PFG 2012 / 6, 0677–0678
Stuttgart, December 2012

3D Reconstruction of Real World Scenes with Low-Cost Hard- and Software

Stereo visualisation and 3D reconstruction of objects together with their environments is an important topic of research but also of interest for entertainment with applications in many areas such as gaming, virtual museums and architecture.

In the last few years 3D cameras and 3D visualisation has progressed enormously. For the first time, the loop from digital stereo image acquisition to visualisation of commercial, scientific and private applications is closed. Besides this, comparatively inexpensive stereo screens and 3D printers are available for everybody.

3D capturing hardware and processing software has been available for several decades. However, their acquisition costs were exorbitantly high. Hence, the use of these systems and methods was limited to larger production companies and a few academic institutions and research institutes.

This situation has changed dramatically in recent years. Today free software is available that allows 3D reconstruction based on uncalibrated consumer-grade sensor systems. An example is Bundler, a 3D software tool based on structure-from-motion (SfM) algorithms where a large number of unordered images can be processed. Previous projects have shown the capabilities of this software where it has been applied on city-scale photo collections, e.g. the Rome in a Day project.

Beside this, digital stereo cameras can now be used for 3D capturing and visualisation. In addition to passive optical systems, some inexpensive active systems, e.g. light measurement and David laser scanner, exist. The Kinect sensor which Microsoft originally provides for the Xbox 360, applies a completely different approach. Measurements can be carried out by comparative analysis with a reference pattern using a speckle pattern depth approach. KinectFusion is another research topic that is gaining interest of the scientific community as the system focuses on real-time 3D reconstruction of indoor scenes by using a moving depth camera.

A new trend is the implementation of 3D reconstruction on a variety of web-based 3D reconstruction services, e.g. 123D Catch, available also on iPhone and iPad.

All the above-mentioned trends are within the scope of the LC3D conference. The department for Geodesy and Geoinformation Techniques at the Technical University of Berlin (TU) as well as the Department of Computer Science (Computer Vision) of Humboldt University (HU) hosted the first "Low-Cost 3D: Sensors, Algorithms, Applications" workshop which was held on December 6 to 7, 2011 at the TU Berlin. The workshop featured live demonstrations of various systems in the afternoon and continued with a conference at the next day including 13 presentations.

The workshop was organized by Prof. Dr.-Ing. Frank Neitzel (TU Berlin) and Prof. Dr. rer. nat. Ralf Reulke (HU Berlin) supported by their respective staff members. Joint organisers of the workshop were the German Aerospace Center (DLR Berlin-Adlershof), the Fraunhofer Heinrich-Hertz Institute as well as the International Society for Photogrammetry and Remote Sensing (ISPRS) represented by the Image Sensor Technology working group. The Low-Cost 3D workshop has been the first one of its kind that exclusively focused on Low-Cost 3D reconstruction while covering the entire spectrum of facets. The event attracted researchers, developers and users that are interested in various fields of application and the intrinsic potential of such techniques. In spite of the short announcement 80 participants from Germany and abroad attended the workshop which reflects the current interest in this hot topic. About two thirds of the participants were affiliated to Universities or other research facilities, 28% worked in free enterprises while a minority of 5% came from ad-

DOI: 10.1127/1432-8364/2012/0147 1432-8364/12/0147 $ 0.50

ministrative or public authorities, e.g. working in archaeology and forestry management.

Live demonstrations of 13 presenters took place at TU Berlin's Geodätenstand, a geodetic observation platform on the university's roof, and proofed that 3D data acquisition can be cheap. The combination of the icebreaker party and live demos went well, accompanied by a picturesque view at dusk over Berlin. Staff of TU Berlin presented how one can reconstruct geometry by applying Bundler and PMVS2 software, Ruhr-University Bochum showed their low-cost navigation system based on a PMD camera, the TU Clausthal-Zellerfeld demonstrated impressively their implementation of a visualisation of cave systems, Fiagon GmbH, displayed a novel navigation system for dentistry while GFaI tech ltd. presented their new laser scanner FinalScan LR-50.

At the following day 13 talks were held in five session covering aspects from data acquisition, algorithms, geometric considerations and various applications. The spectrum of the presentations was large.

Frank Neitzel (TU Berlin) presented a low-cost UAV system, an octocopter equipped with Canon IXUS 100 IS, for mobile mapping purposes. Subsequently, Konrad Wenzel (University of Stuttgart) presented a multi camera system consisting of panchromatic and near infrared cameras, while a Kinect projector was used to ensure sufficient contrast on the object's surface. Yuan Xu (Humboldt-University Berlin) used a Kinect sensor for motion detection within a soccer playing robot. Jan Böhm (University College London) presented results of his accuracy analysis that he conducted on various natural user interface sensors (Kinect, PrimeSense and Asus). Fabio Remondino (FBK Trento, Italy) showed, apart from results of accuracy analysis, effects of multi view stereo software (Photosynth, Agisoft, Apero/MicMac) for automatic generation of 3D point clouds. Sebastian Vetter (fokus GmbH, Leipzig) demonstrated the versatility of their metigo 3D software on various tasks for object documentation based on image bundles. Afterwards Günter Pomaska (University of Applied Sciences Bielefeld) presented a solution which combined Microsoft's Photosynth with SketchUp 3D for modelling buildings. Nesrine Grati (i3mainz) demonstrated a strategy for improving 3D reconstruction within point clouds of urban scenes by deploying 2D imagery. An optimised work and data flow has been revealed by Mohammed Abdel-Wahab (University of Stuttgart). Wilhelm Hannemann (TU Clausthal-Zellerfeld) took the audience underground while presenting their interactive WebGL visualisation of a spacious cave system that has been reconstructed via Bundler and PMVS2. Thomas Kersten (HCU Hamburg) revealed the potential of multi view stereo software, namely Bundler / PMVS2, Photofly / 123D Catch and Microsoft Photosynth, with examples from architecture, cultural heritage and archaeology, and compared the results with point clouds from terrestrial laser scanning as ground truth. Two final talks from commercial organisations grasped the idea of low-cost from a different perspective. Dirk Kowalewski (navXperience GmbH, Berlin) raised the question "What is the real accuracy of a GNSS antenna?". Andreas Rose (fiagon GmbH, Berlin) presented an optical navigation system for dentistry.

The contributions of Abdel-Wahab et al., Hannemann et al., and Kersten & Lindstaedt are published in this issue of the PFG.

To conclude, every participant experienced the enormous potential of the image based data acquisition techniques for numerous fields of application. While the development of the techniques is at its beginning, first commercial steps are currently undertaken. We believe that these methods will gain a wide distribution. However, aspects of accuracy, embedded quality measures and reliability are yet not suitably solved for commercial applications and have still be investigated thoroughly.

We hope that again fascinating talks on the subject can be heard at the second workshop at the end of 2012 in Berlin-Adlershof.

Frank Neitzel and Ralf Reulke, Berlin

PFG 2012 / 6, 0679–0690
Stuttgart, December 2012

Article

Automated and Accurate Orientation of Large Unordered Image Datasets for Close-Range Cultural Heritage Data Recording

Mohammed Abdel-Wahab, Konrad Wenzel & Dieter Fritsch, Stuttgart

Keywords: automatic image orientation, 3D reconstruction, bundle adjustment, close range, cultural heritage

Summary: Reconstruction of image orientations and geometry from images is one of the basic tasks in photogrammetry and computer vision. A fully automated solution of this task in terrestrial applications is still pending in case of large unordered image datasets especially for close-range and/or low-cost applications. Current solutions require high computational efforts for image networks with high complexity and diversity regarding acquisition geometry. Unlike the methods suitable for landmark reconstruction from large-scale Internet image collections we focus on datasets where one cannot reduce the number of images without losing geometric information of the dataset. Within the paper, an automated pipeline for the reconstruction of reliable and precise camera orientation from unordered image datasets is presented. Results for a close-range cultural heritage application, the example of the Amsterdam project, are shown to demonstrate the performance of the presented pipeline for applications with low cost and high accuracy requirements.

Zusammenfassung: *Automatische und hochgenaue Orientierung von großen, ungeordneten Bildverbänden bei der photogrammetrischen Aufnahme von Weltkulturdenkmälern*. Die Rekonstruktion von Kameraorientierungen und Objektstrukturen aus Bildern ist eine der Hauptaufgaben der Photogrammetrie und der Computer Vision. Eine vollautomatische Lösung für terrestrische Anwendungen mit unregelmäßig angeordneten Bildverbänden, unabhängig vom Kamerasystem (professionell oder Amateuraufnahmen), steht noch aus. Gegenwärtige Lösungen erfordern einen hohen Rechenaufwand für komplexe Bildkompositionen. Im Gegensatz zu Ansätzen zur Landmarkenrekonstruktion mittels Bilddatensätzen des Internets will der vorliegende Beitrag alle verfügbaren Bilder nutzen, um wertvolle geometrische Details nicht zu verlieren. Aus diesem Grund stellen wir einen automatischen Workflow für die Rekonstruktion von reproduzierbarer und präziser Geometrie aus ungeordneten Bildkompositionen vor. Dieser wurde für eine spezielle Anwendung bei der Rekonstruktion der beiden Tympana des Königlichen Palastes in Amsterdam entwickelt und getestet. Die Ergebnisse belegen die Leistungsfähigkeit des Gesamtkonzepts wie auch der im automatischen Workflow realisierten Einzelpakete hinsichtlich niedriger Kosten und hoher Genauigkeit.

1 Introduction

In the past few years, close-range and/or low-cost photogrammetry has become a focus of research especially since cameras enable data acquisition at very low prices, but with high geometric and radiometric quality. Therefore, low-cost multi-camera systems for efficient dense point cloud recording from imagery are particularly suitable for cultural heritage applications, where the requirements regarding acquisition efficiency, flexibility, but also spatial resolution and precision are high. For high resolution data recording in cultural heritage applications, the use of a rig with multiple cameras is beneficial. With one shot mul-

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0148
1432-8364/12/0148 $ 3.00

tiple views enable the dense reconstruction of the object surface using dense image matching methods, such as semi global matching (Hirschmüller 2008). Thus, the high similarity between the imagery is exploited to generate one high resolution point cloud with a low amount of occlusions for each shot. However, for multiple shots the imagery requires an automatic registration method. Therefore, a structure-from-motion (SfM) reconstruction method was developed in the context of an example project of data recording using five industrial cameras. They are mounted on a compact square shaped rig, which enables handheld recording of complex objects. However, for large scale data recording the derivation of accurate orientation for the high number of imagery is a key problem to be solved.

The aim of this paper is to report a pipeline for fully automatic derivation of image orientations by using a divide-and-conquer strategy to speed up the SfM process from general imagery networks without initial orientation values. SfM was originally developed to estimate geometry and camera motion from multiple images of a scene. It is used for the determination of initial values for the final and global bundle adjustment step in our pipeline. Most SfM methods are starting with a small reconstruction, i.e. a pair or triplet of images, and then expanding the bundle incrementally by adding new images and 3D points as in Snavely et al. (2008). Here, each pose estimation and point triangulation is followed by an outlier rejection and a bundle adjustment. Other approaches increase the bundle hierarchically by merging smaller reconstructions (Farenzena et al. 2009). Unfortunately, both approaches require multiple intermediate bundle adjustment results and rounds of outlier removal to minimize error propagation as the reconstruction grows due to the incremental approach. This can be computationally expensive for large datasets. This issue is considered to be solved partially in Farenzena et al. (2009) by the introduction of a local bundle adjustment procedure and in Snavely et al. (2008) by optimizing the system over a graph to order the images and remove obsolete images from the dataset. However, we focus on datasets where one cannot reduce the number of images dramatically without losing a substantial part of the model.

A third solution is the so called partitioning method, which we follow in this paper, as presented in Gibson et al. (2002). Here the key idea is reducing the problem to smaller and better conditioned sub-problems. The main advantage of these methods is not only the equalized error distribution on the entire dataset but also a speed up of the processing time. Recently, Klopschitz et al. (2010) presented a robust and flexible SfM pipeline where they used the image triplets as a base to reason about feature track compatibility and image connectivity. Within the following section the example of the cultural heritage data recording project in Amsterdam is briefly introduced. A description of the reconstruction of orientation and structure is given in section 3. In section 4 the clustering process of the global graph is described, where the large dataset is split into smaller clusters. Within section 5 the incremental reconstruction being performed for each cluster is discussed. Section 6 contains the stitching process of the multiple clusters to one cluster. Furthermore, the global bundle adjustment of the resulting cluster is explained. Experimental results for the presented dataset are shown in section 7, followed by the conclusions in section 8.

2 The Amsterdam Project

In March 2011 the Research Group "Photogrammetry and Computer Vision" of the Institute for Photogrammetry (University of Stuttgart) had an industrial contract to collect photos for a very dense 3D point cloud generation of the two Tympana of the Royal Palace in Amsterdam (Fig. 1). Each tympanum covers a triangular shape area of about 25 m in width by 5 m in height containing a relief with complex surface geometries such as statues. For this purpose it was planned to use a multi-camera system incorporating a fully automated pipeline for image orientation and dense matching methods. Thus, a method capable of processing very large image datasets with high accuracy and sufficient time was required here. The first comprehensive reports about the project are presented in Fritsch et al. (2011) and WENZEL et al. (2011). The sensor, as shown in Fig. 1 right, consists of four cameras used for the dense image matching and one camera with a larger field of view for the registration of multiple shots. The four cameras for the dense image matching have a resolution of 5 Megapixels and are equipped with lenses with a focal length of 8 mm. They are arranged in a square with the size of 7.5 cm by 7.5 cm on a solid aluminum bar to provide a stable relative orientation. The fifth camera with a resolution of 2 Megapixels, equipped with a lens with 4.7 mm focal length, is installed between the lower two matching cameras. An aluminum frame is surrounding the cameras in order to protect them from damage. Several mounts for the connection to tripods and arms for a flexible use are installed at this frame. The Microsoft Kinect is attached at the top with the pattern projector at the same height like the cameras to minimize occlusions.

In order to derive a point cloud with a sampling of 1 mm on the object and sub-mm accuracy, the presented sensor was employed for the data recording on scaffolding. Within 9.5 days about 2,000 stations were acquired leading to a total amount of about 10,000 images. First, to achieve a complete coverage, the images were acquired in nadir direction in a meandering pattern within each of the three levels of scaffolding. Then convergent shots have been captured to complete surfaces which were occluded or not covered. Ground control points measured by tachymetry provided the transformation to the global coordinate system.

