vcs (*right*).

Fig. 7: Mapping results of the centroid point method (*left*) and the degree of the overlapping areas (*right*). The blue crosses are the centroids of the orange buffer polygons.

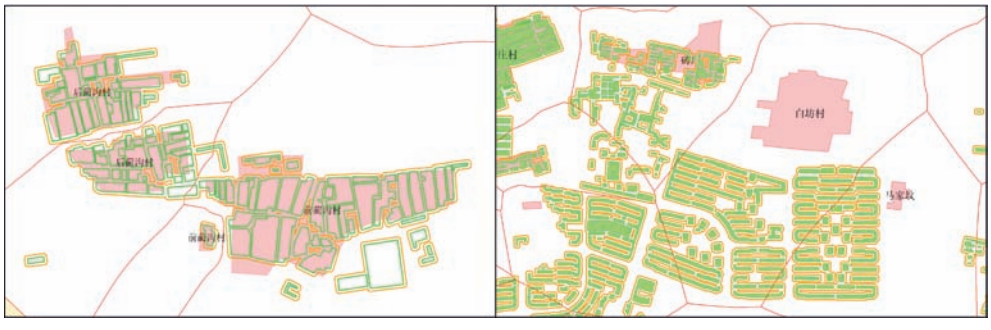

Fig. 8: Wrong assignment of corresponding residential areas due to aggregation (*left*); different spatial situation due to different acquisition time (*right*).

With this matching of bps to Vcs the buildings inside the bps are also assigned.

Furthermore, the degree of overlap between bps and Vcs and ras can give an indication to the quality of the correspondence found. The following analysis steps can be performed: first of all, the degree of overlap of bp and Vc gives an indication to the containment in a cell and also to the fact, whether the buildings have been correctly aggregated. If the degree of overlap of bp and Vc is smaller than 50 % this means that

half of the bp-area is located in one or more other Vcs. This could indicate that the buffer distance had been set too large (see Fig. 8, left). The degree of overlap of bp and ra indicates, how well these two areas fit; the closer the value is to 1, the better is the overlap and thus also the reliability of the assignment of the attributes.

For rapidly changing areas, this approach can not identify corresponding buildings correctly because the spatial situation has changed dramatically: a lot of new buildings have been built, which cannot reliably be assigned to the old residential areas. However, the measure of degree of overlap between both bp and Vc and bp and ra can give an indication and help to perform a visual inspection of the quality of the automatically determined correspondences easily.

5.2 *Attribute Transfer*

The results of the matching are corresponding building objects and small scale residential objects. The residential areas of 1 : 50.000 have the attribute place name, which can now be transfered to the buildings 1 : 10.000, in order to enrich their mere geometry. The transfer of the attributes can be achieved by joining their attribute tables. The quality of the overlap can be taken into account, in order to qualify the reliability of the transferred attributes.

6 Discussion of Results and Future Work

The presentation describes ongoing work on semantic data interoperability. The need of integrating data sets of different origin and different granularity is evident, especially for a data reuse. The work on semantic integration is in a very early stage. In the future, there will be a focus mainly on the improvement of the identification of semantic correspondences by an extended analysis, mainly by automating this process. To this end, additional attributes to the comparison of the data, like shape parameters, will be introduced. Furthermore the analysis of correspondences between the object classes on lowest level must be extended, because equivalent concepts may be expressed at different levels in the hierarchy of the ontology. The analysis will be automated using association rules from Data Mining (Agrawal & Srikant 1994).

The matching of objects of two data sets with dissimilar granularity is more successful, if they share a similar up-to-dateness. In the analysed example of an urban or suburban area in China the changing of the areas was so rapid and far reaching, that the proposed three-step-approach was not suitable for some parts of the test area; it seems to be more appropriate for rural areas, where changes do not happen so often. To improve the results in urban and suburban areas the introduction of additional information about possibly partitioning objects, e. g. road and river networks, is essential. Another issue is a refined aggregation process, which should take the structural similarity of the buildings (large industrial vs. small residential buildings) into account.

Acknowledgements

This research is funded by the German Science Foundation (DFG) and National Nature Science Foundation of China (NSFC): 40620130438. The support is gratefully acknowledged.

References

Agrawal, R. & Srikant, R., 1994: Fast algorithms for mining association rules. – Proceedings of the 1994 International Conference for Very Large Data Bases (VLDB'94), 487–499.

Anders, K.-H., 2004: Parameterfreies hierarchisches Graph-Clustering-Verfahren zur Interpretation raumbezogener Daten. – Dissertation, Fakultät Luft- und Raumfahrttechnik und Geodäsie der Universität Stuttgart, Stuttgart.

Bel Hadj Ali, A., 2001: Positional and shape quality of areal entities in geographic databases: quality information aggregation versus measures classification. – Proceeding of ECSQARU'2001 Workshop on Spatio-Temporal Reasoning and Geographic Information Systems.

Bishr, Y., 1997: Semantic aspects of interoperable GIS. – Wageningen Agricultural University and International Institute for Aerospace Survey and and Earth Science (ITC), Enschede.

Duckham, M. & Worboys, M., 2005: An algebraic approach to automated geospatial information fusion. – International Journal of Geographical Science **19** (5): 537–557.

Dunkars, M., 2003: Matching of Datasets. – Proceedings of the 9th Scandinavian Research Conference on Geographical Information Science, p. 67–78.

Fonseca, F., Câmara, G. & Monteiro, A.M., 2006: A Framework for Measuring the Interoperability of Geo-Ontologies. – Spatial Cognition and Computation **6** (4): 309–331.

Kokla, M., 2006: Guidelines on Geographic Ontology Integration. – ISPRS Technical Commission II Symposium, p. 67–72.

Mustière, S., 2006: Results of experiments on automated matching of networks at different scales. – Proceedings of the ISPRS Workshop on Multiple Representation and Interoperability of Spatial Data, 92–100.

Rodriguez, A. & Egenhofer, M., 2003: Determining Semantic Similarities Among entity classes from different ontologies. – IEEE Transactions on Knowledge and Data Engineering, **12** (2): 442–456.

Samal, A., Seth, S. & Cueto, K., 2004: A feature-based approach to conflation of geospatial sources. – International Journal of Geographic Information Science **18** (5): 459–490.

Uitermark, H.T., 2001: Ontology-based geographic data set integration. – PhD Thesis, Computer Science Department, University of Twente, Enschede, The Netherlands.

Volz, S., 2005: Data-Driven Matching of Geospatial Schemas. – Cohn, A.G., Mark, D.M. (eds.): Spatial Information Theory. Proceedings of the International Conference on Spatial Information Theory (COSIT '05), Lecture Notes in Computer Science 3693, Springer, 115–132.

Addresses of the Authors:

Dipl.-Ing. Birgit Kieler, Prof. Dr.-Ing. habil. Monika Sester, Leibniz Universität Hannover, Institut für Kartographie und Geoinformatik, Appelstraße 9A, D-30167 Hannover, Tel.: +49-511-762-3588, Fax: +49-511-762-2780, e-mail: {Birgit.Kieler, Monika.Sester}@ikg.uni-hannover.de

Msc. HaiQing Wang, Dr. Jie Jiang, National Geomatics Center of China, 1 Baishengcun, Zizhuyuan, Beijing, 100044, Tel.: +89-10-68483218, Fax: +86-10-68424101, e-mail: {whq, jjie}@nsdi.gov.cn

Manuskript eingereicht: Mai 2007
Angenommen: Juni 2007

Photogrammetrie • Fernerkundung • Geoinformation 6/2007, S. 459–468, 7 figs.

Automatic Extraction of Textures from Infrared Image Sequences and Database Integration for 3D Building Models

Ludwig Hoegner, Holger Kumke, Liqiu Meng & Uwe Stilla, München

Keywords: 3D Reconstruction, Markerless Orientation, Visualization, Sensors, Infrared, Building Model, Database, GIS, CITYGML

Summary: This article presents an approach for extraction of façade textures from low resolution infrared (IR) image sequences and their integration into a database. Two strategies are presented for matching the images and the given 3D model: the first one is based on the determination of the camera parameters using three corresponding points, and the second strategy involves the estimation of planes in the image sequence via computation of homographies. Textures are extracted using an algorithm based on the principles of ray casting to generate partial textures for every visible surface in every single image of the sequence. Furthermore, textures generated from different images yet belonging to the same façade are combined. Different combination strategies are briefly described. The resulting textures are integrated into a GIS database in a format that is conformal to CityGML. Integration and visualisation techniques for those textures are discussed with the intension to extend the CityGML format for multi-textural purposes.

Zusammenfassung: *Automatische Extraktion von Texturen aus Infrarot Bildsequenzen und Datenbankintegration für 3D Gebäudemodelle*. Dieser Beitrag behandelt die Extraktion von Fassadentexturen aus niedrig aufgelösten Infrarot (IR) Bildsequenzen und deren Integration in einer GIS-Datenbank. Für das Matching von Bildern und dem gegebenen 3D-Modell werden zwei Verfahren vorgestellt: die Bestimmung der Kameraparameter mittels drei korrespondierenden Punkten in einem Bild und dem Modell, und die Schätzung von Ebenen in Bildsequenzen durch Homographie. Die Texturen werden unter Verwendung eines Ray Casting-Algorithmus extrahiert, um partielle Texturen für jede sichtbare Oberfläche jedes Einzelbildes der Sequenz zu erzeugen. Die Texturen, die aus verschiedenen Bildern der Sequenz erzeugt werden und zur selben Fassade gehören werden anschließend kombiniert. Verschiedene Strategien für die Kombination werden kurz vorgestellt. Die gewonnenen Texturen werden in eine GIS Datenbank integriert, die auf CityGML basiert. Es werden Integration und Visualisierung diskutiert, um CityGML für die Anforderungen von Mehrfachtexturierung zu erweitern.