3 Orientation Reconstruction Pipeline Overview

Our 3D reconstruction pipeline intends to automatically and accurately process unordered sets of images to determine relative image orientations and a sparse point cloud of tie points without prior knowledge of the scene. The pipeline, as shown in Fig. 2, mainly consists of four processing steps: (1) Employ fast image indexing to avoid costly matching of all possible image pairs, which dominates computational complexity along with the multiple bundle adjustment steps. (2) Generate tie points by means of feature extraction and matching where the required automatic measurements are realized at maximum accuracy and reliability. (3) Building and optimizing a geometry graph based on the image network, whereby the dataset can be split into reliable clusters of neighbouring images that can be processed independently and in parallel within the reconstruction step. (4) Merge all clusters and then finally adjust the full model with integrating the ground control points. A detailed description of the individual processing steps is given in the following sections. In general, camera calibration parameters are not strictly necessary for Euclidean 3D modeling, since self-calibration methods exist. However, if a stable camera with a fixed focal length is used and the values for the interior orientation are determined a priori by standard calibration methods robustness and accuracy are usually greatly improved. Furthermore, also an increase in processing speed is achieved due to the lower dimensionality of the prob-

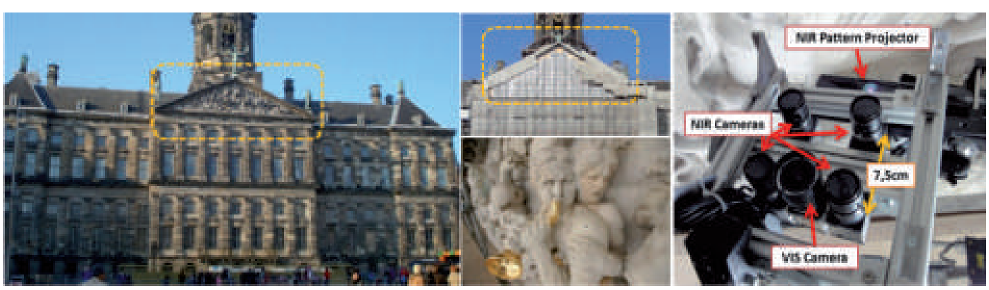


Fig. 1: East tympanum of the Royal Palace of Amsterdam; left: from distance; upper middle: with the scaffold; lower middle: a DSLR colour image of the scene shows the relief containing whole statues (size of visible control point target is 4 x 3 cm²); right: sensor design overview.

lem. Pursuant to that, we prefer to use interior calibration parameters for high accuracy applications where these values can be considered to be stable.

3.1 *Initial Network Geometry Analysis*

This step is designed to accurately and quickly index unordered collections of photos. A *connectivity* matrix is the output of this step and it reveals singleton images and small subsets that should be excluded from the dataset. Finally, it is used to guide the process of pairwise matching (section 3.2). Recent developments regarding this analysis can be distinguished into two major categories according to the type of image representation. Local feature based approaches use quality measures of matched local descriptors while global feature based approaches utilize matching histograms of full images visual words (Aly et al. 2011). In fact, both categories represent the same approach with varying degrees of approximation to improve speed and/or storage requirements. Generally, the first category provides superior recognition performance and the dimensionality is not an issue when only several thousands of images need to be processed. Consequently, we utilise a local feature based method in the pipeline presented in this paper.

For local feature-based indexing, we follow an approach adapted to the method presented in Brown & Lowe (2003) and Farenzena et al. (2009). The first step is the extraction and description of local invariant features from each image by using the SURF (Bay et al. 2006) operator on a downsampled image, e.g. using images with 2 Megapixels resolution. Then all descriptors are stored in a randomized forest of kd-trees, which represents the search dataset, to improve the effectiveness of the representation in high dimensions. Each descriptor in the query image is matched to its nearest neighbours in feature space (we used 10). For that purpose we used the fast library for approximate nearest neighbours FLANN (Muja & Lowe 2009) and the kd-tree implementation in the VLFeat library (Vedaldi & Fulkerson 2008).

Thus, each feature in a query image is initially matched to 10 features of the search dataset. Then we indicate the outliers which have a feature distance more than a certain threshold. We used twice the standard deviation of the distances as threshold value. That gives us a statistical information about the number of matched features between the query image and the remaining images which we can store in a 2D histogram. For more efficiency we used a weighted 2D histogram where the inverse of the distances between each matched feature pair are used as weights. Furthermore, we introduce additional quality measures for possible connections between the query image and the remaining images such as the approximate area of overlap derived from the convex hull of the matched feature points. Finally, the quality measures and the 2D histograms are normalized and summarized to one single quality 2D histogram, which is stored in the *index* matrix (Fig. 3a). Then this *index* matrix is binarized to the *connectivity* matrix, using three thresholds determined empirically, to determine initial probable connections as shown in Fig. 3b. Here any image pair with a quality value more/less than first/second threshold is indicated as connected/disconnected pair. The number of connected images will be compared to the third threshold, which

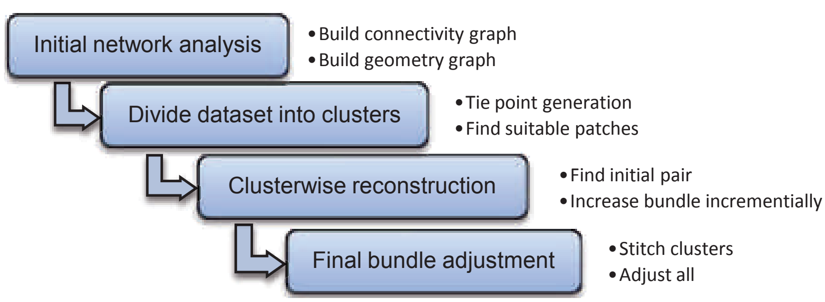


Fig. 2: Flowchart of the presented pipeline.

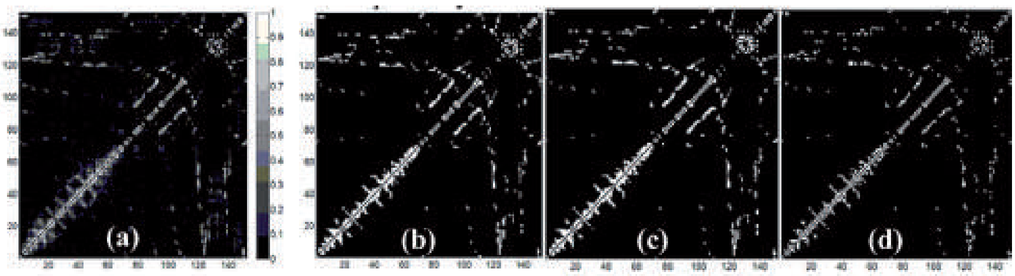


Fig. 3: Top cluster of the east tympanum dataset. The axes refer to the image identifiers in the dataset. (a) Index matrix according to the method presented in section 3.1 with 1457 edges, colour-coded between one/zero indicating connected/disconnected pair; (b) adjacency connectivity matrix before geometry verification, where the number of edges is reduced to after binarization; (c) adjacency connectivity matrix after geometric verification and only 2 edges removed (section 3.2). A white entry in a connectivity matrix indicates that the image pair is connected. (d) Connectivity matrix after the graph optimization step with 396 edges (section 5.1).

refers to the minimum number of images connected to the query image. If this threshold is not met, images with values between the first and second threshold will be added in descending order until the condition is satisfied. We used 0.7, 0.3 and 10 as the threshold values respectively.

3.2 *Pairwise Feature Matching*

Matching each connected image pair is accomplished using the *connectivity* matrix obtained during the previous section. Corresponding 2D pixel measurements are determined between all connected image pairs. Afterwards a weighted undirected graph, we call it *geometry* graph $G_E = (V, E)$ where V is a set of vertices and E is a set of edges, is constructed. Thus, two view relations are encoded such that each vertex refers to an image while each weighted edge presents the overlap between the corresponding image pairs. The weights of the edges are stored according to the number of their common points, w_{ij}^p, and the overlap area, w_{ij}^a, between view i & j. For the computation we follow the approach of Farenzena et al. (2009) where a set of candidate features is matched using a kd-tree procedure based on the approximate nearest neighbour algorithm. This step is followed by a refinement of correspondences using an outlier rejection procedure based on the noise statistics of correct/incorrect matches. The results are then filtered by a standard *RANSAC* based geometric verification step, which robustly computes pairwise relations. Homography and fundamental or essential, in the calibrated case, matrices are used with an efficient outlier rejection rule

Tab. 1: Pseudo code for the clustering approach.

Input:	Geometry graph G_E
Output:	Collection of clusters graph
1.	Set new empty graph (cluster) $G_c = \{\}$
2.	Determine most reliable edge E_{ij} in G_E
3.	Add the vertices V_i, V_j of this edge into G_c
4.	Set $E_{ij} = 0$ in G_E
5.	$\forall V_k$ in G_E connected at least with two vertices V_n, V_m in G_c If w_{nk}^p & $w_{mk}^p \geq \max\left(\frac{1}{2} w_{ij}^p, 200\right)$ & w_{nk}^a & $w_{mk}^a \geq \frac{1}{2} w_{ij}^a$ ● Add V_k into G_c and set E_{nk} & $E_{mk} = 0$ in G_E
6.	Add edges in between inlier vertices in G_c
7.	Set all these edges = 0 in G_E
8.	Repeat steps 5, 6 and 7 until ● if $V_k = 0$ in step 5 ● or if size of G_c = predefined value
9.	Store G_c and repeat all steps until all edges in $G_E = 0$

called *X84* (Hampel et al. 1986) to increase reliability and accuracy. The final output of this step is the *geometry* matrix or graph as illustrated in Fig. 3c. For an in-depth discussion see Farenzena et al. (2009) and Snavely et al. (2008) and references therein.

4 Clustering of the Global Graph

In order to speed up the computation of the incremental reconstruction, we address a fast local optimization instead of a global optimization approach. We divide the dataset into n overlapping clusters, where each one contains a manageable size of images. Thus, a parallelizable process replaces the process of reconstruction of the whole scene at once. This is particularly important since for complex datasets the large number of iterations with the growing number of unknowns can lead to very high computation times for complex datasets. The idea is to start from the most reliable part and use three images as the basic entity to extend each cluster until a predefined size. In practice, we use the workflow as presented in Tab. 1 to identify reliable clusters with the highest mutual compatibility. The idea is to start each cluster G_c from image pair i and j with high overlap in order to ensure a reliable geometry. As shown in Tab. 1, we select an initial pair according to the most reliable edge being identified by its weights, w_{ij}^p and w_{ij}^a within the geometry graph as presented in section 3.2. The graph of this cluster is then extended by the neighbouring edges with a weight (common points) greater than the half of the weight of the initial edge of this cluster or a certain threshold (we used 200 matching points for the presented dataset). We repeat this process until a predefined cluster size is reached or until no more images have sufficient overlap with this cluster. For this we apply thresholds depending on the initial edge weights. Furthermore, only images overlapping with at least two images inside the cluster are considered. While the cluster graph is growing, each used edge is eliminated in the geometry graph. As soon as the cluster graph is finalized, the whole procedure is repeated to find the next cluster until all imagery is covered. The overlap between the final clusters is ensured by considering and removing connections (edges in the graph) only instead of images. Thus, common cameras between the clusters remain.

5 Cluster Reconstruction

Once the clusters are divided as described in the previous section, we can start the reconstruction process for each cluster as follows.

5.1 *Optimization of Cluster Graph*

For each cluster we track the keypoints only over images in this cluster (locally) and store the results in a visibility matrix, which depicts the appearance of points in the images. The results of this step will be the keypoints which have been correctly tracked in at least three images. For more efficiency, we apply a non-maximum suppression filtering approach (Fig. 4) for the tracked points to keep only the points with the highest connectivity. For each image we sort the keypoints in descending order according to their number of projections in other images. Then, the point with the greatest number of projections is visited, followed by an identification and rejection of all nearest neighbour points with a distance less than a certain threshold, e.g. 20 pixels. This step is repeated until the end of the points list. In order to maintain continuity, all points selected in an image must be considered as filtered (fixed) in the following filtering of other images. Filtering is done in order to increase the accuracy but also to reduce the number of obsolete observations. Consequently, the geometric distribution of keypoints is improved, which reduces the computational costs significantly without losing geometric stability.

Once correspondences have been tracked and filtered, we optimize the cluster graph such that we construct a weighted undirected epipolar graph for each cluster G_p containing common tracks. The weight w_{ij} of an edge represents the number of common points between the corresponding image pair. Then we build G_r, the edge dual graph of G_p, where every node in G_r corresponds to an edge in G_p. Two nodes in G_r are connected by an edge if

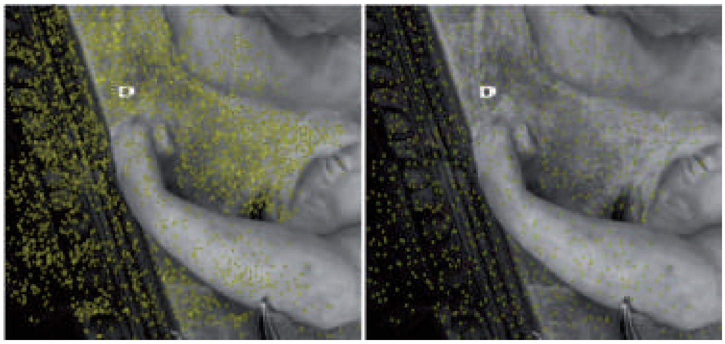


Fig. 4: Point distribution before and after filtering, 3395 & 819 points according to a filtering distance of 40 pixels.

and only if they are sharing an image and 3D points. Thus, each edge represents an image pair with sufficient overlap. Note that even if G_p is fully connected any spanning tree of G_r may be disconnected. This can happen if a particular pairwise reconstruction did not have 3D points in common with another pair. Thus, we use three images as basic geometric entity by using only points that were tracked in at least three images. These points are used to build the graph in order to guarantee full connection for any sub-sequential image. The maximum spanning tree (MST), which maximizes the total edge cost of the final graph, is then computed. The image relation retrieved as $G_p^{\max}$ graph is used for the bundle adjustment. For example, Fig. 6 presents the results of the top cluster of the east tympanum where the previous process reduced the pairwise connection from 600 edges (Fig. 3d) to 396 to orient 150 images.

5.2 *Camera and Geometry Recovery*

Each cluster is processed individually beginning with an initial reconstruction for the two, most suitable images. After this step, orientations and tie points in object space are available for these two images where one image defines the local coordinate system. Within the incremental approach images are added to the existing bundle by triangulating new points, rejecting outliers and performing another iteration of the bundle adjustment. This incremental process is repeated until all images within the cluster are processed.

Reconstruction of the initial pair

The incremental reconstruction step begins with the reconstruction of orientation and 3D points for an initial image pair. The choice of this initial pair is very important for the subsequent reconstruction of the scene. The initial pair reconstruction can only be robustly estimated if the image pair has at the same time a reasonable large baseline for high geometric stability and a high number of common feature points. Furthermore, the matching feature points should be distributed well in the images in order to reconstruct a maximum of initial 3D structure of the scene and to be able to determine a strong relative orientation between the images. Therefore, suitable image pairs should be selected according to the following conditions: the number of matching points is acceptable and the fundamental matrix must explain the matching points far better than homography models. Here we employ the geometric robust information criterion (GRIC) scores to ensure that the criteria are met as used in Farenzena et al. (2009). After that, relative orientation values for this initial pair are estimated by using Nister's implementation of the five point algorithm (Nistér 2004). A two-frame bundle adjustment starting from this initialization is performed to improve the reconstruction.

Adding new images and points

After reconstructing the initial pair additional images are added incrementally to the bundle. The most suitable image to be added is selected according to the maximum number of tracks from 3D points already being reconstructed. Within this step not only this image is added but also neighbouring images that have a sufficient number of tracks as mentioned in Snavely et al. (2007). Adding multiple images at once reduces the number of required bundle adjustments and thus improves efficiency. Next, the points observed by the new images are added into the optimization. A point is added if it is observed by at least two images, and if the triangulation gives a well-conditioned estimate of its location. This procedure follows the approach of Snavely et al. (2007).