1 Introduction

Today, the acquisition and exploitation of thermal imagery is gaining considerable attention. Infrared (IR) images are acquired from different platforms and used for different applications. For instance, satellite images are the common basis for vegetation monitoring (Quattrochi & Luvall 1999), the analysis of urban heat islands (Lo & Quattrochi 2003), or fire detection (Siegert et al. 2004). Airborne IR-systems are applied for detecting stationary vehicles (Stilla & Michaelsen 2002, Hinz & Stilla 2006) and moving objects (Kirchhof & Stilla 2006) or exploration of leakages in district heating systems (Koskeleinen 1992). On the other side, the irradiation of façades is typically recorded with ground cameras in order to get insight into a building's thermal behaviour (Klingert 2006).

Urban areas reveal complex thermal characteristics due to the mutual influences of man-made objects like buildings, roads etc. Aside from object materials, the whole 3D

1432-8364/07/2007/0459 $ 2.50
© 2007 DGPF / E. Schweizerbart'sche Verlagsbuchhandlung, D-70176 Stuttgart

geometric outline of the recorded scene plays an importand role for the heating caused by the sun and radiation and irradiation between buildings. In many urban studies, 3D information has not been exploited for the interpretation of the acquired scenes.

Nichol & Wong (2005) applied simple prismatic building models to investigate the influence of building geometry on urban heat effects. Spaceborne thermal images are used to texturize a digital terrain model. For every building roof an averaged temperature of the corresponding image area is determined. The temperature of building walls are guessed from the roof temperature and the temperature of the surrouding ground. Due to the coarse resolution of the satellite images, a detailed investigation of building surfaces is not possible.

This paper investigates strategies for a detailed texture mapping of building façades to support a thermal inspection. Typically, thermal inspections of building façades are carried out in single images from the observed objects. Larger building parts require several images to be analysed. An integral way of viewing buildings recorded from different images is difficult without combining those images. This problem is getting worse, when images from different cameras or views need to be combined and stored for further processing without any geometric reference.

The main challenge for the extraction and combination of textures refers to the estimation of the exteriour orientation of the camera for the projection of the 3D model into the image plane. In a second step, combined textures have to be integrated in a given database.

1.1 *Automatic Extraction of Textures*

Several techniques for estimating the exteriour orientation have been reported in the literature. A suitable solution depends on the number of images and the number of point correspondences between the IR images and the projected 3D model. The estimation of exterior orientation from a single image works with at least 3 correspondences (3-point algorithm) between image and model (Haralick et al. 1994). Techniques for 4- and 5-point estimation are elicited by Quan & Lan (1999) and Triggs (1999). For 6 and more correspondence points the Direct Linear Transformation (DLT) can be applied (Triggs 1999). Generally, the distribution of the points in object and image space has to be observed to ensure a certain accuracy of the estimated parameters or a solution at all.

Due to the small field of view and the limited spatial resolution of the IR images together with the low level of detail of the given building model, only few point correspondences between IR image and 3D model can be identified. Alternatively, one may first co-register all images of the sequence via tie points (as in stereo photogrammetry (Longuet-Higgins 1981)) and then establish the correspondence of the whole image strip with the 3D model. The exterior orientation of the individual images can then be inferred from the transformation parameters. For façade structures approximately forming a 3D plane, algorithms for homography estimation can be applied to detect planes in image pairs. This enables the direct determination of the orientation of the camera in relation to these planes (Hartley & Zisserman 2000)

1.2 *Database Integration of Textures*

3D city models are widely used in different applications like disaster management, city planning or transmitter placement for mobile communications. The enrichment of building models with additional data related to parts of the building geometry is subject of ongoing research. There exist many proprietary data formats for 3D city models which support textured geometric models and factual data as independent objects, but they do not support factual data assigned to a building model (Kolbe & Gröger 2003). This drawback was first addressed in CityGML (Kolbe et al. 2005, Dörschlag 2006) which is an exchange format based on the Extensible Mark-Up Language XML (W3C 2001). The semantical, topological

and geometrical models were developed to exchange all fundamental building-related informations. Nevertheless, this standard supports only the common data of a building like address, storey height, roof type, year of construction, etc. Here, we propose an extension scheme which allows the integration of multiple image data (e. g. multispectral, multi-aspect, multi-temporal) and additional data (e. g. texture resolution distribution, texture completeness, meteorological data, cartographic signatures).

2 City Model and Data Acquisition

A part of the untextured 3D model selected as test object from the database is shown in Fig. 1. For enriching the building database with thermal textures, IR images were recorded with 50 frames per second at a sampling grid of 320×240 pixels. Two cameras capture the midwave (3–5 µm) and longwave (8–12 µm) infrared band (see the left and middle camera in Fig. 2 and resulting images in Fig. 3). An additional video camera captures the visual information for verification purposes (see the right camera of Fig. 2). The camera position is recorded using a GPS. This is only used, however, for the initial position of the virtual camera since narrow streets and high buildings led to positional errors for lack of sufficient GPS signals. The building models are stored in a database according to the CityGML exchange format for 3D models. Meteorological data are gathered to estimate temperature values from the measured radiation. Additionally, a timestamp was stored for synchronization of meteorological data, camera position and the infrared image sequences.

3 Pose Estimation, Texture Extraction and Texture Combination

3.1 *Position Estimation and Matching between IR Images and 3D Models*

All camera parameters must be known in order to project the 3D model onto the image plane. The interior orientation of the camera is determined by a off-line calibration. The *approximate position of the camera* is recorded during the image acquisition. *Pan, tilt and roll angles of the camera* are estimated from ground control points given by the vertices of the 3D model. Two strategies are described in this section to automate the assignment of correspondence between the 3D model and IR images.

Due to the small field of view, the low spatial resolution of the IR images and the low level of detail of the given building model, only few point correspondences between IR image and 3D model can be identified (Fig. 4a). A straightforward solution is thus the application of the 3-point algorithm proposed by HARALICK et al. (1994). For

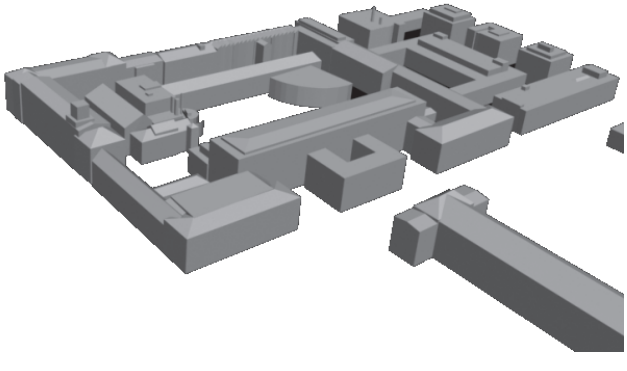

Fig. 1: 3D model of the test site.

Fig. 2: Camera system.

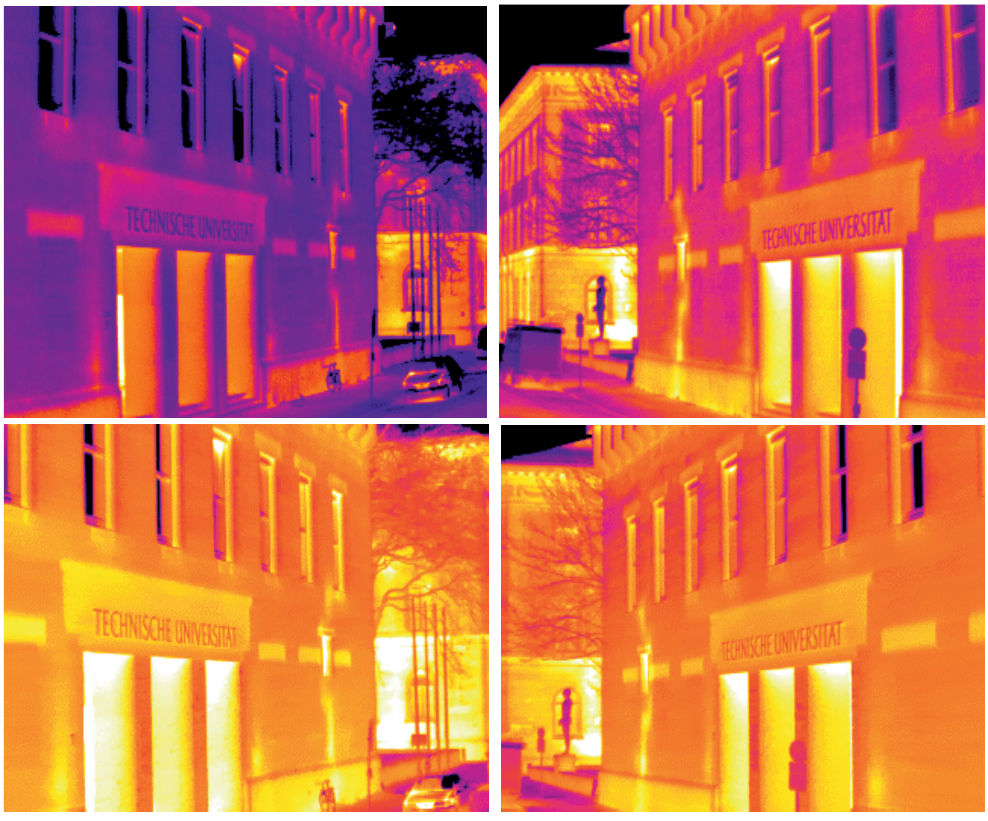

Fig. 3: from upper left to lower right: forward and backward view of midwave IR camera and longwave IR camera.

a given camera position, the image with the corresponding time-code is selected, first. Then, common interest operators are applied to extract salient points and edges. Long horizontal and vertical edges are supposed to be façade edges, thereby including some tolerance values to accommodate perspective effects. Intersections of horizontal and vertical edges finally yield the 2D coordinates of façade corners.

As three points are known in 3D coordinates and should correspond to three points given in 2D coordinates, an estimation of the exterior camera orientation is possible. Every visible polygon is projected onto the image plane transforming its vertices from 3D camera coordinates to pixel coordinates and clipped using Sutherland-Hodgeman polygon clipping. The depth values and texture coordinates for all image pixels covered by the polygon are inferred from the plane equations of the polygon. Every pixel of the final projected scene image is assigned to the smallest depth value from all projected polygon images (z-buffer; Catmull 1974). To this end, an image with polygon ID, depth value and texture coordinates stored for every pixel is available for the projected scene. The intensity values of the IR image are now assigned to the corresponding texture and surface stored in pixels of the projected scene. Missing camera positions of frames are interpolated from frames with known position.

This strategy is feasible if at least three vertices of the model are visible in the image of the virtual camera and the corresponding corners can be found in the IR image. It

might happen that, due to the given characteristics of the IR camera and the situation of image acquisition, the IR image shows only one or two façade corners so that there remain only two edges of the model for finding corresponding edges in the IR image. Two edges are not sufficient to estimate a unique camera orientation, as there are different combinations of camera position and viewing angle that can achieve the overlay of the two lines in the IR image and the model projection.

An alternative strategy, that circumvents the problem of missing vertices in the IR images, is the following:

Instead of calculating the position for individual key frames, the complete image sequence is incorporated into the image orientation procedure, i. e., the relative orientation between all consecutive image pairs is calculated first, before the whole strip is matched to the 3D model. This process can be simplified, since, for homogeneous façade structures that approximately form a plane, homography estimation is suitable to detect planes in image pairs and to infer the orientation of the camera in relation to these planes (Hartley & Zisserman 2000). Since many buildings have an approximately planar façade, the assumtion that significant façade structures lie also in a plane is often valid.

Hence, we establish the correspondences between two subsequent images through Harris interest points and employing a homography matrix H. The estimation of H involves a RANSAC procedure (Fischler & Bolles 1981) to cope with outliers and local extrema in the objective function. For the first image pair, the initial (GPS) camera position is used to estimate a façade plane that is close to the corresponding 3D model surface. For all subsequent image pairs (with frame distance d), a set of homographies is calculated to enable the estimation of the camera trajectory relative to the façade plane. Finally, the trajectory is smoothed by averaging the planes of the homographies. This strategy works well for homogeneous images with only one façade covering most of the image.

3.2 *Texture Extraction from IR Images*

Depending on attitude and trajectory of the camera, some building parts might be invisible due to self occlusion or occlusion from other objects. Thus, a search algorithm determines the visible surface points corresponding to the pixels of the image of the virtual camera:

First, a plane equation is calculated for every surface to receive the depth value and texture coordindates. This plane is then used for ray casting (Foley 1995), where every pixel is projected into the scene and assigned to an intersection point of that particular plane, which has the smallest depth value along the ray from the camera through the pixel. For the intersection point of this plane, the texture coordindates are interpolated from the texture coordinates of the vertices of the plane. The texture coordinates count from $(u, v) = (0,0)$ at the left lower vertex of the surface and are going up to $(u, v) = (1,1)$ at the upper right vertex. The ID of the intersected surface and the texture coordinates of the intersection point are returned to the pixel of the infrared image.

After this 3D to 2D transformation of the model surfaces into the image plane of the IR image, a 2D transformation is carried out to transform the IR image pixels to texture coordinates of the model surfaces. At the end, the intensity values of the IR image have been transformed to texture coordinates of points on the model surfaces.

Another 2D transformation is necessary to transfer the texture points to pixel coordinates for the surface textures. At first, the individual pixel coordinates of the texture are transferred into texture coordinates of the surface. Then, their interpolated values are calculated through a bilinear interpolation. If a pixel has only three surrounding points in u and v direction, the pixel value is interpolated using barycentric coordinates. If the pixel is outside the triangle defined by the three surrounding points, it is outside the visible part of the façade. Pixels with only two or less surrounding texture points also are outside the visible part of the façade and left in black (Fig. 4b).

3.3 *Combination and Storage of Extracted Textures from IR Sequences*

For assembling an entire texture of a façade, its partial textures extracted from the individual images need to be combined, i. e., the intensity values of the combined final texture have to be interpolated from the values of the parital textures. To this end, two methods can be applied (see Fig. 5):

The first method transforms the texture coordinates of every projected IR image into a partial texture image, which in turn is used to interpolate the pixel values of the entire texture (cf. Fig. 5a). This allows to search for corresponding points and edges in the partial textures to correct for displacements. The displacements are caused by inaccuracies in the camera parameter estimation. However, besides residual displacements and inaccurate initial camera position more problems remain: The low resolution of the IR images leads to blurred edges and eventually to jittering edges between adjacent images. In addition, the small field of view and the small distance to the building allows only few edges to be seen, so that this positional inaccuracy is hard to correct.

The second method extracts the texture coordinates directly from the image sequence and combines them into one set of texture coordinates for the purpose of interpolating the pixel values of the entire texture (cf. Fig. 5b). Although it is not possible to correct displacements with this method, a higher interpolation accuracy can be achiev-

Fig. 4: Steps for generating a single texture: a) IR image, b) extracted single partial texture, c) combined final texture. The color table represents the intensity values of one IR band.

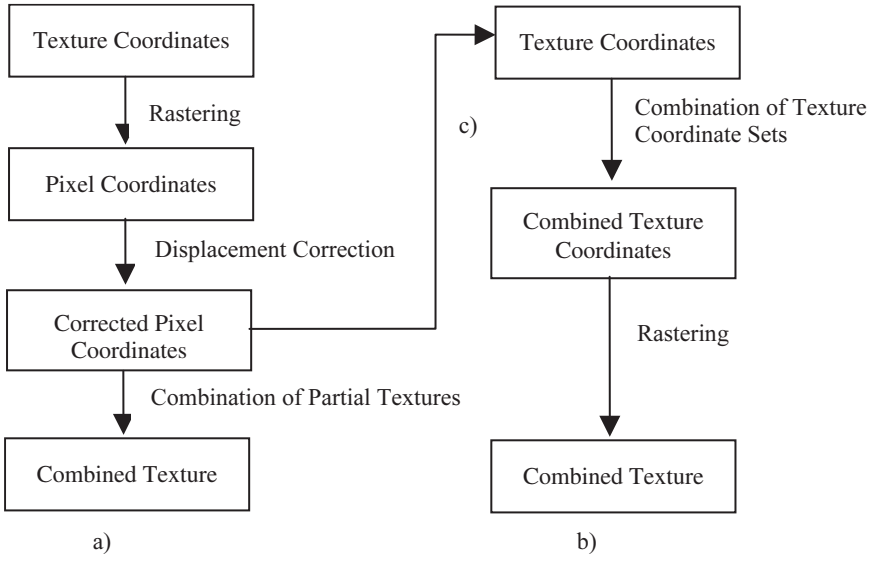

Fig. 5: Texture processing strategies: a) Combination of partial textures, b) Combination of texture coordinate sets, c) Combined strategy using the corrected partial textures to correct the texture coordinate sets.

ed because the original positions of the intensity values estimated during the texture extraction are used for the interpolation and no intermediate step using partial texture images is necessary.

Both methods are combined in the approach in this paper. At first, the partial texture images are used to find displacements between the partial textures like in the first method (cf. Fig. 5a). These displacements are transformed from pixel coordinates to texture coordinates and used to correct the relative positions of the texture points of the partial textures (cf. Fig. 5c). Then, the pixel values of the entire texture are interpolated from the complete set of texture points (cf. Fig. 5b, second method) using bilinear interpolation and barycentric interpolation as described in section 3.2. A combined texture is shown in Fig. 4c.

Before the texture image is stored, the intensity values are quantified for a predefined bit depth (e. g. 8 bit or 24 bit). The absolute temperatures of surfaces are finally calculated by transforming the intensity values to temperature values using meteorological measurements as calibration data. This correction can be realised either on the partial textures or on the interpolated pixel values of the combined texture image.

Some simplifications can be made for a single-pass image sequence in case of a forward-looking oblique view and a moving camera with constant angle between the viewing direction and the façade: If the frame rate is high compared to the velocity of the movement, the differences of two adjacent images of the sequence are very small. For this reason, corresponding points in both surface textures can be determined within a small search area. Then, after matching two partial textures, the intensity values for the combined texture are calculated. It has to be observed that, for an image sequence acquired in the aforementioned way, the partial surface textures have different resolutions. The resolution decreases in the viewing direction of the camera caused by the perspective view (cf. Fig. 6a). During forward recording, each subsequent partial texture has a higher spatial resolution for the overlapping area than its preceding texture, but does not show the complete part of the preceding texture (cf. Fig. 6b).

In an initial step, the first partial surface texture is copied to the entire texture. The second partial surface texture is then copied to the entire texture and overwrites the pixels copied from the first texture that are copresent in the second one. For the first texture, only those pixels not belonging to the second texture remain in the entire texture. This procedure is continued for all partial surface textures of the image sequence. At the end, every part of the entire texture gains the highest resolution of this part in all images from the input image sequence (cf. Fig. 6c).

4 Integration of Thermal Data into CityGML

CityGML is based on the Extensible Mark-Up Language and can be visualised by the Aristoteles 3D Viewer and the commercial program LandXplorer. It supports the extension of project-specific schemes with help of namespaces. In this work the namespaces *xs* and *xAL* are respectively adopted for urban data and address data. An extended namespace *ir* is defined for infrared data. The 3D model of the Technische

Fig. 6: a) Resolution image of a single partial texture (compare Fig. 4a, 4b; bright : high resolution, dark: low resolution), b) Resolution image of a combination of several partial textures, c) Resolution image of the combined texture.

Universität München conforms to the CityGML structure and contains building geometries at the Level of Detail 2 according to the definition in Sig3D.

CityGML describes the building surfaces in terms of boundary representation BRep (Mantyla 1988). BReps are also referred to as ring polygons or linear rings. The material properties are described according to the same standards as VRML, X3D and COLLADA. Among others, CityGML supports the description of radiometrical values that are important for the derivation of the Bidirectional Reflectance Distribution Function (BRDF) of isotrope materials.

The scheme ir is stored in an instance file, with its semantical structure being defined in the *infrared.xsd* file. It contains (i) the meteorological data, (ii) all generated textures from the different wave lengthes that can be optionally visualised, (iii) the resolution textures that can be used to monitor the texture quality and (iv) time stamp. The *EXIF* format stores meta data of camera-specific information in the header of the *TIFF* file, like field of view, wave length, camera type etc.