Sparse bundle adjustment

Once the new points have been added, a bundle adjustment is performed on the entire model. This procedure of initializing a camera orientation, triangulating points, and running bundle adjustment is repeated, until no images observing a reasonable number of points remain. For the optimization we employ the sparse bundle adjustment implementation "SBA" (Lourakis & Argyros 2009). SBA is a non-linear optimization package that takes advantage of the special sparse structure of the Jacobian matrix used during the optimization step in order to provide a computation with reduced time and memory requirements.

6 Stitching of Clusters and Global Adjustment

After the reconstruction of points and orientations for the overlapping clusters the results are merged. Since outlier rejection was performed within the previous steps, the available 3D feature points are considered to be reliable and accurate. Due to the overlap, the clusters have a certain number of points and image orientations in common which enables the determination of a seven-parameter transformation in order to align the clusters into a common coordinate system. The transformed orientations and points are introduced into a common global bundle adjustment of the whole block. If ground control point measurements are available they can be used to improve the bundle stability and to enable georeferencing.

7 Experimental Results

For the derivation of the point cloud, the exterior orientations were derived first using the presented method, where each façade was automatically divided into 6 individually processed patches. Secondly, an additional dense image matching step followed using the obtained orientations. However, since the exposure of all cameras was not synchronized sufficiently, the relative orientation was not sufficiently stable. This is particularly important since the acquisition distance was short and the accuracy requirements were high. Thus, the relative orientation from the calibration was omitted and determined using the SfM process instead. In order to use the relative orientation directly, hardware triggering of the cameras should be used instead of software triggering.

Figs. 5 and 6 depict the results of the tympanum at the west façade where approximately 4000 images are oriented. The first row shows the reconstructed and stitched 6 clusters where the mean reprojection error before merging are around 1 pixel and it is reduced to 0.5 pixel after the final bundle adjustment step. The second row left shows the full sparse cloud of 1.1 million feature points in object coordinates and right the dense point cloud derived by a subsequent dense image matching step with about 1.1 billion points. As presented in section 2, the targets are distributed over the whole object and are used for georeferencing. These control points are captured in 12 independent clusters. The white circles within the targets are detected and measured automatically using an ellipse fit. These image measurements are considered to have an accuracy of about 0.1 pixels.

Up to this point the control points are not used in the bundle and thus do not impact the orientations to be evaluated. Consequently, they can be used to assess the quality of the relative orientations by evaluating the reprojection errors. These results are demonstrated in Tab. 2 rows 2 and 5, for each cluster of both tympana respectively. The root-mean-square of the reprojection errors for each dataset is about 0.3 pixels. At 70 cm distance this corresponds to an error of approximately 0.2 mm for the image scale of this dataset. This is considered to meet the requirements for the later dense surface reconstruction step, where

Tab. 2: Overview of the 6 clusters C1–C6 and their performance of the east and west tympanum. Time = runtime for each cluster, GC point / Projection = identified ground control points per overall count of projections in the image, RMS (pixel) = error of the reprojection, GC = number of ground control points, RMS (mm) = error of a ground control point in object space.

	Cluster Ids.	C1	C2	C3	C4	C5	C6	.
East tympanum	Images	909	526	440	682	412	460	
	3D points	517,360	386,708	300,799	420,020	276,534	258,184	
	Time (hrs)	3.48	1.62	0.61	3.08	0.69	0.78	Mean
	GC points/Proj.	20/163	12/63	11/52	13/80	5/17	9/55	11.67/71.67
	RMS (pix.)	0.39	0.18	0.26	0.26	0.36	0.21	0.28
	GC points	18	6	11	13	5	9	10.33
	RMS (mm)	2.63	2.16	2.18	3.25	2.97	3.21	2.73
West tympanum	Images	862	478	779	995	956	849	
	3D points	318,064	149,349	245,232	318,366	332,613	322,151	
	Time (hrs)	3.47	0.71	3.26	4.21	3.27	3.38	Mean
	GC points/Proj.	24/176	11/65	12/85	15/97	13/95	15/114	15/105.33
	RMS (pix)	0.26	0.26	0.26	0.31	0.26	0.26	0.27
	GC points	20	9	7	14	12	15	12.83
	RMS (mm)	2.24	1.59	1.36	2.06	1.13	1.90	1.71

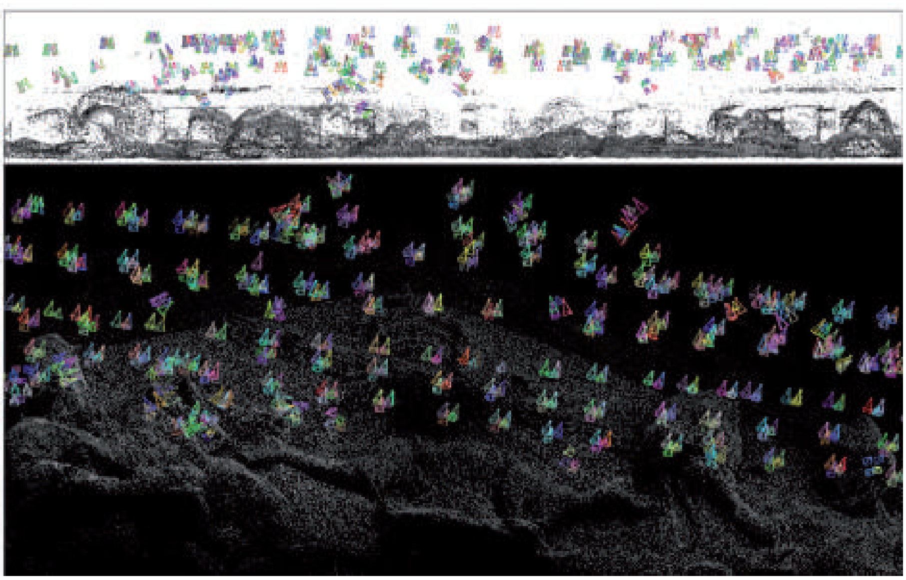


Fig. 5: Reconstructed cameras and point clouds of the tympanum at the west façade. First row: geometry of imagery with close-up area is shown in upper right corner; second row: sparse point cloud resulted from SfM (left) and derived by dense image matching (right).

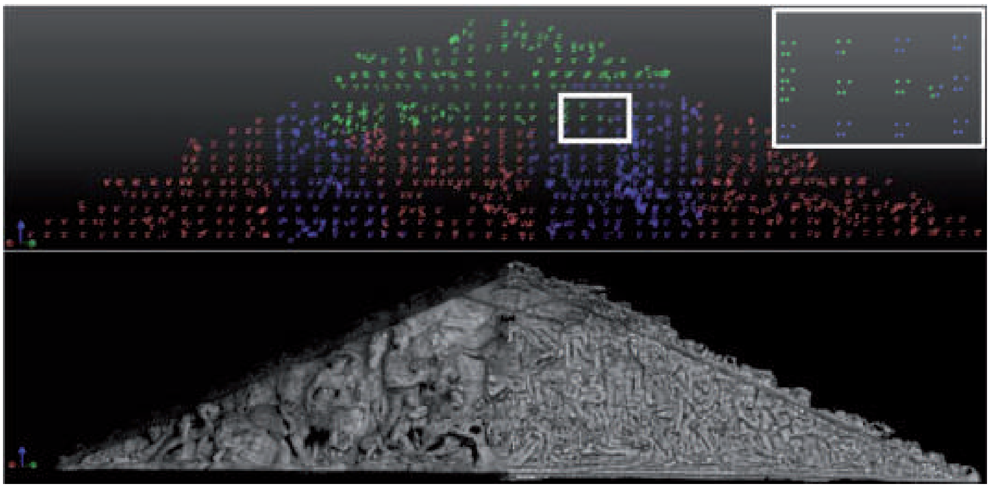


Fig. 6: Image orientations, randomly coloured, and sparse point cloud from SfM for the queen area within the top cluster of the tympanum at the west façade.

a relative accuracy in image space of better than 0.5 pixels is required for a reliable image matching. The control points were measured by reflectorless tachymetry. The mean standard deviation determined by the network adjustment amounts to 1.4 mm in position and 1.6 mm in height. However, many points were occluded and thus could not or only once be measured. However, the RMS values, see the third and the sixth rows in Tab. 2, derived from the residuals after the georeferencing cannot be used for an absolute accuracy assessment. Since a reference measurement with an accuracy of one magnitude higher would be required. In contrast, the accuracy of the tachymetric measurements is in a similar range. Therefore, these values are only used to validate the reconstructed orientations in object space.

8 Conclusions

The presented pipeline for the reconstruction of orientations and surface information is specifically designed for the efficient processing of large datasets with high accuracy requirements. An initial network analysis is used along with other techniques to realize a reasonable processing time while adjusting a stable bundle containing information from a maximum number of images. Thus, it is specifically suitable for large scale photogrammetric applications at low costs. In order to complement the measures of computational efforts it was our goal from the beginning that the whole processing pipeline should run on standard PC enviroments (for example i3 processors).

References

Aly, M., Munich, M. & Perona, P., 2011: Indexing in Large Scale Image Collections: Scaling Properties and Benchmark. – IEEE WACV 2011 Hawaii.

Bay, H., Ess, A., Tuytelaars, T. & Van Gool, L., 2006: SURF: Speeded Up Robust Features. – Computer Vision and Image Understanding: 346–359.

Brown, M. & Lowe, D., 2003: Recognizing panoramas. – 9th International Conference on Computer Vision **2:** 1218–1225.

Farenzena, M., Fusiello, A. & Gherardi, R., 2009: Structure and Motion Pipeline on a Hierarchical Cluster Tree. – ICCV Workshop on 3-D Digital Imaging and Modeling: 1489–1496.

Fritsch, D., Khosravani, A.M., Cefalu, A. & Wenzel, K., 2011: Multi-Sensors and Multiray Reconstruction for Digital Preservation. – Fritsch, D. (ed.): Photogrammetric Week '11: 305–323, Wichmann Verlag, Berlin/Offenbach.

Gibson, S., Cook, J., Howard, T., Hubbold, R. & Oram, D., 2002: Accurate Camera Calibration for Off-Line, Video-Based Augmented Reality.

– IEEE and ACM International Symposium Mixed and Augmented Reality: 37–46.

Hampel, F., Rousseeuw, P., Ronchetti, E. & Stahel, W., 1986: Robust Statistics: The Approach Based on Influence Functions. – Series in probability and mathematical statistics, Wiley, New York.

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

Klopschitz, M., Irschara, A., Reitmayr, G. & Schmalstieg, D., 2010: Robust Incremental Structure From Motion. – 3DPVT 2010.

Lourakis, A. & Argyros, A., 2009: SBA: A Software Package for Generic Sparse Bundle Adjustment. – ACM TOMS **36** (1): 1–30.

Muja, M. & Lowe, D., 2009: Fast Approximate Nearest Neighbors with Automatic Algorithmic Configuration. – VISAPP 2009.

Nistér, D., 2004: An efficient solution for the five-point relative pose problem. – IEEE PAMI **26** (6).

Snavely, N., Seitz, S. & Szeliski, R., 2007: Modeling the World from Internet Photo Collections. – International Journal of Computer Vision **80** (2): 189–210.

Vedaldi, A. & Fulkerson, B., 2008: VLFeat: An Open and Portable Library of Computer Vision Algorithms, http://www.vlfeat.org/ (28.8.2012).

Wenzel, K., Abdel-Wahab, M. & Fritsch, D., 2011: A Multi-Camera System for Efficient Point Cloud Recording in Close Range Applications. – LC3D workshop 2011: Berlin.

Address of the Authors:

M.Sc.Eng. Mohammed Abdel-Wahab, Dipl.-Ing. Konrad Wenzel and Prof. Dr.-Ing. Dieter Fritsch, University of Stuttgart, Institute for Photogrammetry, Geschwister-Scholl-Str. 24D, D-70174 Stuttgart, Tel: +49-711-68584117, Fax: +49-711-68583297, e-mail: {mohammed.othman}{konrad.wenzel}{dieter.fritsch}@ifp.uni-stuttgart.de

Manuskript eingereicht: Mai 2012
Angenommen: August 2012

PFG 2012 / 6, 0691 – 0700
Stuttgart, December 2012

Article

Zustandsdokumentation ausgedehnter untertägiger Hohlraumsysteme

Wilhelm Hannemann, Tilman Brock & Wolfgang Busch, Clausthal-Zellerfeld

Keywords: structure-from-motion, digital archaeology, subsurface objects, 3d reconstruction, photo-realism

Summary: *Documentation of conservation state in large-scale subsurface objects.* Scene reconstruction by structure-from-motion algorithms has received a great deal of attention due to the freely available software *Bundler*. We describe a workflow mainly based on *Bundler*, *PMVS2* and *Meshlab* for a three-dimensional documentation of conservation state in subsurface objects. The presented procedure is also suitable for locations difficult to access. Specific challenges in subsurface objects need to be addressed. This contribution describes the employed photo recording and processing procedure. As a result, photo-realistic textured models are exported in *WebGL* format, which is viewable in recent web browsers. Experiments were performed in the abandoned ore mine "Thurm Rosenhof" in Clausthal-Zellerfeld, the karst cave "Bielshöhle" in Rübeland and a drilling site of a potash mine. Two water wheel cambers and the connecting (and partly water-filled) adit of "Thurm Rosenhof" were photographically documented. 3D geometry and texture was reconstructed based on these images. The karst cave provides a second test object, which is characterized by a more irregular geometry and varying passage dimensions. The blast hole drilling site in potash mining provides an object where reference data from laser scanning is available, enabling comparisons to the data derived with the structure-from-motion based approach. Median point distance between both models was found to be 8 mm.

Zusammenfassung: Die dreidimensionale Rekonstruktion von Objekten mittels des Structure-from-Motion Algorithmus hat durch die frei erhältliche Software *Bundler* vermehrt Beachtung gefunden. Dieser Beitrag beschreibt einen Workflow zur dreidimensionalen Zustandsdokumentation untertägiger Hohlraumsysteme, welcher zu großen Teilen auf *Bundler*, *PMVS2* und *Meshlab* basiert. Das beschriebene Vorgehen kann auch an schwer zugänglichen Stellen eingesetzt werden. Untertägige Objekte stellen bei der Aufnahme spezielle Herausforderungen. Im Folgenden wird sowohl die Foto-Aufnahme vor Ort als auch der Workflow zur Prozessierung beschrieben. Als Endergebnis entstehen fotorealistisch texturierte Modelle, die in vielen gängigen Webbrowsern per *WebGL* dargestellt werden können. Experimente erfolgten in dem stillgelegten Erzbergwerk „Thurm Rosenhof" in Clausthal-Zellerfeld, der durch Verkarstung entstandenen Bielshöhle in Rübeland sowie einem Bohrort im Kali-Bergbau. Zwei Radstuben und die sie verbindende (und teilweise wasserführende) Rösche in „Thurm Rosenhof" wurden in mehreren Befahrungen fotografisch aufgenommen und aus diesen Digitalfotos die Geometrie und Textur rekonstruiert. Mit der Karsthöhle stand ein verwinkeltes und in den Raumdimensionen stark variierendes Testobjekt zur Verfügung. Der Bohrort im aktiven Kalibergbau wurde wegen zur Verfügung stehender Laserscanner-Referenzmessungen als weiteres Objekt ausgewählt. Ein Vergleich mit den Ergebnissen des Structure-from-Motion Ansatzes ergab für die punktweisen Abstände zwischen beiden Modellen einen Median von 8 mm.

1 Einleitung

Untertägige Objekte, wie Höhlen und Hohlräume des Altbergbaus, stellen im Gegensatz zu übertägigen Objekten besondere Anforderungen an eine dreidimensionale Erfassung, nicht nur durch den Wegfall der Nutzung von GNSS. Der Zugang kann technisch schwierig und beengt sein, ebenso die Situation vor Ort. Die Objekte sind teilweise sehr verwinkelt, so

DOI: 10.1127/1432-8364/2012/0149 1432-8364/12/0149 $ 2.50

dass sie für optische Verfahren nur kurze Zielweiten zulassen. Einige Bereiche können wasserführend sein, Luftfeuchtigkeit und Temperatur können stark schwanken. Die Anbringung von Zielmarken ist ggf. durch Auflagen des Natur- und Denkmalschutzes nicht möglich. Eine Stromversorgung oder Beleuchtung ist vor Ort meist nicht vorhanden. Durch diese Besonderheiten ist u.a. auch die Verwendung großer und empfindlicher Messinstrumente eingeschränkt.

Die Möglichkeit einer nahezu automatischen Rekonstruktion einer dreidimensionalen Punktwolke allein aus einer Menge von sich überlappenden Digitalfotos durch Software-Pakete wie Bundler (Snavely et al. 2006) oder APERO (Pierrot-Deseilligny & Cléry 2011) bot sich als Ergänzung zu bisherigen Verfahren wie z.B. der Verwendung von Distometern mit Inklinationsmessung und Kompass zur Höhlenkartierung (Heeb 2008) oder Messungen mit klassischer Tachymetrie bzw. Laserscanning (für Präzisionsmessungen) an. Das Mitführen einer Digitalkamera mit Blitzgerät ist fast überall möglich. Daher wurden Experimente mit dieser Vorgehensweise in Höhlen und Bergbau-Objekten praktisch durchgeführt (Hannemann et al. 2010, Hannemann et al. 2011).

Die verwendeten Software-Pakete zur Offline-Berechnung der dreidimensionalen Punktwolke werden in vielen Anwendungsbereichen eingesetzt, u.a. zur Visualisierung in der Archäologie (Ducke et al. 2011) oder zur Auswertung von Luftbildern autonom fliegender Systeme/UAVs (Klonowski & Neitzel 2011). Darüber hinaus existieren z.B. in der Robotik Software-Lösungen zur echtzeitfähigen Umgebungsrekonstruktion, welche teilweise weitere Sensoren mit einbeziehen (Engelhard et al. 2011). Hierfür ist allerdings ein leistungsstarker mobiler Rechner notwendig.

In diesem Beitrag werden zunächst die in den Experimenten verwendete Ausrüstung und der Prozessierungsablauf beschrieben. Die eingesetzten Open-Source-Programme und daran vorgenommene Modifikationen werden kurz beschrieben, ebenso eigene Ergänzungen. An Hand von mehreren praktischen Beispielen räumlich ausgedehnter untertägiger Hohlraumsysteme wird die praktische Durchführbarkeit demonstriert.

2 Vorgehensweise

2.1 *Verwendete Ausrüstung*

Durch die in der Einleitung angesprochenen Besonderheiten muss das mitgeführte Instrumentarium möglichst robust und klein sein. Es sollte relativ unempfindlich gegenüber Luftfeuchtigkeits- und Temperaturschwankungen sein und sich ggf. auch wasserdicht verpacken lassen. Die minimale Ausrüstung besteht verfahrensbedingt aus einer Digitalkamera, einer geeigneten Lichtquelle und einem mitgeführten Referenzobjekt zur Bestimmung des Maßstabs.

Als Digitalkamera ist erfahrungsgemäß eine robuste Spiegelreflexkamera gut geeignet. Je nach Objekt ist zusätzlich die Verwendung einer robusten Outdoor-Kompaktkamera sinnvoll. Als Beleuchtung erwies sich statt der Verwendung eines Blitzgeräts bei starker Luftfeuchtigkeit oder Staubentwicklung alternativ die Nutzung von batteriegetriebenen (LED-)Scheinwerfern als sinnvoll. Eine Kombination von Fotoserien verschiedener Kameras ist je nach Beschaffenheit der einzelnen aufgenommenen Objektabschnitte denkbar.

In den durchgeführten Experimenten wurde folgende Ausrüstung eingesetzt:

- digitale Spiegelreflexkamera (Nikon D50) mit 5 Megapixel Bildsensor,
- lichtstarkes Normalobjektiv (Nikkor 35 mm – F/1.8), zusätzlicher abnehmbarer Weitwinkel-Vorsatz (x 0.45),
- Blitzgerät (Sigma EF-500 DG ST),
- akkubetriebene Videoleuchte, LED-Scheinwerfer,
- Nivellierlatte (4 m, klappbar), Gliedermaßstab,
- Laserdistanzmessgerät (Leica D8).

2.2 *Aufnahme vor Ort*

Die Aufnahme der Fotos erfolgt von verschiedenen frei gewählten Standpunkten in der Regel freihändig. Dabei ist zu beachten, dass jeder Teil des Objekts auf mehreren Fotos von unterschiedlichen Kamerapositionen aus abgebildet ist. Benachbarte Fotos sollten überlappende gemeinsame Bereiche enthalten. Da für eine erfolgreiche Rekonstruktion der Objektgeometrie weniger auf Ästhetik der Aufnahmen als auf eine gute Abdeckung des Objekts geachtet werden muss, können durchaus mehrere hundert Bilder innerhalb einer Stunde aufgenommen werden.

2.3 *Prozessierungsablauf*

Die Prozessierung erfolgt in weiten Teilen automatisch mit wenigen notwendigen manuellen Eingriffen. Einige der im Folgenden aufgelisteten Teilschritte sind optional:

1. (optional) Entzerrung der Fotos nach erfolgter Kamera-Kalibrierung,
2. Berechnung markanter Punkte/Features,
3. Matching der markanten Punkte/Features,
4. Rekonstruktion der Kamerapositionen und -parameter (ggf. mit manueller Vorgabe von Startbildern),
5. Berechnung der dichten 3D-Punktwolken (ggf. manuelle Aufteilung in Objektabschnitte),
6. (optional) Glättung der Punktnormalen,
7. Berechnung einer Dreiecksvermaschung (ggf. Vorgabe der Berechnungstiefe),
8. (optional) manuelles Zuschneiden der Dreiecksvermaschung,
9. Reduzierung der Dreiecksanzahl,
10. Texturatlas-Erstellung,
11. Projektion der Fotos auf die Dreiecksvermaschung,
12. (optional) Transformation mehrerer Modelle in ein gemeinsames Koordinatensystem,
13. (optional) Maßstabskorrektur,
14. (optional) Konvertierung des texturierten Gesamtmodells in WebGL bzw. PDF.

Der Ablauf ist für eine automatisierte Prozessierung soweit möglich durch Shell-Scripte umgesetzt. Folgend werden die einzelnen Arbeitsschritte und die verwendeten Programme detaillierter erläutert:

Schritt 1: Eine Kalibrierung des verwendeten Kamerasystems liefert bessere Startwerte für die Bestimmung der Kameraparameter und kann damit Schritt 4 sowohl beschleunigen als auch die Ergebnisse verbessern. Zur Kalibrierung wird die „Camera Calibration Toolbox for MATLAB“ (Bouguet 2010) verwendet. Die anschließend durchgeführte Entzerrung der Fotos erfolgt mit dem Programm *Bundle2PMVS* aus *Bundler* (Bundler 2012).

Schritt 2: Die Berechnung der markanten Punkte und ihrer Deskriptoren erfolgt mittels des SIFT-Algorithmus (Lowe 2004); verwendet wird die Implementation aus *libsiftfast* (Libsift 2012).

Schritt 3: Das Matching der Punkte aus Schritt 3 wird mit einem gepatchen *KeyMatchFull* aus *Bundler* durchgeführt. Die im Programm hinzugefügte Angabe einer Teilmenge an zu matchenden Fotos ermöglicht die Aufteilung der Berechnung auf mehrere Computer. Ein „forken“ des KeyMatch-Prozesses erlaubt zusätzlich eine Parallelisierung auf Multicore-Maschinen ohne übermäßig zusätzlichen Hauptspeicher zu nutzen. Ergänzend wäre auch die Nutzung von GPUs zur Beschleunigung der Berechnung möglich (Wu 2007).

Schritt 4: Die Rekonstruktion der Kamerapositionen und -parameter erfolgt mit *Bundler*. Teilweise ist eine manuelle Vorgabe der Startbilder für die iterative Rekonstruktion nötig, um ein gutes Ergebnis zu erhalten. Über diese Vorgabe können auch einzelne Abschnitte des Objektes parallel berechnet werden; hierbei ist eine spätere Überführung in ein gemeinsames Koordinatensystem zu beachten. Eine weitere Möglichkeit zur Beschleunigung der Berechnung ist die Verwendung einer Ausschlussliste, um als Zwischenschritt bereits vollständig rekonstruierte Bereiche aus der weiteren Berechnung herauszunehmen und damit die interne Bündelausgleichung zu beschleunigen. Falls die Fotos vor der Berechnung entzerrt wurden, kann zunächst auf die Berechnung der Verzeichnung verzichtet werden. Nach Beendigung der Berechnung sollte auf dem Zwischenergebnis neu aufgesetzt werden, diesmal mit *Bundler*-interner Schätzung der Verzeichnung.

Schritt 5: Die Berechnung der in den weiteren Schritten verwendeten dichten Punktwolken erfolgt mittels *PMVS2* (Furukawa & Ponce 2007). *PMVS2* (PMVS2 2012) ermöglicht die Angabe eines „levels“, um die Fotos vor der Berechnung in ihrer Auflösung zu verkleinern. Es werden dabei je nach Auflösung teilweise unterschiedliche, sich ergänzende

Bereiche abgedeckt. Daher hat sich als sinnvoll herausgestellt, die Punktwolken für verschiedene Auflösungen zu berechnen und anschließend in einem gemeinsamen Datensatz zusammenzufassen.

Schritt 6: Das umfangreiche Programm *Meshlab* (MESHLAB 2012) ermöglicht u.a. eine Glättung der Punktnormalen anhand benachbarter Punkte. Dies verbessert ggf. das Ergebnis des nachfolgenden Schrittes, da *PMVS2* die Normalen anhand der Kamerapositionen abschätzt.

Schritt 7: Die Berechnung einer Dreiecksvermaschung aus der Punktwolke erfolgt mittels *Poisson Surface Reconstruction* (POISSON 2012). Hierbei gibt die mit Normalen versehene 3D-Punktwolke die Lösung einer Poisson-Gleichung vor. Die Dreiecksvermaschung enthält nicht notwendigerweise die Eingangspunkte, sondern nähert sich ihnen nur an (KAZHDAN et al. 2006). Das Ergebnis ist ein „wasserdichtes" Mesh ohne Lücken.

Schritt 8: Die Dreiecksvermaschung kann in Meshlab zugeschnitten werden. Große Dreiecke sollten ggf. entfernt werden, da in Schritt 7 auch Bereiche ohne Punkte gefüllt werden, um ein „wasserdichtes" Mesh zu erhalten.

Schritt 9: Die Anzahl der Dreiecke wird zur Reduzierung der Datenmenge und Beschleunigung einer späteren Darstellung in *Meshlab* mit dem Filter „Quadric edge collapse" verringert.

Schritt 10: Die Berechnung des Texturatlas erfolgt mittels ABF++ (SHEFFER et al. 2005). Das ist z.B. mit der Software Graphite (ALICE 2012) möglich.

Schritt 11: Die Fotos werden unter Verwendung der in Schritt 4 rekonstruierten Parameter auf das Mesh bzw. in den Texturatlas projiziert (eigene Implementation). Es wird je Textur-Pixel ein abstandsgewichteter Mittelwert aller RGB-Farbwerte der projizierten Fotos berechnet, da meist mehrere Fotos für einen Bereich des Objekts vorhanden sind (Abb. 1). Die Texturierung kann auf eine manuelle Vorauswahl gut geeigneter Fotos begrenzt werden. Eine Verbesserung des Verfahrens zur Texturierung ist zurzeit in Arbeit, da durch die Mittelwertbildung eine stellenweise unscharfe Textur entsteht.

Schritt 12: Falls das Objekt aus mehreren Teilmodellen zusammengesetzt ist, müssen diese in ein gemeinsames Koordinatensystem überführt werden. Hierzu werden die Teilmodelle über die in Schritt 2 berechneten markanten Punkte der Modelle fusioniert, indem mittels *RANSAC* (random sample consensus) die Parameter einer 3D-Helmert-Transformation berechnet werden. Alternativ ist auch die manuelle Auswahl von Passpunkten möglich, um daraus anschließend die Transformationsparameter zu bestimmen.

Schritt 13: Zur Maßstabskorrektur muss ein in der Realität bekanntes Längenmaß im Modell abgegriffen und das Modell entsprechend skaliert werden. Das Abgreifen kann direkt in der 3D-Punktwolke erfolgen. Um genauere Ergebnisse zu bekommen, wurde die Berechnung von Punktkoordinaten aus manu-

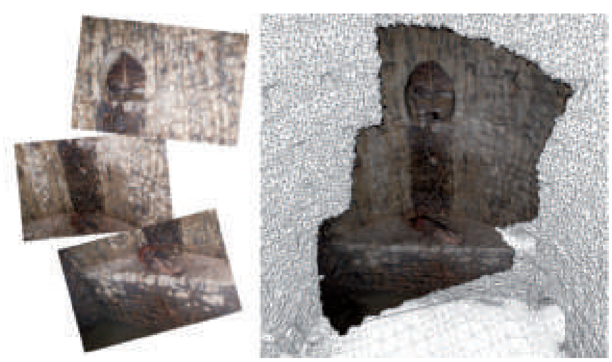


Abb. 1: Texturprojektion.

ell identifizierten Punkten in den Originalfotos implementiert. In mehreren Fotos werden Anfangs- und Endpunkt der gesuchten Länge markiert und anschließend über die bekannten Projektionsmatrizen der Fotos mittels Ausgleichungsrechnung die 3D-Koordinaten der Längenreferenz, und daraus wiederum die Länge der gesuchten Strecke in Modelleinheiten ermittelt. Über den so gewonnenen Maßstabsfaktor kann das Modell dann metrisch skaliert werden.