For each texture, its corresponding meteorological data at the recording time can be used to derive the absolute temperatures from the stored texture intensity. This allows, for example, the inverstigation of the influence of weather conditions on the thermal behavior of building façades. A completeness factor for each façade texture is included in the *ir* scheme for the purpose of showing the coverage percentage of the surface by its infrared texture. With the exception of this completeness factor no information about the position of the missing parts is stored. Furhtermore the texture resolution is stored for every pixel in a resolution image (cf. Fig. 6c). CityGML supports the data transfer between the Oracle 9i database and the visualisation software.

The open source program Aristoteles 3D Viewer developed by Dörschlag & Drerup (2007) and written in Java 3D is adopted for the visualisation of thermal information. This software allows retrieval, manipulation and storage of CityGML files in a 3D environment. Aristoteles supports standard schemes. New plug-in tools can be created to visualise surfaces with multi-textures, temporal changes in texture sequences, geometrical structures, and their semantic attributes.

5 Discussion

We have introduced a solution to automatically extract textures from low resolution IR image sequences for a given 3D building model and to integrate these textures into a database. The estimation of the exterior orientation of the virtual camera for the projection from GPS data is insufficient. The two possible ways for correcting the orientation are limited to certain environmental parameters. The 3-point camera correction fails in case of a low number of corresponding points in the 3D model view and the IR image. The homography-based orientation estimation works pretty well for large plane surfaces but has revealed some weaknesses for heterogeneous façades.

To get a texture for visualisation, a raster image has to be generated which may cause errors in the position of the measured values and the intensity value itself. The research for the influence of the type of combination of the textures, which can be conducted before and after rasterization and before or after the quantification of the values for the texture image, remains a substantial task in future work.

6 Outlook

Midwave thermal infrared and longwave thermal infrared normally reveal similar but not exactly the same thermal structures. Structures reflecting the sunlight are better visible in midwave infrared. However, this spectrum is strongly affected by weather changes. For example, the measured radiance changes immediately as soon as the sun is covered by clouds. This fact has inspired us to investigate the change of thermal behaviour over time with combined infrared datasets. The recorded scenes of our test site

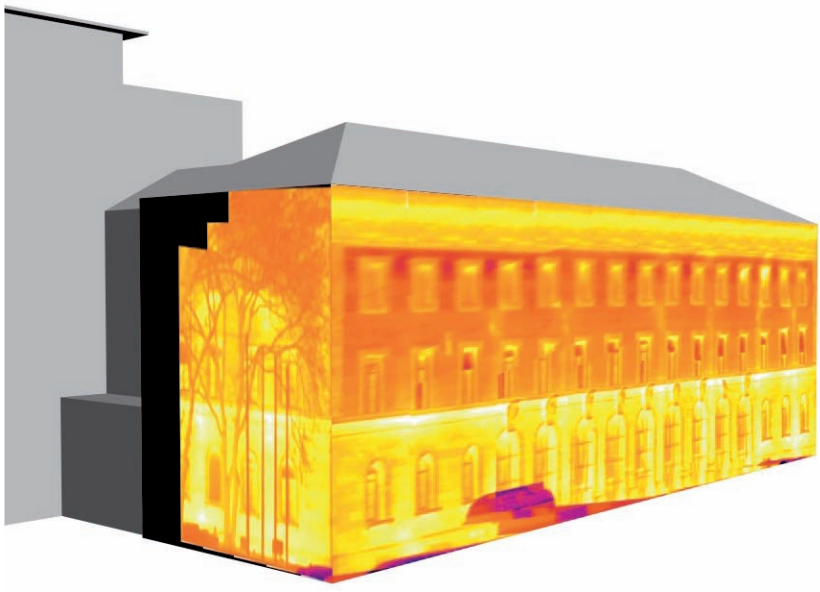

Fig. 7: Final texture on 3D building model.

in different wavelengths at different times and under different viewing angles allow us to compare the different behaviours of geometric structures and surface attributes (e. g. heating pipes, surface temperature) in midwave and longwave infrared. Whenever a thermal structure is detected in one texture, it can be inspected in other textures of the same façade from different times and wavelengths, even if it may be invisible in other structures. This allows verifying the detected structure and understanding its thermal behaviour.

In the current work, edges and vertices of the 3D model were used to find correspondences to image structures. The existing 3D model corresponds to the LOD2. That is, only building edges and roof structures are represented, but structures on the façade are not included. This causes problems in finding correspondences between the model and the image because only a small part of the objects is visible in the image due to the small field of view and short distance to the observed objects. One possible solution is to combine 3D model parts with surface structures detected from available textures of both visible and infrared domains so that additional edges and points for the matching can be generated. In particular, we will investigate the feasibility of using small structures extracted from textures of visible domain to refine the position estimation and matching process.

The existing techniques of 3D computer graphics that have been mainly focused on geometries and textures from visible domains need to be extended to include special symbolization and animations to display thermal information. We also attempt to highlight the time dependent thermal behaviour of small structures in different textures of the same façade. Instead of striving for the visualization of classified façade temperature based on the general knowledge about the typical material radiation by different weather condition, we concentrate on highlighting the distinctive characters of thermal structures such as heat bridges, heat insulation, and windows in relation to their surrounding.

Acknowledgements

The authors thank Dr. Clement, Dr. Schwarz, and Mr. Kremer of FGAN-FOM,

Ettlingen, for equipment and their assistance during the recording campaign.

References

Aristoteles 3D Viewer. – http://131.220.71.208/index.php/Aristoteles (accessed 31.03.2007).

Catmull, E.E., 1974: A Subdivision Algorithm for Computer Display of Curved Surfaces. – Phd dissertation, University of Utah, Salt Lake City.

CityGML – Exchange and Storage of Virtual 3D City Models. – http://www.citygml.org/ (accessed 15.05.2007).

Extensible Mark-Up Language. – http://www.w3.org/XML/ (accessed 10.05.2007).

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

Foley, J.D., 1995: Computer Graphics: Principles and Practice. – Addison-Wesley.

Haralick, R.M., Lee, C.N., Ottenberg, K. & Nolle, M., 1994: Review and Analysis of Solutions of the 3-point Perspective Pose Estimation Problem. – International Journal of Computer Vision **13** (3): 331–356.

Hartley, R.L. & Zisserman, A., 2000: Multiple View Geometry in Computer Vision. – Cambridge University Press.

Hinz, S. & Stilla, U., 2006: Car detection in aerial thermal images by local and global evidence accumulation. – Pattern Recognition Letter **27:** 308–315.

Kirchhof, M. & Stilla, U., 2006: Detection of moving objects in airborne thermal videos. – ISPRS Journal of Photogrammetry and Remote Sensing **61:** 187–196.

Klingert, M., 2006: The usage of image processing methods for interpretation of thermography data. – 17th International Conference on the Applications of Computer Science and Mathematics in Architecture and Civil Engineering, Weimar, Germany, 12–14 July 2006.

Kolbe, T., Gröger, G. & Plümer, L., 2005: CityGML – Interoperable Access to 3D City Models. – In: Oosterom, P. van, Zlatanova, S. & Fendel, E.M. (Eds.): Proceedings of the 1st International Symposium on Geoinformation for Disaster Management. – Springer.

Koskeleinen, L., 1992: Predictive maintenance of district heating networks by infrared measurement. – Proceedings SPIE **1682:** 89–96.

Lo, C.P. & Quattrochi, D.A., 2003: Land-Use and Land-Cover Change, Urban Heat Island Phenomenon, and Health Implications: A Remote Sensing Approach. – Photogrammetric Engineering & Remote Sensing **69:** 1053–1063.

Longuet-Higgins, H.C., 1981: A computer algorithm for reconstructing a scene from two projections. – Nature **239:** 133–135.

Mantyla, M., 1988: Introduction to Solid Modeling. – W. H. Freeman & Co.

Nichol, J. & Wong, M.S., 2005: Modeling urban environmental quality in a tropical city. – Landscape and Urban Planning **73:** 49–58.

Quan, L. & Lan, L.D., 1999: Linear n-point camera pose determination. – IEEE Transactions on Pattern Analysis and Machine Intelligence **21**(8): 774–780.

Quattrochi, D.A. & Luvall, J.C., 1999: Thermal infrared remote sensing for analysis of landscape ecological processes: methods and application. – Landscape Ecology **14:** 577–598.

Siegert, F., Zhukov, B., Oertel, D., Limin, S., Page, S.E. & Rielay, O., 2004: Peat fires detected by the BIRD satellite. – International Journal of Remote Sensing **25** (16): 3221–3230.

Sig3D – Special Interest Group 3D. – http://www.ikg.uni-bonn.de/sig3d (accessed 15.05. 2007).

Stilla, U. & Michaelsen, E., 2002: Estimating vehicle activity using thermal image sequences and maps. – International Archives of Photogrammetry and Remote Sensing **34** (4).

Triggs, B., 1999: Camera pose and calibration from 4 or 5 known 3D points. – International Conference on Computer Vision.

VRML 97 – ISO /IEC 14772-1:1997. – http://www.web3d.org/x3d/specifications/vrml/ISO-IEC-14772-VRML97/part1/concepts.html (accessed 05.05.2007).

Addresses of the Authors:

Dipl.-Ing. Ludwig Hoegner, Prof. Dr.-Ing. Uwe Stilla, Technische Universität München, Fachgebiet Photogrammetrie und Fernerkundung, Arcisstr. 21, D-80290 München, Germany, Tel.: +49-89-289-22680, -22670, Fax: +49-89-2809573, e-mail: Ludwig.Hoegner@bv.tu-muenchen.de, stilla@tum.de

Dipl.-Ing. Holger Kimke, Prof. Dr.-Ing. Liqiu Meng, Technische Universität München, Lehrstuhl für Kartographie, Arcisstr. 21, D-80290 München, Germany, Tel.: +49-89-289-22837, -22825, Fax: +49-89-2809573, e-mail: Holger.Kumke@bv.tum.de, Liqiu.Meng@bv.tum.de

Manuskript eingereicht: Mai 2007
Angenommen: Juni 2007

Berichte von Veranstaltungen

"Optical 3-D Measurement Techniques" und "3D Virtual Reconstruction and Visualization of Complex Architectures"

vom 9. bis 13. Juli 2007 in Zürich

Die achte Konferenz für *„Optical 3-D Measurement Techniques"* (Optical 3-D) wurde vom 9. bis 12. Juli 2007 an der Eidgenössischen Technischen Hochschule (ETH) Zürich ausgetragen. Zu der Veranstaltung hatten Prof. Dr. A. Grün vom Institut für Photogrammetrie und Geodäsie der ETH und Prof. Dr. H. Kahmen vom Institut für Geodäsie und Geophysik der Technischen Universität Wien eingeladen. Zudem wurde die Konferenz von Fachorganisationen der ISPRS, IAG und FIG unterstützt. Die achte dieser alle zwei Jahre jeweils im Wechsel in Zürich und Wien ausgetragenen Konferenz hatte gegenüber ihren Vorgängerveranstaltungen ein deutlich ausgeweitetes Themenspektrum, welches von hoch aufgelösten Satellitendaten und digitalen Luftbildkameras bis zu industriellen Streifenprojektionssystemen und forensischen Anwendungen reichte. Dieses breite Themenspektrum begründen die Veranstalter durch die zu beobachtende Konvergenz der Technologien, die zunehmend vom Maßstab unabhängig zum Einsatz kommen.