Schritt 14: Als Darstellungsformat bieten sich wegen der weiten Verbreitung geeigneter Viewer sowohl *WebGL* als auch PDF-Dateien mit eingebettetem 3D-Inhalt an. Über einen Export aus *Meshlab* kann *LaTeX*-Code und aus diesem dann PDF-Dateien erzeugt werden. Zur Darstellung in WebGL-fähigen Browsern, z.B. *Mozilla Firefox* oder *Google Chrome*, wird auf die Bibliothek *three.js* (ThreeJS 2012) zurückgegriffen. Die Modelle müssen für die Verwendung mit *three.js* einmalig in ein Präsentationsformat konvertiert werden. Bewegung und Blickwinkelveränderung im Modell erfolgen mittels Maus (Blickrichtung) und Tastatur (Bewegung). Einzelne Positionen und Blickwinkel können über Tastenkürzel angesteuert werden. Die Beleuchtung der Szene kann dynamisch angepasst werden.

3 Experimente in untertägigen Objekten

3.1 *Radstuben der Grube „Thurm Rosenhof“*

Als ein räumlich ausgedehntes Altbergbauobjekt wurden die untertägigen Rosenhöfer Radstuben der ehemaligen Grube „Thurm Rosenhof“ in Clausthal-Zellerfeld gewählt. Hier wurde seit Mitte des 16. Jahrhunderts bis in das 20. Jahrhundert hinein silberreiches Erz abgebaut. Ein in jeder Radstube installiertes Wasserrad diente der Energiegewinnung für den Grubenbetrieb. Das Wasser wurde über horizontale Gänge bis zur Tagesoberfläche geführt. Die am westlichen Ortsrand gelegene „Runde Radstube“ ist seit einigen Jahren touristisch erschlossen und über ein neu eingebautes Treppenhaus begehbar. Sie hat einen Durchmesser von etwa 10 m bei einer maximalen Tiefe von 24 m (Abb. 2). Von dieser Radstube aus ist über die ehemalige Abfallrösche (wasserführender Gang) die derzeit nicht touristisch zugängliche „Ovale Radstube“ erreichbar. Sie ist rund 12,50 m hoch und liegt ca. 10 m unter der heutigen Geländeoberfläche.

3.1.1 Aufnahme vor Ort

In vier jeweils ca. zweistündigen Befahrungen wurden die Radstuben und die sie verbindende Rösche mit insgesamt rund 2600 Fotos aufgenommen. Die „Runde Radstube“ konnte von der eingebauten Treppe aus hinreichend gut aufgenommen werden; die Messung einer Referenzstrecke für die spätere Maßstabsbestimmung war problemlos möglich. Um Erfahrungswerte für Objekte zu sammeln, bei denen diese einfache Möglichkeit nicht gegeben ist, wurde eine Seilstrecke bis zur Sohle (dem Boden) der Radstube aufgebaut (Abb. 2). Die vertikalen Abstände der vom Seil aus aufgenommenen Fotos wurden näherungsweise festgehalten, um später den Maßstab abschätzen zu können.

Die von der „Runden Radstube“ wegführende Abfallrösche ist bis zu einem Meter Tiefe mit Wasser gefüllt. Um einhändig mit der Spiegelreflexkamera arbeiten zu können, wurde die Autofokus-Funktion verwendet. Helmlampen reichen als Beleuchtung für den Autofokus aus.

Ab dem Kreuzungspunkt zur Ovalen Radstube folgt eine Kriechstrecke bis in die Ovale Radstube. Auch hier war es von Vorteil, die Kamera einhändig bedienen zu können. Der Weitwinkelvorsatz diente in dieser Strecke zusätzlich als Schutz für das Kameraobjektiv.

In der Ovalen Radstube konnte von der Sohle aus eine hinreichende Anzahl Fotos erstellt werden.

3.1.2 Prozessierung und Ergebnisse

Die Prozessierung der Daten erfolgte wie in Abschnitt 2.3 beschrieben. Zeitaufwendig waren sowohl das Matching als auch die Berechnung mit *Bundler*. Das Matching wurde innerhalb einiger Tage auf einem Mehrprozessorsystem berechnet. Obwohl die Bündelausgleichung parallel in drei unterschiedlichen

Abschnitten gerechnet wurde, dauerte die Berechnung mehrere Tage. Dabei entstanden die Modelle der Radstuben jeweils mit angrenzendem Abschnitt der Rösche und ein komplettes Modell der Rösche.

Die Visualisierung des Gesamtmodells für den aufgenommenen Bereich der Rosenhöfer Radstuben erfolgte in einem Geoinformationssystem. Abb. 3 zeigt eine perspektivische Ansicht dieses Modells.

Zur Darstellung der Ergebnisse kann es je nach Zielsetzung sinnvoll sein, die Textur ein- oder auszublenden. In Abb. 4 ist links die Geometrie als schattiertes Dreiecksnetz zu se-

Abb. 2: Runde Radstube mit Seilstrecke.

hen, rechts daneben das texturierte Modell. In der schattierten Ansicht sind die unterschiedlichen Bauabschnitte der „Runden Radstube“ anhand der Größe der verwendeten Mauersteine gut zu erkennen; auch die Auflager der ehemals eingezogenen Balken sind sichtbar. In der entsprechenden texturierten Ansicht ist der Oberflächenzustand, Moosbewuchs usw., erkennbar.

In Abb. 5 ist die Rösche in einer perspektivischen Ansicht zu sehen. In der Kriechstrecke sind zwei kleine Abzweige zu erkennen. Diese sind wenige Meter lang und nur bäuchlings liegend zugänglich.

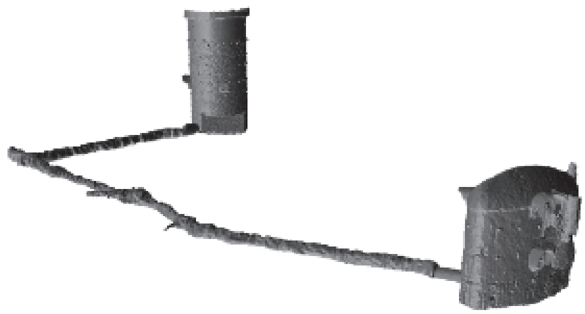


Abb. 3: Perspektivische Ansicht Radstuben und Rösche.

Abb. 4: Runde Radstube, Geometrie und Textur.

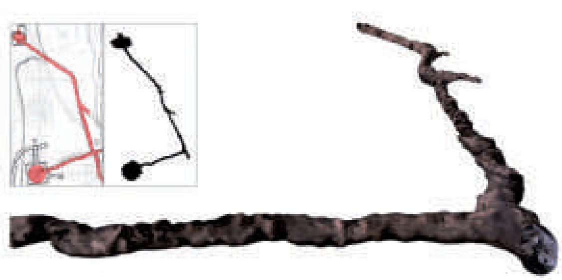


Abb. 5: Perspektivische Ansicht der Rösche (groß); Lageskizze und Aufsicht 3D-Modell (links oben).

3.2 *Bielshöhle bei Rübeland*

Als weiteres Testobjekt wurde der vordere Teil der ehemaligen Schauhöhle „Bielshöhle“ bei Rübeland im Harz ausgewählt. Die Höhle bietet bedingt durch den natürlichen Entstehungsprozess der Verkarstung eine unregelmäßige und somit „anspruchsvolle“ Geometrie. Abb. 6 zeigt einen nur bäuchlings kriechend zugänglichen Bereich, den sogenannten Fledermausgang.

Der insgesamt ca. 40 m lange Zugang konnte zusammen mit der sogenannten Teilungshalle in zwei jeweils ca. dreistündigen Befahrungen aufgenommen werden. Der Zugang besteht aus dem Eingangsbereich, dem Schlangengang und dem davon abzweigenden parallelen Fledermausgang. Es entstanden insgesamt 2000 Digitalfotos. In Abb. 7 ist eine Aufsicht auf das errechnete Modell zu sehen. Darin ist der Rundschluss aus Schlangengang, Fledermausgang und Teilungshalle zu erkennen. Dass dieser nahezu 100 m lange Rundweg während der Prozessierung ohne Zwang bündig geschlossen werden konnte, spricht für die gute Qualität des Verfahrens, zumal der Rundschluss in einem sehr verwinkelten Objektteilbereich mit nur jeweils sehr kleinen einsehbaren Bereichen (Abb. 6) erfolgte.

Abb. 6: Fledermausgang, kriechend befahren.

Abb. 7: Bielshöhle, Darstellung mit WebGL.

3.3 *Bohrort im Kali-Bergbau*

Bei den in Zusammenarbeit mit einem Bergbauunternehmen durchgeführten Versuchen zur Erstellung dreidimensionaler Modelle untertägiger Abbaue konnte ein Vergleich mit parallel erstellten Aufnahmen eines Laserscanners durchgeführt werden (Thiele et al. 2011).

Für einen Bohrort mit ca. 15 m Länge und Breite und einer Höhe von ca. 3 m standen sowohl eine wie oben beschriebene fotobasierte Rekonstruktion als auch ein Laserscan (Laserscanner: Faro Focus 3D) zur Verfügung. Zum Vergleich der beiden Modelle wurde die fotobasierte Rekonstruktion mit 1,5 Mio. 3D-Punkten in das Modell (Koordinatensystem) des Laserscans (27 Mio. 3D-Punkte) überführt. Hierzu wurde die Punktwolke des Laserscans auf eine zylindrische Fläche projiziert, um anschließend mit dem SIFT-Algorithmus markante Punkte zu detektieren und beide Auswertungen zu fusionieren. Die mittels *RANSAC* bestimmte 7-Parameter-Transformation zwischen den Koordinatensystemen erfolgte über 372 Punktpaare. Die mittlere Klaffung ergab sich zu 1,1 mm.

Der Median der Abweichungen zwischen den 1,5 Mio. Punkten aus der fotobasierten Rekonstruktion zu den jeweils räumlich nächsten Punkten des Laserscans beträgt 8 mm. 67% der Punkte liegen maximal 11,2 mm von den räumlich nächsten Punkten des Laserscans entfernt. Abb. 8 zeigt das Histogramm der Abweichungen.

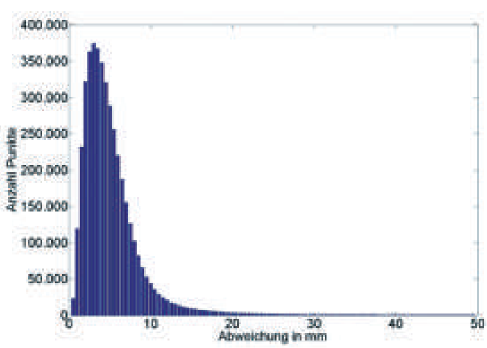


Abb. 8: Histogramm der Abweichungen fotobasierte 3D-Rekonstruktion / Laserscan beim Bohrort im Kalibergbau.

4 Zusammenfassung und Fazit

Es wurde anhand mehrerer Objekte die Anwendbarkeit eines fotobasierten Rekonstruktionsverfahrens zur Erstellung dreidimensionaler Modelle untertägiger Hohlraumsysteme demonstriert. Die zur Aufnahme vor Ort notwendige Ausrüstung ermöglicht auch unter beengten räumlichen Verhältnissen die Erzeugung detailreicher und qualitativ ausreichender 3D-Modelle.

Das gezeigte Verfahren zur Erstellung von 3D-Modellen lässt sich mit relativ geringem Zeitaufwand vor Ort anwenden. Lediglich die automatisierte Auswertung der Daten benötigt auf handelsüblichen Computern je nach Umfang der Aufnahmeserien mehrere Tage. Die verwendete Ausrüstung ist sehr kostengünstig. Auch schwierig zugängliche oder verwinkelte Bereiche sind erfassbar, solange sie von Menschen begehbar sind. Dies wurde am Beispiel eines Altbergbauobjektes und einer größeren, durch Verkarstung entstandenen Höhle gezeigt. Ein dort rekonstruierter Rundweg deutet darauf hin, dass für die untersuchten Objekte auch über große Strecken eine geringe Winkel- und Streckenabweichung erzielt werden kann.

Das verwendete Verfahren zur Texturierung des 3D-Modells aus den Originalfotos kann noch verbessert werden. Bei der Rekonstruktion der Kamerapositionen ist noch zu untersuchen, welche Vorteile sich durch die Verwendung blickwinkel-unabhängiger Features (Wu et al. 2008, Wu 2011) in untertägigen Objekten ergeben. Auch kommerzielle Programme wie *PhotoScan* (Agisoft 2012) bieten eine Alternative.

Die Präsentation mit WebGL (Abb. 7) ist eine einfache Möglichkeit, um ohne dedizierte 3D-Viewer-Software und ohne Browser-Plugins die gewonnenen Modelle für die Öffentlichkeit zu präsentieren. Für einige der hier aufgeführten Beispiele sind entsprechende WebGL-Umsetzungen unter IGMC (2012) abrufbar. Bei der Darstellung der Modelle mit WebGL sind weitere Entwicklungen denkbar. Zur verbesserten Zustandsdokumentation ist eine Annotationsmöglichkeit an beliebigen Modellpositionen wünschenswert. Dadurch würde es möglich werden, schnell auf die dem Modell zugrunde liegenden Detailfotos, Messwerte oder andere ortsbezogene Informationen zuzugreifen, die bisher wenn überhaupt nur über komplexe Desktop-Applikationen zugänglich sind.

Literaturverzeichnis

AGISOFT, 2012: http://www.agisoft.ru/products/photoscan/ (26.5.2012).

ALICE, 2012: http://alice.loria.fr/index.php/software/3-platform/22-graphite.html (26.5.2012).

BOUGUET, J.-Y., 2010: Camera Calibration Toolbox for Matlab. – http://www.vision.caltech.edu/bouguetj/calib_doc/ (29.8.2011).

BUNDLER, 2012: http://phototour.cs.washington.edu/bundler/ (26.5.2012).

DUCKE, B., SCORE, D. & REEVES, J., 2011: Multiview 3D reconstruction of the archaeological site at Weymouth from image series. – Computers & Graphics **35** (2), ISSN 0097-8493.

ENGELHARD, N., ENDRES, F., HESS, J., STURM, J. & BURGARD, W., 2011: Real-time 3D visual SLAM with a hand-held RGB-D camera. – RGB-D Workshop on 3D Perception in Robotics at the European Robotics Forum, Västerås, Sweden.

FURUKAWA, Y. & PONCE, J., 2007: Accurate, Dense, and Robust Multi-View Stereopsis. – IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, ISBN 1-4244-1180-7.

HANNEMANN, W., BROCK, T., HEBEL, H.-P. & BUSCH, W., 2010: Photogrammetrische Hohlraumerfassung einer Radstube. – Tagungsband Altbergbaukolloquium **10:** 43–57, TU Bergakademie Freiberg, VGE Verlag.

HANNEMANN, W., BROCK, T. & BUSCH, W., 2011: Fotobasierte 3D Zustands- und Geometriedokumentation altbergbaulicher Objekte. – Tagungsband Altbergbaukolloquium **11:** 34–45, TU Wrocław, Polen, VGE Verlag.

HEEB, B., 2008: Paperless Caving – An Electronic Cave Surveying System. – **4th** European Speleological Congress, Vercors, Frankreich.

IGMC, 2012: http://www.igmc.tu-clausthal.de/index.php?id=foto3d (25.8.2012).

KAZHDAN, M., BOLITHO, M. & HOPPE, H., 2006: Poisson Surface Reconstruction. – **4th** Eurographics Symposium on Geometry Processing, Cagliari, Italien.