Die Konferenz gliederte sich in insgesamt 18 Sitzungen, von denen 14 jeweils parallel in zwei so genannten technischen Sitzungen abgehalten wurden. Vier Sitzungen wurden als Plenumssitzung veranstaltet zu den Themenschwerpunkten *Range Imaging, Digital Airborne Large-format Cameras, Terrestrial Laser Scanning* und *Body Modelling and Crime scene investigations*. Hinzu kamen noch zwei Postersitzungen. Erstmalig wurde eine eigene Sitzung den Entwicklungen der digitalen Photogrammetrie in Russland gewidmet. Der Veranstaltung wohnten 173 Teilnehmer aus 35 Nationen bei. Der Konferenzband umfasst 101 Beiträge. Den sozialen Höhepunkt der Veranstaltung bildete eine Exkursion zu den Rheinfällen mit Konferenzdinner auf Schloss Laufen.

Die Konferenz wurde eröffnet mit einem eingeladenen Vortrag von Prof. Roland Siegwart zum Thema „UAVs – Status und Erwartungen" (UAV – Unmanned Aerial Vehicle), der über unbemannte Luftfahrzeuge referierte und insbesondere über die beeindruckende Miniaturisierung bei den so genannten „Micro-UAVs" berichtete. Im Laufe der Konferenz konnten sich die Teilnehmer auch in einer praktischen Demonstration von der Tauglichkeit eines UAVs für die Vermessung überzeugen. Die Veranstaltung widmete sich in gleich zwei Sitzungen dem Thema *Range Imaging*, also der direkten dreidimensionalen Bildgewinnung mittels flächenhafter Laufzeitmessung. Den Veranstaltern war es gelungen, mehrere der namhaften Hersteller einzuladen, um den Teilnehmern einen umfassenden Überblick zu geben. Die Vortragenden aus Wissenschaft und Industrie informierten dabei über den Stand der Entwicklung, Einsatzbereiche und Potenziale der Technologie. Anwendungen in der Messtechnik stehen noch im Hintergrund, stattdessen dominieren Anwendungen in der automobilen Sicherheitstechnik und Multimedia-Anwendungen wie beispielsweise eine neuartige Benutzerschnittstelle für Spielkonsolen.

Aus den zahlreichen interessanten Vorträgen der Konferenz blieben einige Beiträge in besonderer Erinnerung. So zum Beispiel der Beitrag von U. Tempelmann zum neuen Kameramodell der Leica ADS40, welcher über Neuerungen an der Hardware und das damit verbundene neue Kalibrierverfahren berichtete. K. Mechelke beschrieb das Genauigkeitsverhalten der neuesten Generation von terrestrischen Laserscannern basierend auf aufwändigen Testverfahren und gab damit einen herstellerübergreifenden Vergleich. B. Leibe stellte Verfahren zur bildbasierten Erstellung eines dreidimensionalen Stadtmodells aus fahrzeuggetragenen Kameraaufnahmen vor, das zur Unterstüt-

zung in Fahrzeugnavigationssystemen dienen soll.

Anwendungen und Technologien aus dem Bereich der Denkmalpflege und der Dokumentation von Kulturgütern wurden auf der *„Optical 3-D"* nicht präsentiert. Diese Themen wurden in die Konferenz *„3D Virtual Reconstruction and Visualization of Complex Architectures"* (3D-Arch) verlagert, die vom 12. bis 13. Juli am selben Veranstaltungsort abgehalten wurde. Die enge Verzahnung beider Veranstaltungen wurde durch eine gemeinsame Sitzung betont. Die *3D-Arch* wurde gemeinsam organisiert von den Verantwortlichen der ISPRS Arbeitsgruppe V/4: Sabry El-Hakim (NRC, Ottawa, Kanada), Fabio Remondino (ETH Zürich) und Jan Böhm (Universität Stuttgart) und der Arbeitsgruppe V/2: Pierre Grussenmeyer (INSA, Straßburg, Frankreich) und Klaus Hanke (Universität Innsbruck) sowie von Lorenzo Gonzo (ITC-IRST, Trento, Italien).

Der Workshop gab einen guten Eindruck über den momentanen Stand der Forschung auf dem Gebiet der Modellierung, sowohl mit rein bildbasierten Ansätzen als auch mittels Laserscanning für Konstruktionsanwendungen in der Architektur und im Denkmalschutz. Mit 68 Teilnehmern konnte sich die Veranstaltung, die erst zum zweiten Mal nach 2005 abgehalten wurde, bereits erfolgreich etablieren. Der Zusammenschluss der Workshops der ISPRS Arbeitsgruppen V/4 – *Virtual Reality and Computer Animation* und V/2 – *Cultural Heritage Documentation* war gelungen, da sich Anwendungen und Technologien gut ergänzten. In sieben Sitzungen und einer Poster-Sitzung wurden insgesamt 30 Beiträge vorgestellt, die auch online unter http://www.commission5.isprs.org/3darch07 verfügbar sind.

L. Borgeat eröffnete die Veranstaltung mit seinem eingeladenen Vortrag zur Visualisierung großer Datensätze. In einer Softwaredemonstration zeigte er die Leistungsfähigkeit eines modernen Visualisierungssystems auf, das selbst bei immensen Datenmengen mittels automatischer Darstellung im geeigneten Level-of-Detail (LOD) noch flüssige Interaktion ermöglicht. M. Pollefeys präsentierte die Arbeiten seiner Gruppe zur Echtzeitrekonstruktion von Stadtmodellen aus Videodaten. Der Vortrag brachte den Zuhörern den aktuellen Forschungsstand auf dem Gebiet Computer Vision nahe, wo Echtzeitfähigkeit und Automatisierung im Vordergrund stehen. Demgegenüber stellte S. Linsinger die Ergebnisse seines Projekts zur Rekonstruktion des Cuvilliés-Theaters in München vor. Die beeindruckende Punktdichte und Detailliertheit des Modells beruhen auf 10-monatiger Datenaufnahme mittels Laserscanning und Streifenprojektion sowie intensiver manueller Nachbearbeitung. Dieser Bericht aus der praktischen Anwendung stellte den Gegenpart zu den vollautomatischen Ansätzen von Computer Vision dar und belegte die Ausgewogenheit der Veranstaltung zwischen Forschung und praktischer Anwendung.

Jan Böhm, Stuttgart

Hochschulnachrichten

Universität Zürich

Dipl.-Geogr. Benjamin Kötz promovierte am 02.08.2006 am Geographischen Institut (Fernerkundung/RSL) der Universität Zürich mit der Arbeit „*Estimating Biophysical and Biochemical Properties over Heterogeneous Vegetation Canopies – Radiative Transfer Modeling in Forest Canopies Based on Imaging Spectrometry and LIDAR*" zum Dr. sc. nat. 1. Gutachter: Prof. Dr. Klaus Itten, Zürich, 2. Gutachterin: Dr. Britta Allgöwer, Zürich, 3. Gutachter: Dr. Frédéric Baret, INRA, Avignon (F)

Kurzfassung:
Vegetation steuert wichtige Ecosystem-Prozesse, die für den Austausch von Energie und Masse innerhalb der terrestrischen Biosphäre verantwortlich sind. Eine umfassende Beschreibung der Vegetationsbestände ist folglich notwendig, um die heterogene und dynamische terrestrische Biosphäre systematisch zu beobachten.

Obwohl die Erdbeobachtung durch Fernerkundung detaillierte Messungen von der Erdoberfläche liefert, sind verlässliche Datensätze über den Vegetationszustand der Erde in seiner räumlichen Verteilung und zeitlichen Veränderung nach wie vor eine Herausforderung. Im Vergleich zu der heterogenen und drei-dimensionalen Beschaffenheit der Vegetation sind die generell verfügbaren unabhängigen Messungen der Erdbeobachtung limitiert. Die Ableitung der Vegetationseigenschaften von Erdbeobachtungen wird so zu einem unterbestimmten System. Für eine verbesserte Ableitung von Vegetationseigenschaften muss daher die Anzahl von unabhängigen Informationsquellen erhöht werden.

Die detaillierten Messungen der beiden Erdbeobachtungssysteme Bildspektrometrie und LIDAR liefern solche unabhängigen Informationen. Die Informationsdimensionen, aufgenommen durch die zwei Sensorsysteme, enthalten relevante Daten für unterschiedliche Aspekte einer umfassenden Vegetationsbeschreibung. Die Informationsdimension eines LIDAR (LIght Detection And Ranging) liefert direkte Messungen der vertikalen Bestandesstruktur einschließlich der Bestandeshöhe. Die spektrale Informationsdimension der Bildspektrometrie enthält dagegen Informationen über die biophysikalischen und biochemischen Bestandeseigenschaften.

Die vorliegende Dissertation konzentriert sich auf die kombinierte Auswertung der unabhängigen Dimensionen der Erdbeobachtung durch Bildspektrometrie und LIDAR basierend auf Strahlungstransfermodellierung. Der Ansatz der Strahlungstransfermodellierung beschreibt explizit den Zusammenhang zwischen dem Fernerkundungssignal und Vegetationseigenschaften. Ein Strahlungstransfermodell (RTM – Radiative Transfer Model) berücksichtigt die einfallende Strahlung, Sensorspezifikationen und die physikalischen Prozesse, welche den Strahlungstransfer innerhalb eines Bestandes beherrschen. Zwei solcher physikalisch basierten RTM werden eingesetzt, um die Signale eines Bildspektrometers und eines large footprint LIDAR (LIDAR mit weiter Beleuchtungszone) separat zu beschreiben. Die kombinierte Invertierung dieser zwei RTM führte zu einer synergetischen Auswertung der unabhängigen Informationsdimensionen der beiden Erdbeobachtungssysteme.

Die entwickelte Methode bietet nicht nur einen Algorithmus von höherer Verlässlichkeit, sondern liefert auch eine erweiterte Beschreibung des Vegetationsbestandes. Es wurden zuverlässige und quantitative Schätzungen von Bestandeseigenschaften erzielt, die sowohl die horizontale und vertikale Bestandesstruktur als auch die Biochemie der Belaubung eines Nadelwaldes umfassten. Die Resultate des vorgeschlagenen Ansatzes zeigen das Potenzial von unabhängigen Informationsdimensionen der Erdbeobachtung und von Strahlungstransfermodel-In-

vertierung für die Ableitung von quantitativen Vegetationseigenschaften.

Technische Universität Wien

Dipl.-Ing. Zvonko Biljecki promovierte am 27.06.2007 am Institut für Photogrammetrie und Fernerkundung der Technische Universität Wien mit der Arbeit „*Konzept und Implementierung eines Kroatischen Topographischen Informationssystems*“ zum Dr. techn.
1. Gutachter: Prof. Dr. Norbert Pfeifer,
2. Gutachter: Prof. Dr. Gottfried Konecny.

Kurzfassung:
Ziel dieser Doktorarbeit sind das Konzept und die Implementierung eines topographischen Informationssystems (TIS) aus Sicht des aktuellen Standes der Photogrammetrie. Ein TIS kann als computerbasiertes System mit Datenverbindung zur Landschaft definiert werden (K. Kraus, Photogrammetrie, Band 3, Wien, 2000), welches erfasst und weiterverarbeitet, speichert und reorganisiert, modelliert und analysiert, sowie alphanumerisch und grafisch darstellt. Ein TIS setzt sich aus Datenbank und Anwendungssoftware zusammen.

In der Praxis wird mit vorhandenen Plänen aus Papier und unvollständigen digitalen Datensätzen mit heterogener Qualität gearbeitet. Dagegen wird hier eine neue Herangehensweise vorgeschlagen, die einen optimalen Aufbau eines landesweiten digitalen topographischen Informationssystems darstellt. Diese wurde im Zusammenhang mit dem Aufbau eines Informationssystems der Republik Kroatien untersucht.

Die Anwendung des konzeptionellen Modellierens, der Normierung und der formalen Spezifikation von geometrischen und topologischen Darstellungen, beruhend auf dem Konzept erfolgreicher Entwicklungen von TIS unter Berücksichtigung von Benutzerbedürfnissen, sind die grundlegenden angewandten Methoden zur Lösung der gegebenen Aufgabe. Diese beinhaltet das Erstellen des Konzepts eines TIS sowie dessen Einführung, welche neben ihrem wissenschaftlichen Bestandteil ausdrücklich durch organisatorische und praktische Aspekte gekennzeichnet ist. Die wissenschaftliche Komponente beinhaltet GML (Geography Markup Language) und objektorientierte Herangehensweise beim Modellieren der Daten und der funktional orientierten Modellierung der realen Welt. Der Autor wird unterstützt von einer Vielzahl von Benutzern aus der höchsten staatlichen geodätischen Behörde, die zuständig für die Verwaltung und Fortführung geodätischer und raumbezogender Systeme im Sinne der Gesetzgebung zur hoheitlichen Vermessung und Verwaltung des Liegenschaftswesens ist.

Im Weiteren werden die bedeutendsten verwirklichten Schritte zur Implementierung dieses TIS dargestellt, welche die umfangreichsten staatlichen Aktivitäten innerhalb der nationalen räumlichen Dateninfrastruktur der letzten zehn Jahren waren. Die vertragliche Vereinbarung umfasst Maßnahmen zum „Erstellen eines topografischen Informationssystems der Republik Kroatien – CROTIS“, Aufstellen eines Teams zur Implementierung dieses TIS mit entsprechendem Expertenwissen, Erstellen eines Datenmodells in Zusammenarbeit mit den Endanwendern, Erfassung der Daten bezüglich des erstellten konzeptionellen Modells und Etablierung einer topografischen Datenbank in Begleitung von hunderten bereits eingeführten Projekten mit Blick auf die Aktivitäten bezüglich des Erstellens des CROTIS / STOKIS.

Das konzeptionelle Modellieren basiert auf Abstraktion, d. h. Schwerpunktbildung bzw. Erkennen von Ähnlichkeiten zwischen Objekten der realen Welt unter vorläufiger Vernachlässigung ihrer Unterschiede. Abstraktion wird verwendet, um das Abbild der realen Welt, der Objekte und ihrer Verbindungen in eine Hierarchie von Abstraktionen, d. h. in die Kombination von Aggregationen und Generalisierungen, zu zerlegen. Wie andere Geoinformationssysteme Europäischer Staaten berücksichtigt auch das Kroatische TIS ISO-Standards. Ziel ist es, Standards und Grundregeln für das Modellieren eines grafischen und alphanumerischen Kennzeichensystems herzustellen, welches für das Definieren, Strukturie-

ren, Ersetzen, Kodieren, Umwandeln und Übertragen räumlicher Daten bestimmt ist. Geometrische und alphanumerische Daten bilden ein komplexes System, das in einer klaren Art die Geometrie und die Topologie der geografischen Objekte darstellt. Die Komplexität wird bestimmt durch die Anwendungen der Benutzer. Die geoinformatisch funktional orientierte Vorgehensweise in CROTIS beschleunigt den Prozess der Umwandlung der Realität in die gewünschte Informationsform und liefert eine gesamte Abbildung des erstellten Datenmodells in Form einer strukturierten Basis. Die Einrichtung eines funktionalen, nicht-analogen kartografischen Betriebes ist die grundlegende Eigenschaft dieses neuen Konzeptes.

Analoge kartografische Vorgehensweisen können in der Darstellung der Realität nicht die Erwartungen moderner Geoinformationssysteme erfüllen. Luftbilder liefern entweder unzureichende oder überflüssige Daten in Anbetracht aller Grundlagen und Standards der kartografischen Darstellung. Funktional orientierte Abbildungen steigern, unter Berücksichtigung des konzeptionellen Modells, das Qualitätsniveau der Daten erheblich. Die Realität in all ihrer Komplexität kann nur unter Erkennung dieser Vielschichtigkeit in einem befriedigenden Maße dargestellt und verarbeitet werden. Der grundlegende Leitfaden zur funktional orientierten Modellierung gibt der Darstellung und Erfassung der Daten ein höheres Gewicht mit größerem Einfluss in der räumlichen Erschließung. Hauptverkehrsstraßen, sonstige Kommunikationswege, Gebäude und ihre grundlegenden Ansätze, sind wesentliche Objektklassen.

Gegenwärtig wird CROTIS abgestimmt mit den aktuellsten ISO- und OGC-Standards, um Prozeduren zur Verteilung topografischer Daten, zum Zweck der Speicherung und des Datenaustausches zu entwickeln, in welchen GML benutzt wird. Eine neue Herangehensweise des Datenmodellierens, ein neues Schema für den Datenaustausch und ein neuer Datenkatalog wurden mit dieser Untersuchung eingeführt. Das Modellieren wurde zum Zweck der Verbesserung des Datenmodells gemäß den Standards getätigt, welche Methoden und Eigenschaften zur Ansicht der Objektmodelle und Objektmerkmale liefern. UML (Unified Modelling Language) wurde zur formalen Beschreibung des Datenbestandes benutzt. Das logische Datenmodell in UML ermöglicht die Implementierung eines GIS einschließlich der Datenbank. Die automatische Generierung eines GML-Anwendungsschemas wurde aus UML heraus entwickelt. Die Datenbeschreibung ist im Datenkatalog beinhaltet, welche automatisch aus dem Datenmodell resultiert.

Das grundlegende Prinzip im Erstellen von CROTIS ist funktional orientiertes, konzeptionelles Modellieren und formales Spezifizieren geometrischen und topologischen Modellierens auf konzeptioneller Ebene, das eine leichtere und erfolgreichere Entwicklung und Realisierung komplexer Geoinformationssysteme unterstützt. Diese wissenschaftlichen Vorgehensweisen wurden jüngst innerhalb einiger Workshops getestet und überprüft, welche mit mehr als dreißig der wichtigsten Nutzer organisiert wurden. Diese Vorgehensweise, sowie die vollständige Kompatibilität mit ISO-Standards, liefert eine vollständige und rasche Integration des Kroatischen Geoinformationssystems, als eine der wichtigsten Schnittstelle des NSDI, in europäische (INSPIRE) und globale räumliche Dateninfrastrukturen. „Die geoinformatische Vorgehensweise im Erstellen der Realität mit reduzierter Quantität weniger bedeutender Daten, mit gleichzeitiger Betonung ihrer funktionalen Notwendigkeit, begleitet durch die Nutzung objektorientierter Systeme und moderner informatischer Infrastruktur“ entspricht den Merkmalen von CROTIS.

Mitteilungen der DGPF

Spende zur Nachwuchsförderung in der DGPF

Im Juli 2007 hat Herr Dr. Andreas Busch, Bundesamt für Kartographie und Geodäsie, der DGPF eine Spende von 1875 Euro überlassen, die der Nachwuchsförderung zugute kommen soll. Dr. Busch erhielt im Juni 2007 den diesjährigen Carl-Pulfrich-Preis und spendet unserer Gesellschaft sein damit verbundenes Preisgeld. Die DGPF bedankt sich herzlich für diesen Beitrag und wird ihn zweckgebunden verwenden.