KLONOWSKI, J. & NEITZEL, F., 2011: Mobile 3D-Mapping with a Low-Cost UAV System. – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XXXVIII** (1/C22), Conference on Unmanned Aerial Vehicle in Geomatics; Zürich, ISSN 1682-1777.

LIBSIFT, 2012: http://sourceforge.net/projects/libsift/ (26.5.2012).

LOWE, D.G., 2004: Distinctive image features from scale-invariant keypoints. – International Journal of Computer Vision **60** (2), ISSN 0920-5691.

MESHLAB, 2012: http://meshlab.sourceforge.net/ (26.5.2012).

PMVS2, 2012: http://grail.cs.washington.edu/software/pmvs/ (26.5.2012).

PIERROT-DESEILLIGNY, M. & CLÉRY, I., 2011: APERO, an Open Source Bundle Adjustment Software for Automatic Calibration and Orientation of a Set of Images. – ISPRS Commission V Symposium, Image Engineering and Vision Metrology, Trento, Italy.

POISSON, 2012: http://www.cs.jhu.edu/~misha/code/poissonrecon/ (26.5.2012).

SHEFFER, A., LÉVY, B., MOGILNITSKY, M. & BOGOMYAKOV, A., 2005: ABF++: fast and robust angle based flattening. – ACM Transactions on Graphics **24**.

SNAVELY, N., SEITZ, S.M. & SZELISKI, R., 2006: Photo tourism: Exploring photo collections in 3D. – ACM Transactions on Graphics, SIGGRAPH, Boston, Massachusetts, ISBN 1-59593-364-6.

THIELE, J.C., HANNEMANN, W. & BUSCH, W., 2011: VR-Methoden in der Ausbildung bei K+S Aktiengesellschaft. – Tagungsband Energie und Rohstoffe **2011**, Beitrag des Markscheidewesens: 209–216, TU Bergakademie Freiberg, Papierflieger Verlag.

THREEJS, 2012: https://github.com/mrdoob/three.js/blob/master/readme.md (26.5.2012).

WU, C., 2007: SiftGPU: A GPU Implementation of Scale Invariant Feature Transform (SIFT), http://cs.unc.edu/~ccwu/siftgpu/ (29.8.2011).

WU, C., CLIPP, B., LI, X., FRAHM, J.-M. & POLLEFEYS, M., 2008: 3D Model Matching with Viewpoint Invariant Patches (VIPs). – CVPR 2008.

WU, C., 2011: VisualSFM: A Visual Structure from Motion System, http://www.cs.washington.edu/homes/ccwu/vsfm/ (9.1.2012).

Adresse der Autoren:

WILHELM HANNEMANN, TILMAN BROCK & Prof. Dr. WOLFGANG BUSCH, Institut für Geotechnik und Markscheidewesen, TU Clausthal, 38678 Clausthal-Zellerfeld, Tel.: +49-5323-72-2294, Fax: +49-5323-72-2479, e-mail: {wilhelm.hannemann}{tilman.brock}{wolfgang.busch}@tu-clausthal.de

Manuskript eingereicht: Mai 2012
Angenommen: August 2012

PFG 2012 / 6, 0701 – 0710
Stuttgart, December 2012

Article

Graph-Based Analysis of Pedestrian Interactions and Events Using Hidden Markov Models

Florian Burkert & Richard Bamler, München

Keywords: event detection, pedestrian surveillance, aerial image sequences, hidden Markov models, UAV

Summary: In this paper, we present an improved approach for the analysis of pedestrian interaction in crowded and cluttered scenes from aerial image sequences. Related work is limited to the detection of an undeclared abnormal event with regard to the common behaviour or to the detection of specific simple events incorporating only up to two trajectories. In our approach, event detection in pedestrian groups is done by detecting universal motion interaction patterns between pairs of pedestrians in a graph-based framework. Event detection is performed by analyzing the temporal behaviour of the motion interaction, which is represented by edges in the graph, by means of hidden Markov models (HMM). Temporarily disappearing edges in the graph can be compensated by an HMM buffer which internally continues the HMM analysis even if the corresponding pedestrians depart from each other awhile. Experimental results show the potential of our graph-based approach for event detection. The used datasets contain UAV image sequences in which an instructed pedestrian group simulates meaningful group behaviour and an aerial image sequence in which pedestrians approach a soccer stadium.

Zusammenfassung: *Graphenbasierte Ereignisdetektion von Fußgängerinteraktion mittels Hidden Markov Modellen*. In diesem Beitrag wird eine verbesserte Methode für die Detektion von Fußgänger-Interaktion in dichten und unstrukturierten Szenen aus Luftbildsequenzen vorgestellt. Bislang bestehende Arbeiten beschränken sich auf die Erkennung von unnormalen Ereignissen im Allgemeinen oder auf die Erkennung von einfachen Ereignissen, welche nur von bis zu zwei Personen durchgeführt werden. In der hier vorgestellten Methode wird Ereignisdetektion in Personengruppen vollzogen, wofür die Bewegungsinteraktion zwischen benachbarten Personenpaaren in einem graphenbasierenden System analysiert wird. Das zeitliche Verhalten der Bewegungsinteraktion wird mittels Hidden Markov Modellen (HMM) ausgewertet. Zeitlich unbeständige Kanten im Graph werden mit Hilfe eines HMM-Puffers abgefangen, welcher die Auswertung intern weiterführt, wenn sich das einer Kante zugehörige Personenpaar kurzzeitig voneinander entfernt. Es werden Ergebnisse präsentiert, welche das Potential der vorgestellten Methode zur Ereignisdetektion aufzeigen. Die verwendeten Datensätze beinhalten UAV-Sequenzen, welche Gruppenbewegungen eingewiesener Testpersonen beinhalten, und Luftbildsequenzen, welche Fußgänger vor einem Fußballstadion zeigen.

1 Introduction

The main objective of this work is event detection in crowds by robustly analyzing a pedestrian interaction graph using Hidden Markov Models in which the edges represent motion interaction patterns between pedestrians.

The huge amount of surveillance data requires automatic or at least semi-automatic interpretation. Consequently, research in crowd analysis has been intensified in the last decades in order to support human surveillance operators. In addition to purely image based crowd analysis techniques, crowd models from psychology, physics or from the nature have to be incorporated into more sophisticated surveillance systems (Zhan et al. 2008, Butenuth et al. 2011). Aerial imagery pro-

DOI: 10.1127/1432-8364/2012/0150 1432-8364/12/0150 $ 2.50

vides a wide overview over a scene and can, therefore, ideally be used to extract trajectories of pedestrians which can then be used for event detection.

Numerous publications indicate the importance of crowd analysis. Zhan et al. (2008) present a survey which recapitulates contributions to object detection, tracking and event detection. The main input data for event detection are either trajectories or optical flow. Event detection systems using optical flow are able to detect abnormal events in high density crowds after observing the common behaviour for some time (Adam et al. 2008, Andrade et al. 2006, Mehran et al. 2009). However, individual behaviour cannot be inferred by optical flow and no classification of the type of the unusual event is made besides of flow-specific characteristics. In scattered and medium-dense scenes, the analysis of discrete trajectories is preferred because of the visibility of individuals. For the analysis of discrete trajectories, hidden Markov models (HMM) (Rabiner 1989) have often been applied in the past, which have originally been developed for speech recognition. Several specialisations of HMM have been developed and utilized for event detection, such as coupled HMM (CHMM) (Oliver et al. 2000) or switched dynamical HMM (SD-HMM) (Nascimento et al. 2010). Nascimento et al. (2010) classify recurring human trajectories in busy scenes by concatenating a given set of low level models using switched dynamical hidden Markov models (SD-HMM). Human trajectory mining is performed in the work of Calderara & Cucchiara (2010) by clustering frequent behaviours of pedestrians using different similarity measures. Kuettel et al. (2010) automatically learn spatio-temporal dependencies of moving agents and show experimental results from traffic scenes. However, by classifying or mining recurring trajectories, only a very stringent model containing some trajectory clusters can be built which is not flexible enough to cope with individual and spontaneous motion patterns in cluttered scenes that do not match recurring paths. Learning of recurring trajectories can also be used for the detection of unusual events (Basharat et al. 2008, Hu et al. 2006, Porikli & Haga 2004). Unusual events can only be detected if enough data about usual events is available beforehand. To this end, the scene has to provide specific conditions which can be followed by the majority of the observed objects, like entrance doors for pedestrians or driving lanes for vehicles.

We overcome the limitations of the related work. We use manually generated trajectories in order to be able to draw significant information about individuals' motion behaviour. The analysis of the entire scene is achieved by modelling all pedestrians in a graph at each frame. We calculate four extended motion features adapted from Burkert & Butenuth (2011) to deduce six universal motion patterns for each pair of trajectories. The motion patterns which describe the motion interaction of a pair of pedestrians correspond to the edges in the graph. The sequential behaviour of the motion patterns is analyzed by HMM. We focus on the detection of simple and universal motion patterns which allows us to interpret resulting large scale clusters of motion patterns but also individual events in the scene. At this level, findings from social and traffic sciences such as the social force model (SFM) Helbing & Molnár (1995) can be used. We show experimental results for our event detection system based on a dataset acquired by an unmanned aerial vehicle (UAV) in which an instructed group of pedestrians fulfills meaningful scenarios of group behaviour. Another real-world dataset contains an airborne image sequence in which pedestrians approach a soccer stadium.

The outline of this paper is as follows: In section 2 we introduce the terminology of HMM we use in this paper. Section 3 describes our system for robust, graph-based event detection. In section 4 we show experimental results and in section 5 we conclude and discuss future work.

2 Terminology of Hidden Markov Models (HMM)

A hidden Markov model (HMM) is a probabilistic model which is represented by a directed acyclic graph. A HMM shows the simplest form of a dynamic Bayesian network. The system underlying the HMM is a Markov chain of hidden states. At each time step, an observable output of the model is generated which only depends on the current hidden state.

2.1 *Parameters of an HMM*

An HMM provides clear Bayesian semantics and is defined by the following set of parameters (Rabiner 1989):

- A set of N possible hidden states $\{s_1, s_2, \ldots, s_N\}$, the state at time t is denoted as q_t.
- A set of M possible observations $\{v_1, v_2, \ldots, v_M\}$, the observation at time t is denoted as o_t.
- The transition probability matrix A with its elements a_{ij} denoting the transition probabilities from s_i to s_j

$$a_{ij} = P(q_t = s_j \,|\, q_{t-1} = s_i) \tag{1}$$

$$\sum a_{ij} = 1 \forall i, \quad a_{ij} \geq 0 \tag{2}$$

- The observation probability distribution $b_j(v_k)$ for an observation v_k

$$b_j(v_k) = P(o_t = v_k \,|\, q_t = s_j) \tag{3}$$

$$\sum_k b_j(v_k) = 1 \, \forall i, \quad b_j(k) \geq 0 \tag{4}$$

- The initial probabilities π_i that s_i is the initial state

$$\pi_i = P(q_1 = s_i) \tag{5}$$

$$\sum_i \pi_i = 1, \quad \pi_i \geq 0 \tag{6}$$

From this set of parameters the transition probability matrix A, the observation probabilities $b_j(v_k)$ and the initial probabilities π_i can be subsumed under a parameter λ which characterizes an HMM.

2.2 *Inference in HMM*

The inference in HMM to find the most probable sequence of hidden states $\{q_1,q_2,\ldots,q_T\}$ is performed by using the corresponding given sequence of observations $\{o_1,o_2,\ldots,o_T\}$, where T is the length of the sequence. This problem can be solved by the *Viterbi algorithm* (Rabiner 1989) which is used if a complete and terminated sequence of observations is available. For our problem of event detection in real-time, which operates as the image sequence is acquired, filtering has to be used instead of the Viterbi algorithm. Filtering is used for computing the probability distribution over the hidden states $\{s_1,s_2,\ldots,s_N\}$ at a certain time step t, given the sequence of observations up to this time step $\{o_1,o_2,\ldots,o_{t-1},o_t\}$. Filtering can efficiently be solved by the *forward algorithm* (Rabiner 1989). The forward algorithm is appropriate for our task because it does not depend on an already terminated sequence and, thus, can be iteratively applied at every new frame of the image sequence.

The forward algorithm employs forward variables $\alpha_t(s_i)$,$1 \leq i \leq N$ which are calculated at each time step t for every hidden state s_i. Thus, the forward variables are defined as

$$\alpha_t(s_i) = P(o_1, o_2, \ldots, o_t, q_t = s_i \,|\, \lambda) \,. \tag{7}$$

$\alpha_t(s_i)$ is the probability to produce the observation sequence up to t and to reach state s_i at time t, given the HMM λ. At the first time step $t = 1$, the initialization of the forward algorithm is realized by

$$\alpha_1(s_i) = \pi_i b_i(o_1) \,. \tag{8}$$

The initialization of the forward variables depends on the initial probabilities π_i and the first observation o_1. At further time steps t, $2 \leq t \leq T$, the recursion of the forward algorithm is performed by

$$\alpha_t(s_j) = b_j(o_t) \sum_i \alpha_{t-1}(s_i) \, a_{ij} \,. \tag{9}$$

The recursion step depends on the observation o_t and on all forward variables of the previous time step $\alpha_{t-1}(s_i)$, multiplied by their probabilities of transition to state s_j.

3 Graph-Based Event Detection

In order to analyze motion interaction patterns in crowds we create a pedestrian interaction graph which contains all pedestrians of a scene. The analysis is done for existing edges in the graph by HMM-based event detection which is robust to fluctuant and partially departing pedestrians.

3.1 *Pedestrian Motion Model for Pedestrian Interaction*

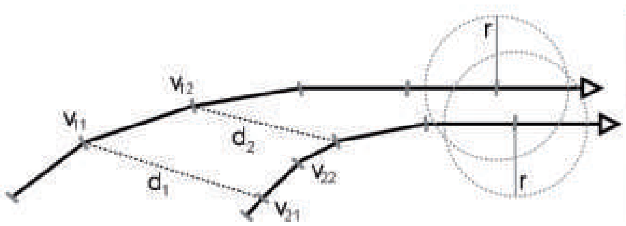
The motion model consists of four motion features which are refined and extended versions of those defined in Burkert & Butenuth (2011). The motion features are the observations for the HMM because they can ideally be used to infer six universal motion patterns. The motion patterns are treated as the events which have to be detected and are then used to interpret a crowd's behaviour.

Motion Features

We use four motion features which are calculated from trajectory pairs at every time step. A pair of trajectories is defined by two pedestrians i and j which are sufficiently close to each other such that a significant motion interaction takes place. The method to specify significant motion interaction is described in section 3.2. The four motion features for two pedestrians i and j are the sum of the velocities v_i+v_j, the variation of the distance Δd, the average pedestrian density $\mu(\rho_i+\rho_j)$ and the normalized scalar product of both motion direction vectors s. The motion features serve as the observations in the HMM. Fig. 1 (left) depicts two trajectories illustrating the motion features v_i+v_j, Δd and $\mu(\rho_i+\rho_j)$. The velocity v_i of a pedestrian is calculated at each time step using the frame rate and the covered distance since the last time step. The variability of the distance is defined as $\Delta d=d_t / d_{t-1}$, with d_{t-1} being the distance at time step t-1 and d_t being the distance at time step t. Thus, $\Delta d > 1$ for an increasing distance and $\Delta d < 1$ for a decreasing distance. The local pedestrian density ρ_i is calculated by the inverse of the area of the corresponding cell in a Voronoi diagram, which is constructed from the pedestrian locations at each frame (Steffen & Seyfried 2010). By using a Voronoi diagram, the local pedestrian density can be calculated instead of using a fixed area in which the number of pedestrians is counted. Only for those pedestrians which are located at the border of groups, the density is calculated by counting the number of pedestrians within a specified radius r and relating it to the area. The reason for this exception is that Voronoi cells at the border of a point cluster will receive very large or infinite area (Fig.1, right). We further introduce a fourth motion feature which is the normalized scalar product s of both motion direction vectors. s receives values up to 1 for pedestrians walking in parallel (Fig. 1, left), values of about 0 for orthogonal vectors, and up to -1 for pedestrians walking in opposite directions. Thus, s complements the feature Δd by describing the type of the distance variation. Δd only gives the absolute variation of the distance between two pedestrians but does not specify the directions, in which the distance variation is fulfilled. In summary, the feature vector is [v_i+v_j, Δd, $\mu(\rho_i+\rho_j)$, s].

Motion Patterns

We define six motion patterns which occur when pedestrians are close to each other, namely *Standing*, *Queueing*, *Walking*, *Running*, *Diverging* and *Converging*. These simple and universal motion patterns are the basis for our event detection system and define the type of the motion interaction between neighbouring pedestrians. The speed of pedestrians



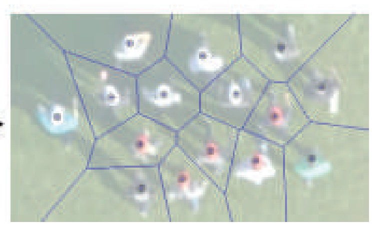


Fig. 1: Motion features derived from trajectories; left: v_{it} is the velocity of trajectory i at time t, d_t is the distance at frame t and r is the radius in which the pedestrian density at a group boundary is computed; right: Voronoi diagram of pedestrian locations, pedestrian density is defined as the inverse cell size.

can be variable at the motion patterns *Converging* and *Diverging* and is at most 0.1, 0.3, 2.0 and 4.5 m/s for the motion patterns *Standing*, *Queueing*, *Walking*, *Running*, respectively. The variability of distance explicitly leads to a statement if two pedestrians approach or depart from each other, independent of the motion direction. Therefore, the variability of distance is about 1 for *Standing* and the parallel motion patterns *Queueing*, *Walking* and *Running*. Consequently, $\Delta d > 1$ for *Diverging* and $\Delta d < 1$ for *Converging*. The pedestrian density is variable but rather low when pedestrians are standing or moving naturally and can reach high values up to 5 pedestrians per m^2 in dense queuing areas. The normalized scalar product of the motion direction vectors is variable for standing pedestrians because a slight motion into a random direction always occurs and no pedestrian will stand completely still. The motion patterns *Queueing*, *Walking* and *Running* are characterized by pedestrians walking in parallel, so the scalar product is close to 1. The scalar product emphasizes parallel scenarios in which the variability of distance might misleadingly suggest diverging or converging motion patterns because of different velocities.

Usually, training data from the real world surveillance scenes is used to learn HMM offline. However, we do not use data from surveillance cameras to learn the HMM for event detection because we focus on cluttered scenes which occur at big events. Learning from real world data always relies on frequently recurring motion paths within the scene of interest, which we cannot assume to be available for any place where a big event may take place. Instead, we use synthetic data representing the motion patterns to learn the HMM. The training data are generated by moving agents which follow rules of motion depending on the description above. The training data consist of 900–1200 feature vectors per motion pattern which were calculated from normally distributed simulated trajectories. From these feature vectors, the mean values are used to derive the feature vector of each motion pattern. Based on the central limit theorem, the feature vectors are approximately normally distributed. For the estimation of the hidden state q_t which corresponds to the motion pattern, a grid-based ML estimation is used. The initial probabilities π_i are assumed to be uniformly distributed. The transition probabilities a_{ij} are set manually in a way to reflect the fact that human motion is very variable, so there is no regular scheme if a pedestrian might stop, turn left or turn right after walking straight. The values used for the transition probabilities a_{ij} are thus nearly identical, with deviations from a uniform distribution in the range of 0.05 to 0.10. For example, these deviations model that after *Standing*, *Running* is less probable than *Walking*. By setting the transition probabilities a_{ij} in the way just described, we overcome the limitation that there exists no real world or synthethic training data which represents realistic transition probabilities between the motion patterns in our approach.

3.2 *Analysis of a Pedestrian Interaction Graph*

Event detection in a scene is performed based on a pedestrian interaction graph in which nodes represent pedestrians and edges represent motion interaction between pairs of pedestrians. The motion interaction is modelled as pairwise motion patterns which are analyzed using HMM. The graph is capable of robustly changing its topology because it is dynamically updated at every frame.

Edge Weight for Graph Simplification

The number of edges and the computational cost for the analysis of the graph is $N \cdot (N\text{-}1)/2$ for a number of N pedestrians. To overcome this high computational cost, we introduce edge weights based on a Gaussian function including the pedestrian density to significantly reduce the number of edges in the graph. Thus, only edges representing significant motion interactions between directly adjacent pedestrians are considered in the graph. The weight function $w_{ij}(d,\rho)$ with $0 \leq w_{ij} \leq 1$ between two nodes i and j is defined as

$$w_{ij} = \exp\left(-\frac{d_{ij}^{\ 2}}{\frac{1}{2\rho}} \right). \tag{10}$$

The weight w_{ij} depends on the distance d_{ij} between the pedestrians representing the nodes i and j and on the density ρ which is given by pedestrians per m^2. The weight function is a Gaussian function with height 1, μ=0 and $\sigma = 1/\left(\sqrt{\rho}\cdot 2\right)$. Thus, the weight w_{ij} receives high values for directely adjacent pedestrians i and j where a significant motion interaction is supposed to take place. At high pedestrian densities the weight w_{ij} is only high between pedestrians that have a distance of a few decimeters, whereas at low densities the weight w_{ij} can be high even if adjacent pedestrians have an offset of several meters. We introduce a threshold which is applied to the weights in order to determine which edges in the graph are to be constructed and which edges are omitted.

Framewise Graph Updating

Our graph-based approach for event detection in crowds represents dynamic behaviour of pedestrians. To this end, the pedestrian interaction graph is capable of changing its topology at every frame depending on the new arrangement of pedestrians in the scene. Fig. 2 shows an example of four pedestrians and their trajectories. Particular time steps are represented by dotted lines in between the trajectories. Figs. 2 a), 2 b) and 2 c) show the corresponding graph with four nodes. The topology changes by inserting a new edge between nodes 2 and 3 because of the decreased distance between the corresponding pedestrians. The density is supposed to be constant. The width of the edge connecting nodes 2 and 3 increases in Fig. 2 c) as a consequence of the increased weight w_{23}.

There are three preconditions of how our system deals with the sequential interaction analysis, depending on the configuration of the graph in the previous step. The first case is that an edge existing in the previous time step will further exist, such that the corresponding interaction analysis can be continued. The second case is that two pedestrians are converging and the weight w_{ij} exceeds the threshold. In this case, a new edge is generated and the analysis of this interaction is started. The third case is that two pedestrians diverge and the weight w_{ij} falls below the threshold. In this case, the corresponding edge is removed from the pedestrian interaction graph.

Robust Event Detection

The interaction analysis between a pair of pedestrians is performed by HMM. The forward algorithm is used to derive the type of motion pattern for each pair of pedestrians, for which a common edge in the graph exists. When applying the forward algorithm, the motion features serve as observations and the motion patterns serve as the hidden states of the HMM. Edges in the graph can arise or disappear during the sequence because of the dynamic behaviour of the crowd described in the previous section. Pedestrians do not move in a linear way but tend to slightly deviate to the left or right while walking. Therefore, the interaction analysis bears the risk of being interrupted for some frames if pedestrians depart from each other for a short time only. To overcome this risk and to achieve a robust sequential analysis of the motion interaction, an HMM buffer is used when analyzing the pedestrian interaction graph. The HMM buffer is internally activated for a specific interaction when the weight of the corresponding edge decreases below the threshold. At this point, the recursion of the forward algorithm would be terminated if no HMM buffer was used. However, the recursion is internally continued for a user-defined maximal number of frames in the HMM buffer. In the case that the two cor-

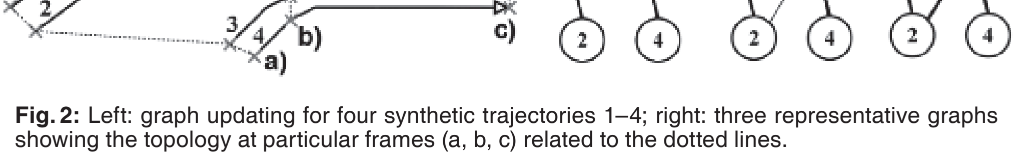


Fig. 2: Left: graph updating for four synthetic trajectories 1–4; right: three representative graphs showing the topology at particular frames (a, b, c) related to the dotted lines.

responding pedestrians approach each other and the weight increases again, the consistently and eternally analysed motion interaction is loaded from the HMM buffer and the result is subsequently added to the corresponding interaction analysis. Hence, the temporarily omitted corresponding edge of the graph is subsequently constructed. Thus, no fragments of the corresponding pedestrian interaction can arise. If the weight does not increase again after the defined number of frames, the corresponding interaction analysis is terminated and the edge is finally deleted.

4 Experimental Results

We present experimental results for robust pedestrian interaction analysis using two datasets with different camera platforms and different scenes. The datasets are described in the next section. Afterwards the results show significant scenes and demonstrate the robustness of our approach.

4.1 *Datasets*

The first dataset used contains image sequences captured from a UAV octocopter platform. The images were taken from a flying height of 85 m with a Panasonic DMC-LX3 camera, resulting in a ground resolution of 1.5 cm. The frame rate of the image sequences is 1 Hz. The captured scenes contain more than 10 different scenarios of pedestrian group behaviour in different complexity levels. The pedestrians were instructed about the scenarios in advance; however, the information was reduced to a minimum in order to preserve natural behaviour. The captured scenes contain simple scenarios such as commonly walking pedestrians but also complex scenes like a bottleneck, crossing pedestrian groups at different velocities or an escaping situation. The second dataset is an image sequence taken by an airborne camera platform of the German Aerospace Center (DLR). The image sequence contains 16 frames taken at a frame rate of 2 Hz. The ground resolution is 0.15 m at a flying height of 1000 m. The area of interest is a soccer stadium where thousands of people are approaching the gates.

For the experimental results we use pedestrian trajectories which were generated manually from the image sequences because we focus on realistic trajectories and the potential of our graph-based event detection system for realistic pedestrian behaviour. However, our event detection system is able to deal with possibly incomplete automatically generated tracklets because the pedestrian interaction graph can deal with changing topology in a straightforward way.

4.2 *Event Detection Results*

The event detection results for the UAV dataset are shown in Figs. 3 and 4, including a colourbar which depicts the colour labels for the edges corresponding to the motion patterns. In Fig. 3, a group of pedestrians passes a narrow bottleneck (frames 3, 6, 9 and 12 are shown). Our event detection system successfully detects the typical motion interaction characteristics. Neighbouring pedestrians *Converge* and *Walk* towards the bottleneck, which is illustrated by orange and blue edges in frames 3 and 6. Pedestrians who have passed the gap *Diverge* and *Walk* ahead, whereas the pedestrians at the back of the group have to *Queue* for a while in frame 9. In Fig. 4, a corridor scenario of two walking groups walking in opposite directions is successfully detected. This scenario is characterized by two approaching and internally *Converging* groups in which the backmost pedestrians again have to *Queue* because of the narrowness of the corridor. The formation of lanes, which has already been investigated in Helbing et al. (2001), can be confirmed by the linearly arranged motion pattern *Walking* and the oppositely arranged motion patterns *Converging* and *Diverging* (frames 8 and 11).

The results for the soccer stadium sequence are presented in Figs. 5 and 6. Our event detection system is robust in the case that interacting pedestrians depart from each other for a short time by applying the HMM buffer. This robustness is exemplified in Fig. 5: the top row contains three consecutive frames of a pedestrian interaction graph where edges between node 7 and nodes 5 and 6 are not present in the middle frame. This graph was produced without the HMM buffer. The bottom row shows the graph which was produced with the HMM buffer. Here, the corresponding edges exist such that a continuous and robust analysis of the interaction between the pedestrians 5, 6 and 7 can be performed. In Fig. 6, *Queueing* pedestrians are successfully detected, which is displayed by the yellow edges during the whole sequence. Some pedestrians are passing the queue and perform multiple interactions due to freedom of motion. During the sequence, the density in the narrow area between

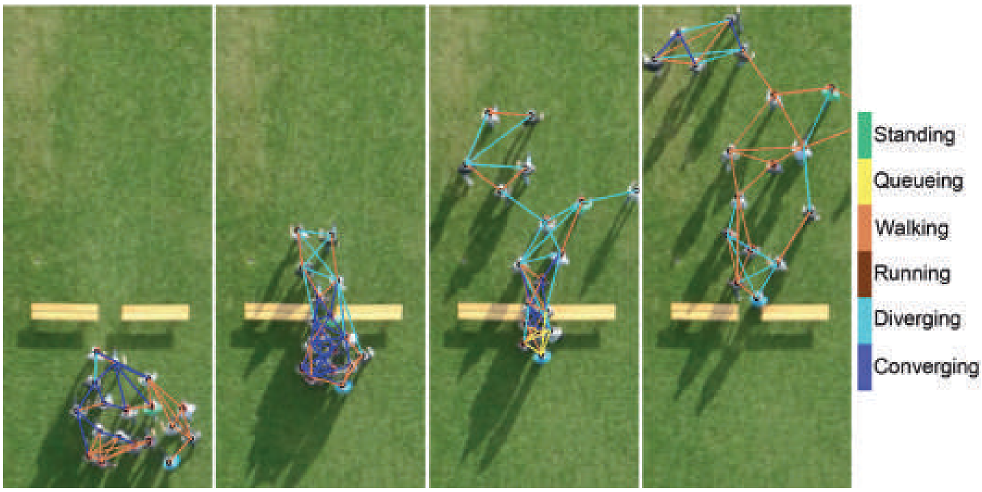


Fig. 3: Event detection result (frames 3, 6, 9 and 12) for the UAV dataset showing a bottleneck scenario in the upwards walking pedestrian group.