Oldenburg, August 2007

Thomas Luhmann, Präsident

Buchbesprechungen

Adrijana Car, Gerald Griesebner, Josef Strobl (Hrsg.), 2007: Geospatial Crossroads @ GI_Forum. Proceedings of the First Geoinformatics Forum Salzburg, Herbert Wichmann Verlag.

Seit 1994 hat sich die AGIT in Salzburg zweifelsohne als die größte und wichtigste Tagung im Bereich der angewandten Geoinformatik im deutschsprachigen Raum etabliert. Dies wird nicht zuletzt durch die hohe Anzahl von rund 1100 Teilnehmern im Jahr 2007 wieder einmal eindrucksvoll belegt. Inzwischen hat sich die Bedeutung der AGIT auch über Österreich, Deutschland und die Schweiz hinaus herumgesprochen. Die Nachfrage aus dem Ausland ist so groß geworden, dass sich die Veranstalter entschlossen haben, 2007 parallel zur klassischen AGIT erstmals ein thematisch ähnlich konzipiertes Symposium – das „Geoinformatics Forum" (kurz: „GI_Forum") – in englischer Sprache anzubieten.

Aus den rund 40 Tagungsbeiträgen wurden – teilweise mit full review – 22 Aufsätze ausgewählt und im vorliegenden, englischsprachigen Band zusammengestellt. Die Internationalität wird durch Autoren aus 14 Nationen bzw. drei Kontinenten abgebildet, wobei allerdings deutsch-sprachige Verfasser für knapp die Hälfte der Beiträge verantwortlich oder daran beteiligt sind.

Thematisch gesehen ist dieses Buch, genauso wie die bekannten AGIT-Bände, sehr breit gefächert. So gibt es Beiträge zu Datenerfassungsverfahren, GI-Technologien, raumbezogenen Analysen, Geostatistik, dynamischen Modellierungen und Simulationen, Standardisierungen, Geodateninfrastrukturen oder Ontologien. Aufgrund der heterogenen Autorenschaft, die hauptsächlich aus Universitäten, aber auch aus Wirtschaft und Verwaltung stammt, ergibt sich ein Gemisch aus Grundlagenarbeiten und anwendungsorientierten Beiträgen. Nicht zuletzt aus Platzgründen sollen und können die einzelnen Aufsätze hier nicht bewertet werden.

Mit dieser breit gefächerten Zusammenstellung wird der vorliegende Band dem Schlüsselwort „Crossroads" (Kreuzungen) im Titel gerecht, auch wenn der Leser in den einzelnen Beiträgen die Kreuzungspunkte selbst identifizieren und miteinander zu einem Netzwerk verknüpfen muss. Leider wird dieser Prozess durch die alphabetische Sortierung der Aufsätze nach Autorennamen (anstelle nach Themengruppen) nicht besonders gut unterstützt. Andererseits ergibt sich auch das typische Problem von Tagungsbänden, dass für einen Leser meist nur einzelne Beiträge von Interesse sind bzw. ein Überblick über andere Themenbereiche aufgrund der Spezialisierung der Auf-

sätze oft nicht möglich ist. Es ist also „für jeden etwas dabei", eine konkrete Zielgruppe für das Buch ist allerdings schwer zu definieren.

Abschließend stellt sich die Frage nach der Zukunft des englischsprachigen „GI_Forums": Die weitere qualitative und quantitative Entwicklung wird sicherlich darüber entscheiden, ob der bilinguale Ansatz, der 2007 durch zwei getrennte Bücher abgebildet wurde, aufrecht erhalten wird, oder ob alle Beiträge zu einem großen Band verschmelzen werden. Beide Ansätze haben ihre Vor- und Nachteile. Unabhängig hiervon ist aus Sicht des Rezensenten zu wünschen, dass sich der internationale Anteil an der Gesamtveranstaltung noch verstärken wird, was den Besuch in Salzburg nicht nur für deutschsprachige Teilnehmer bestimmt noch attraktiver machen wird.

Jochen Schiewe, Osnabrück

Thomas Jekel, Alfons Koller, Josef Strobl (Hrsg.), 2007: Lernen mit Geoinformation II, Herbert Wichmann Verlag.

Das Buch *Lernen mit Geoinformation II* ist ein Tagungsband, der anlässlich der AGIT 2007 entstanden ist. Inhaltlicher Schwerpunkt des Buches, das aus 22 Fachbeiträgen und einem sehr interessanten Editorial besteht, ist der Einsatz von GIS an Schulen im deutschsprachigen Raum. Das Buch ist in vier Kapitel untergliedert. Das erste Kapitel führt zuerst in Strategien und Konzepte ein. Ausgehend von einer europäischen Perspektive bekommt der Leser anhand mehrerer Beiträge viele Hinweise, wie GIS in Schulen und anderen Bildungseinrichtungen eingeführt werden kann. Dabei gibt es selbstverständlich nicht d i e Lösung, vielmehr gilt es zu unterscheiden, für welche Altersklasse, in welcher Schulart und in welchem Fach und mit welchem Ziel GIS in der Schule eingesetzt werden soll. Gerade die Frage „Wieso eigentlich GIS im Unterricht?" – zur Unterstützung der Raumkompetenz oder zum Erlernen moderner Methoden und Werkzeugen raumbezogener digitaler Datenverarbeitung, wird intensiv beleuchtet.

Aus den Beiträgen wird auch deutlich, dass GIS in der einen oder anderen Form mittlerweile in jeder Schule eingesetzt werden kann. Einen besonderen Beitrag dazu liefern Google Earth und artverwandte digitale Globen und Fachinformationssysteme. Diese bilden das zweite Kapitel. Die Fachbeiträge zeigen zum einen die Fälle verfügbarer Informationen auf und verdeutlichen zum anderen anhand experimenteller Untersuchungen, wie gut sich Schüler mit digitalen Globen zurechtfinden können und wie sich diese in den Unterricht integrieren lassen. In Kapitel drei werden vor allem praktische Beispiele für verschiedene Altersstufen und Fächer vorgeführt. Der theoretische Anteil in den Beiträgen ist sehr unterschiedlich, aber die Mischung gibt vor allem für Lehrer, die über den Einsatz von GIS im Unterricht nachdenken, eine Menge Anregungen.

Das letzte Kapitel geht auf E-Learning im Bereich GIS in der universitären Lehre ein, sei es im Fern- oder im Direktstudium. Neben der Untersuchung, wie die Selbstorganisation im Fernstudium bei GIS-Projekten funktioniert, wird auch derjenige auf seine Kosten kommen, der ein didaktisches Konzept für eine GIS-Lehrveranstaltung benötigt. Allerdings musste an der ETH Zürich offensichtlich wieder ein neuer Begriff für GIS erfunden werden (Geographische Informationswissenschaft und -technologien). Nun denn. Der sicherlich interessanteste kurze Beitrag aus dem letzten Block berichtet über das Projekt eduGI, das den europäischen Austausch von E-Learning Kursen aus dem GIS-Bereich vorantreiben soll.

Als Fazit des Buches kann man festhalten, dass nicht jeder Beitrag ein Highlight ist. Es enthält aber viele gute Hinweise für (Hochschul-) Lehrer, die GIS an Schulen einführen wollen und einen Aufruf, dass sich die Universitäten auch bei der Verbreitung von GIS an Schulen als regionale Kristallisationskerne engagieren müssen.

Görres Grenzdörfer, Rostock

Veranstaltungskalender

2007

5.–7. November: **Joint International Symposium and Exhibition on Geoinformation and International Symposium on GPS/GNSS 2007** in **Johor Bahru**, Malaysia. Auskünfte: Md Nor Kamarudin, e-mail: md.nor@fksg.utm.my und isg.gnss07@fksg.utm.my, http://www.isg-gnss07.com

12.–16. November: 28th **Asian Conference on Remote Sensing ACRS2007 in Kuala Lumpur**, Malaysia. Auskünfte: Conference Secretariat, e-mail: acrs2007@macres.gov.my, http://www.macres.gov.my/acrs2007

14. November: 18th **IEEE/ISPRS/OGC GEOSS workshop "Environmental Disasters & Tsunami"** in **Kuala Lumpur**, Malaysia. Auskünfte: Prof. Ian Dowman, e-mail idowman@ge.ucl.ac.uk, http://www.macres.gov.my@acrs2007.

15.–17. November: **ISPRS** WG VII/7 **Workshop** in **Riaydh**, Saudi-Arabien. Auskünfte: Dr. Sultan Al Sultan, Tel.: +966-60-0050, e-mail: rsensing_2004@yahoo.com

20.–23. November: **ISPRS** WG I/6 **Workshop on Earth Observation Small Satellites for Remote Sensing Applications** in **Kuala Lumpur**, Malaysia. Auskünfte: Mazlan Hashim, Tel.: +60-7-553-0873, Fax: +60-7-55-6163, e-mail: mazlan@fksg.utm.my, http://www.commission1.isprs.org/wg6

26.–30. November: **VAST2007 – Taking up the Challenge – Future technologies to empower heritage professionals** in Brighton, UK, e-mail: VAST2007@brighton.ac.uk, http://aranea.brighton.ac.uk/vast2007/

29./30. November: 2nd **International Conference on Geospatial Semantics "GeoS 2007"** in **Mexico-City**. Auskünfte: Marco Moreno-Ibarra, e-mail: marcomoreno@cic.ipn.mx, http://www.geosco.org/

4.–6. Dezember: 3rd **International Conference "Earth from Space – the Most Efficient Solutions"** in **Moskau**. e-mail: conference@scanex.ru, http://www.transparentworld.ru/conference/

12.–14. Dezember: **ISPRS** WG IV/8 **International Workshop on 3D Geo-Information**: *Requirements, Acquisition, Modeling, Analysis, Visualization* **"3D GeoInfo07"** in **Delft**, Niederlande. Auskünfte: Sisi Zlatanova, Tel.: +31-15-278-2714, Fax: +31-15-278-2745, e-mail: S.Zlatanova@tudelft.nl, 3Dgeoinfo07@tudelft.nl, http://www.3d-geoinfo-07.nl/

17.–19. Dezember: **Trans Asiatic GIS Society Conference on GIS "b-GIS@Asia"** in **Trivandrum**, Indien. Auskünfte: Trans Asiatic GIS Society, e-mail: bgisasia.info@tagsasia.org, http://www.tagsasia.org/bgisasia/

2008

22.–25. Januar: 3rd **International Conference on Computer Vision Theory and Applications** in **Funchal**, Madeira, Portugal. e-mail: secretariat@visapp.org http://www.visapp.org/

27.–31. Januar: **Electronic Imaging 2008, San Jose,** California, USA. http://electronicimaging.org/

30.–31. Januar: **Oldenburger 3D-Tage – Optische 3D-Messtechnik – Photogrammetrie – Laserscanning.** e-mail: iapg@fh-oldenburg.de. http://www.fh-oow.de/institute/iapg/workshop/

30. Januar – 01. Februar: WG I/3, I/4 & ICWG V/I **International Calibration & Orientation Workshop "EuroCOW 2008"** in **Barcelona,** Spanien. Auskünfte: Mrs. Eva Hernandez, e-mail: info@eurocow.org, http://www.eurocow.org

20.–22. Februar: **IRSPS** WG I/3, II/3, IV/3 & **EuroSDR Workshop on "Geosensor Net-**

works" in **Hannover.** Auskünfte: Prof. Monika Sester, e-mail: Monika.Sester@ikg.uni-hannover.de

25.–29. Februar: 10th **International Conference for Spatial Data Infrastructure (GSDI-10)** in **St. Augustine,** Trinidad. e-mail: onsrud@gsdi.org, http://www.gsdi.org/gsdi10/

05.–07. März: **EARSeL SIG** Joint Workshop **"Remote Sensing – New Challenges of High Resolution"** in **Bochum**. Auskünfte: Gesine Böttcher, EARSeL Sekretariat e-mail: secretariat@earsel.org, http://www.sig-urs-2008.de.

16.–18. April: 14th **IAPR International Conference on Discrete Geometry for Computer Imagery** in **Lyon,** Frankreich. e-mail: dgci2008@liris.cnrs.fr, http://liris.cnrs.fr/dgci2008/

23.–26. April: **Gemeinsame Jahrestagung 2008 von DGPF und DGfK** in Oldenburg.
– 28. Wissenschaftlich-Technische Jahrestagung der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.; DGPF (www.dgpf.de).
– 56. **Deutscher Kartographentag**; Deutsche Gesellschaft für Kartographie e.V., DGfK (www.kartographentag.net).

12.–14. Mai: 6th **International Conference on Computer Vision Systems** in **Santorin**, Griechenland. e-mail: chair@icvs2008.info, http://icvs2008.info/

14.–19. Juni: **FIG XXXI General Assembly & Working Week** in **Stockholm**. Auskünfte durch: FIG Office, e-mail: fig@fig.net, www.fig.net/events/2008/fig_2008_stockholm.pdf

24.–26. Juni: **IEEE Conference on Computer Vision and Pattern Recognition** in **Anchorage,** Alaska, USA. http://vision.eecs.ucf.edu/

3.–11. Juli: **XXI ISPRS Kongress** in **Beijing**, China. Auskünfte: Prof. Chen Jun (Congress Director), e-mail: congressdirector@isprs2008-beijing.org oder loc@isprs 2008-beijing.org, http://www.isprs2008-beijing.org/

13.–20. Juli: 37th Scientific Assembly of the **Committee on Space Research & Associated Events – COSPAR 2008**, "50th Anniversary Assembly" in **Montreal**, Kanada. Auskünfte: COSPAR Secretariat, Tel.: +33-1-44-767510, e-mail: co-spar@cosparhq.cnes.fr

4.–9. August: **GEOBIA 2008 – Pixels, Objects, Intelligence: "Geographic Object Based Image Analysis for the 21st Century"** in **Calgary**, Kanada. Auskünfte: Geoffrey J. Hay, Tel.: +1-403-220-4768, e-mail: gjhay@ucalgary.ca, http://www.ucalgary.ca/GEOBIA

12.–18. Oktober: **ECCV 2008 – European Conference on Computer Vision** in **Marseille**, Frankreich. http://eccv2008.inrialpes.fr

2009

16.–19. März: **ISPRS** WG VIII/12, 6th **EARSeL SIG IS Workshop "IMAGING SPECTROSCOPY: Imaging Spectroscopy: Innovative tool for scientific & commercial environmental applications"** in **Tel-Aviv,** Israel. Auskünfte: Prof. Eyal Ben-Dor, e-mail: bendor@post.tau.ac.il, http://www.earsel6th.tau.ac.il/

Zum Titelbild

Große Chinesische Mauer

Das Titelbild von Hongshan Zhao zeigt einen Abschnitt der Großen Chinesischen Mauer „Badaling" bei Beijing.

Die Große Chinesische Mauer besteht aus einer Reihe von Stein-, Ziegel- und Bodenbefestigungen, die sich über 7000 Kilometer vom *Shanhai Pass* („der Pass, an dem die Berge das Meer treffen") im Osten bis zum *Jiayu Pass* im Westen erstrecken. Bezüglich Länge, Oberfläche und Masse ist sie die weltweit längste und größte jemals von Menschen gebaute Struktur.

Die Chinesen begannen Wallbefestigungen zur *Frühlings-* und *Herbstperiode* um das 8. Jahrhundert v. Chr. gegen Angriffen der Nomaden aus dem Norden zu bauen. Die „berühmte Mauer" wurde während der *Qin Dynastie* (221–206 v. Chr.) gebaut. Sie verband die vorhandenen Mauern entlang der Nordgrenze des ersten chinesischen Reiches zu einem einzigen System, das als Große Mauer bezeichnet wird.

Die Große Mauer wurde von Zeit zu Zeit erneuert und verfeinert. Die am besten erhaltenen Abschnitte stammen aus der *Ming Dynastie* (1368–1644). Unter der Regierung des letzten Reiches, der *Qing Dynastie* (1644–1911), gingen die chinesischen Grenzen über die Mauern hinaus. Insbesondere wurde die Mongolei an das Reich angeschlossen, so dass die Große Mauer ihre Funktion als Befestigung verlor.

Lioiu Meng, München und Jun Chen, Beijing, China, e-mail: meng@bv.tum.de, ChenJun@nsdi.gov.cn

Do you want to join the No. 1 provider of digital aerial sensors? Do you want to be part of a group with sales of over USD 1 billion? This position allows an exceptional candidate to join the unique provider of both large format digital aerial cameras and airborne LiDAR, with more sensors sold then any other supplier. The Sales Manager will be part of the Airborne Sensor sales team based in Heerbrugg, Switzerland. He/She will be responsible for all airborne sensor hardware and related software sales within the Middle East and English speaking Africa. The purpose of the position is to greatly increase profitable sales and foster relationships with governmental departments and private aerial service providers and end-users of airborne sensors. It is the intention to be the premier provider of aerial systems in this territory.

Sales Manager Airborne Sensors

Duties & Responsibilities:
- Top line sales and gross margin responsibility for agreed budget achievement
- Relationship holder with all current and prospective service providers and end users
- Implementing strategic and operation goals of Digital Imaging business unit
- Ensuring Leica Geosystems is 'front of mind' in geospatial/airborne sensor industry
- Participation in all suitable events and exhibitions
- Answering Tenders and creating of finished tender documentation
- Completing RFIs (requests for information) from customers
- Customer & Visitor Support in Heerbrugg
- Pre Sales Marketing Efforts
- Any other responsibilities that may be assigned

Qualification/Experience:
- Excellent knowledge of Geomatics, Photogrammetry, Remote sensing, Aerial Sensors, or strong knowledge and track record in selling digital solutions
- Experience in high value, long sales cycle selling
- Experience in high level governmental sales, preferably in Middle East and or Africa
- Good understanding of technical issues relating to both hardware and software
- Strong interpersonal relation building and relationship selling skills
- Good problem solving skills
- Drive and Initiative
- Flair for Sales (and Marketing)
- Adheres to Leica values
- Good Social Skills

Education:
Educated to degree standard or equivalent sales experience. A qualification in Photogrammetry would be an advantage, as would a Business or Marketing qualification.

Language:
English besides own mother language is a must. German would be an advantage.

Travel Requirements:
It is anticipated that this position will have 40 - 50% travel.

Computer Skills:
Very good working knowledge of MS Word, Excel, PowerPoint. Knowledge of IBM Lotus Notes preferable.

If you are interested in this vacancy, please contact Mr. Juergen Schluderbacher, Human Resources Manager, Tel No. +41 71 727 4730 or email Juergen.Schluderbacher@leica-geosystems.com.

Leica Geosystems AG
Human Resources
Heinrich-Wild-Strasse
CH-9435 Heerbrugg
Switzerland

www.leica-geosystems.com

- when it has to be **right**

Leica Geosystems

Neuerscheinungen

Thomas Jekel, Alfons Koller & Josef Strobl (Hrsg.), 2007: Lernen mit Geoinformation II. XI, 240 Seiten, kartoniert. ISBN 978-3-87907-460-0, Herbert Wichmann Verlag, Heidelberg.

Adrijana Car, Gerald Griesebner & Josef Strobl (Hrsg.), 2007: Geospatial Crossroads @ GI_Forum. Proceedings of the First Geoinformatics Forum Salzburg. VIII, 198 Seiten, kartoniert. ISBN 978-3-87907-461-7, Herbert Wichmann Verlag, Heidelberg.

John Fryer, Harvey Mitchell & Jim Chandler (Eds.), 2007: Applications of 3D Measurement from Images. 384 pages, liberally illustrated with a colour section and DVD hardback, ISBN 978-1870325-69-1, Whittles Publishing, Dunbeath, Caithness, Scotland.

Arthur L. Allan, 2007: Principles of Geospatial Surveying. 472 pages, softback illustrated, ISBN 978-1904445-21-00, Whittles Publishing, Dunbeath, Caithness, Scotland, www.whittlespublishing.com.

Jonathan Iliffe, & Roger Lott, 2008: Datums and Map Projections, 2nd edition. 192 pages, softback liberally illustrated with colour, ISBN 978-1904445-47-0, Whittles Publishing, Dunbeath, Caithness, Scotland, www.whittlespublishing.com.