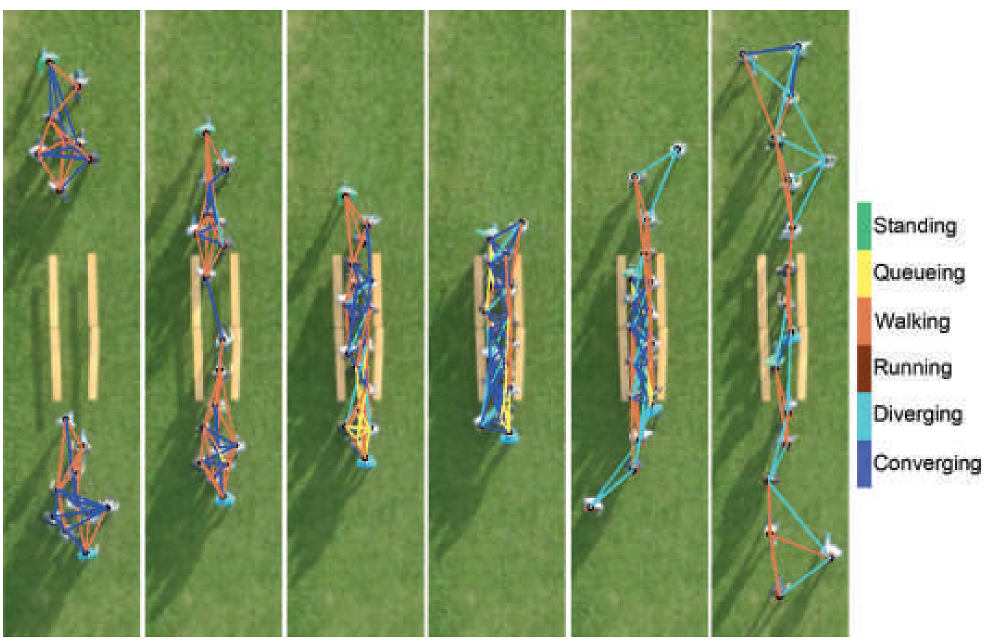


Fig. 4: Event detection result (frames 2, 5, 8, 11 and 14) for the UAV dataset showing a corridor scenario between two antiparallel walking pedestrian groups.

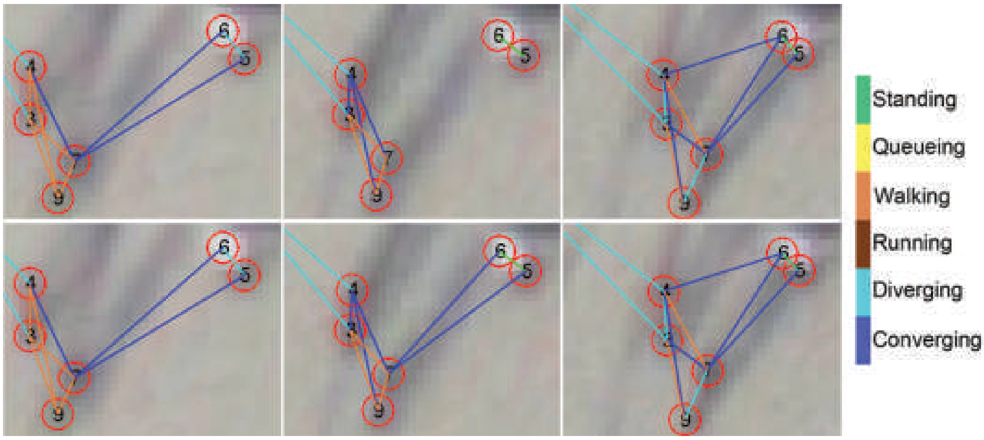


Fig. 5: Top: pedestrian interaction graph without HMM buffer; bottom: graph with HMM buffer.

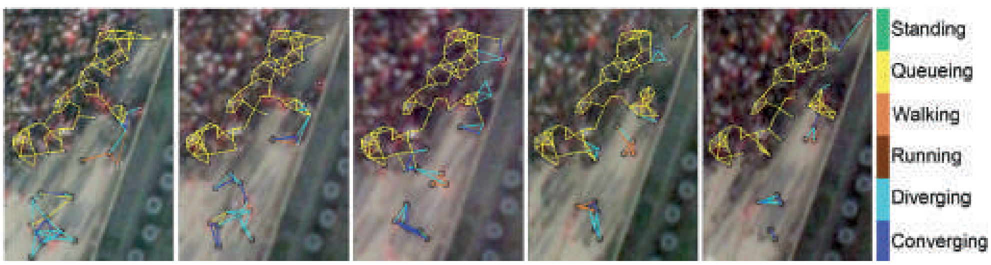


Fig. 6: Event detection result (frames 2, 6, 10, 14 and 16) for the soccer stadium dataset showing pedestrians passing a queue in a narrow area by a wall.

the queue and the wall on the right rises and the velocity decreases, which is illustrated by more and more *Queueing* patterns in this area.

The results demonstrate the potential of our system for graph based event detection by analyzing motion interaction in groups of pedestrians. Using HMM, the sequential behaviour of motion interaction between two pedestrians can reliably be analyzed. The six motion patterns we have defined represent human motion behaviour in a simple manner, such that areas of homogeneous behaviour as well as specific behaviour of only a few pedestrians can be inferred. In some frames outliers of the predominant motion pattern for a continuous pedestrian interaction arise. This is caused by the individual freedom of motion of pedestrians. Single and momentary variations of motion interaction in crowded scenes have to be expected. The low frame rate of the UAV dataset is sufficient to derive motion features which are used for event detection.

5 Conclusions

In this paper, we present a new approach for the analysis of pedestrian interaction and events in crowded scenes. We construct a pedestrian interaction graph in which nodes represent all pedestrians in a scene and edges represent motion interaction between neighbouring pedestrians. The graph can change its topology during the sequence and is robust against fluctuating and partly departing pedestrians. A set of six universal motion patterns is defined, describing the type of interaction which is then detected by HMM. We extend the motion features of previous work by the scalar product of motion directions and a refined density calculation, serving as observations for the HMM. We use a new UAV dataset for the evaluation of our event detection system, as well as an aerial dataset of a soccer stadium. The promising results show the potential of our approach to interpret pedestrian motion behaviour. Future work aims at a higher level analysis of the graph for an automatic and probabilistic detection of complex events in pedestrian groups.

Acknowledgements

We thank the German Aerospace Center (DLR) for providing the image sequence and for funding this work and also the International Graduate School of Science and Engineering (IGSSE) of Technische Universität München for supporting this study.

References

Adam, A., Rivlin, E., Shimshoni, I. & Reinitz, D., 2008: Robust Real-Time Unusual Event Detection using Multiple Fixed-Location Monitors. – IEEE Transactions on Pattern Analysis and Machine Intelligence **30** (3): 555–560.

Andrade, E.L., Blunsden, S. & Fisher, R.B., 2006: Hidden Markov Models for Optical Flow Analysis in Crowds. – **18th** International Conference of Pattern Recognition, Hong Kong, 460–463.

Basharat, A., Gritai, A. & Shah, M., 2008: Learning Object Motion Patterns for Anomaly Detection and Improved Object Detection. – IEEE Conference on Computer Vision and Pattern Recognition: 1–8, Anchorage, USA.

Burkert, F. & Butenuth, M., 2011: Event Detection based on a Pedestrian Interaction Graph using Hidden Markov Models. – Stilla, U., Rottensteiner, F., Mayer, H., Jutzi, B. & Butenuth, M. (eds.): Photogrammetric Image Analysis. – Lecture Notes in Computer Science **6952:** 271–283, Springer, Berlin, Germany.

Butenuth, M., Burkert, F., Borrmann, A., Kneidl, A., Hinz, S., Schmidt, F., Sirmacek, B. & Hartmann, D., 2011: Integrating Pedestrian Simulation, Tracking and Event Detection for Crowd Analysis. – International Conference on Computer Vision Workshops, MSVLC, Barcelona, Spain.

Calderara, S. & Cucchiara, R., 2010: People Trajectory Mining with Statistical Pattern Recognition – **23rd** IEEE Conference on Computer Vision and Pattern Recognition Workshops, San Francisco, USA.

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

Helbing, D., Molnár, P., Farkas, I.J. & Bolay, K., 2001: Self-Organizing Pedestrian Movement. – Environment and Planning B: Planning and Design **26:** 361–383.

Hu, W., Xiao, X., Fu, Z., Xie, D., Tan, T. & Maybank, S., 2006: A system for learning statistical motion patterns. – IEEE Transactions on Pattern Analysis and Machine Intelligence **28** (9): 1450–1464.

Kuettel, D., Breitenstein, M.D., Van Gool, L. & Ferrari, V., 2010: Learning Object Motion Patterns for Anomaly Detection and Improved Object Detection. – **23rd** IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, USA.

Mehran, R., Oyama, A. & Shah, M., 2009: Abnormal Crowd Behaviour using Social Force Model. – **22nd** IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, USA.

Nascimento, J.C., Figueiredo, M.A.T. & Marques, J.S., 2010: Trajectory Classification using Switched Dynamical Hidden Markov Models. – IEEE Transactions on Image Processing **19** (5): 1338–1348.

Oliver, N.M., Rosario, B. & Pentland, A.P., 2000: A Bayesian Computer Vision System for Modeling Human Interactions. – IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (8): 831–843.

Porikli, F. & Haga, T., 2004: Event Detection by Eigenvector Decomposition using Object and Frame Features. – **17th** IEEE Conference on Computer Vision and Pattern Recognition Workshops, Washington, DC, USA.

Rabiner, L.R., 1989: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. – IEEE **77** (2): 257–286.

Steffen, B. & Seyfried, A., 2010: Methods for measuring pedestrian density, flow, speed and direction with minimal scatter. – Physica A: Statistical Mechanics and its Applications **389** (9): 1902–1910.

Zhan, B., Monekesso, D.N., Remagnino, P., Velastin, S.A. & Xu, L., 2008: Crowd Analysis: A Survey. – Machine Vision and Applications **19** (5–6): 345–357.

Address of the Authors:

Dipl.-Ing. Florian Burkert and Prof. Dr. Richard Bamler, Technische Universität München, Lehrstuhl für Methodik der Fernerkundung, D-80333 München, Tel.: +49-89-289-22643, Fax: +49-89-289-23202, e-mail: florian.burkert@bv.tum.de, richard.bamler@dlr.de

Manuskript eingereicht: März 2012
Angenommen: August 2012

PFG 2012 / 6, 0711 – 0726
Stuttgart, December 2012

Article

Lithological Mapping of Dahab Basin, South Sinai, Egypt, using ASTER Data

Adel Omran, Tübingen, Michael Hahn, Stuttgart, Volker Hochschild, Tübingen, Ahmed El-Rayes & Mohamed Geriesh, Ismailia, Egypt

Keywords: lithological mapping, band ratio stacking, classification, ASTER, Dahab basin, Sinai, accuracy assessment

Summary: The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) has gained importance for lithological mapping over the last decade. Its suitability for creating a lithological map of the Dahab basin in south eastern Sinai in Egypt is studied and presented in this paper. For classification and discrimination of different rock types spectral features, in particular band ratios, are used. A new band ratio stacking with a false colour composite created by the ratio stack image (8/5, 4/8, 11/14) is proposed for differentiation between younger granitoids and older granitoids in the southern and central part of the study area. Band ratios 7/6 and 6/4 have turned out to be very suitable for discriminating between cambrian rocks and upper cretaceous rocks in the northern part of study area. Field investigations at different locations of the study area have been carried out to aid in the interpretation and analysis of the ratio stack images. Together with available geological maps the ground truth data is taken into account for selecting training areas and for creating a new map using maximum likelihood classification. An accuracy assessment of the classification result with respect to the regional geological map created by EGSMA (1994) and the local area maps of Hassen et al. (2007) and El Masry et al. (2003) indicates overall accuracies between 83% and 94%. An achievement of this study is a lithological map which extends the EGSMA map by adding some rock units such as ring-dykes at Wadi Ferani and metasediment, acidic metavolcanics and basic metavolcanics in Wadi Saal and Wadi Ramthy.

Zusammenfassung: *Lithologische Kartierung des Dahab Beckens im Süd-Sinai (Ägypten) mit ASTER-Daten*. Mit dem Sensorsystem ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) werden seit Februar 2000 Daten aufgezeichnet, die einen Beitrag zur detaillierten lithologischen Kartierung leisten können. In diesem Artikel wird die Eignung von ASTER für die Herstellung einer lithologischen Karte des Dahab Beckens im Südosten des Sinai in Ägypten untersucht. Zur Klassifizierung und Diskriminierung verschiedener Gesteinsarten werden aus den 14 ASTER-Bändern ausgewählte spektrale Merkmale (Indexbilder) verwendet, wofür einzelne Bänder zueinander ins Verhältnis gesetzt werden. Die Indexbilder der Bandverhältnisse 6/7 und 4/6 haben sich als sehr geeignet für die Unterscheidung zwischen kambrischen Felsen und oberen Kreidefelsen im nördlichen Teil des Untersuchungsgebietes erwiesen. Mit den zu einem Falschfarbenbild zusammengesetzten Bandverhältnissen (8/5, 4/8, 11/14) wird eine bislang zur lithologischen Kartierung nicht verwendete Kombination vorgeschlagen, die sich zur Differenzierung zwischen den jüngeren und älteren Granitoiden im südlichen und zentralen Teil des Untersuchungsgebietes besonders eignet. Feldbegehungen des Untersuchungsgebietes dienten sowohl dem besseren Verständnis der zu interpretierenden Bilder als auch der stichprobenhaften Erfassung von Ground-Truth-Daten für die Klassifizierung. In einem für die im Dahab Becken vorkommenden Gesteinsarten optimierten Prozess wird eine neue lithologische Karte durch Maximum-Likelihood-Klassifizierung der Indexbilder und eine entsprechende Nachbearbeitung erstellt. Vorhandene geologische Karten (EGSMA 1994, Hassen et al. 2007, El Masry et al. 2003) wurden für die Auswahl von Trainingsgebieten, insbesondere auch zur Validierung der erstellten Karte herangezogen. Die Beurteilung der neuen Karte erfolgte anhand von Konfusionsmatrizen, die auf

 www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0151 1432-8364/12/0151 $ 4.00

eine Übereinstimmung mit den vorhandenen Karten zwischen 83% und 94% hindeuten. Die neue Karte erweitert zudem die amtliche geologische EGSMA-Karte um Ring-Dykes im Wadi Ferani und um Metasedimente sowie um saure und basische Metavulkanite in den Wadis Saal und Ramthy.

1 Introduction

Although Remote Sensing techniques have opened new ways for mapping lithology over the past three decades, the geological maps of the Sinai Peninsula were created based on conventional ground surveys with suitable field observations. A standard procedure was to probe along traverse lines at regular intervals and to plot this point information on a topographic map. The final geological maps then have been created by interpolating the point information and cartographic post processing. By interpolating the sparse point data in a mapping area certain errors are unavoidable and lead to inaccuracies in the map. With remote sensing data the mapping procedures have undergone significant changes. The availability of high resolution multispectral and hyperspectral data has further increased the potential of remote sensing in delineating the lithological contacts and geological structures into greater detail and with better accuracy (Drury 1987).

A main purpose of the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) mission is to extend the understanding of local and regional phenomena on the earth surface and its atmosphere. Goals of geologic research using ASTER are summarized by Gomez et al. (2005). They put the focus on the "study the geologic phenomena of tectonic surfaces and geologic history through detailed mapping of the Earth topography and geological formation". For the discussion of the importance of the different ASTER wave bands the before mentioned authors refer to early work of Knipling (1970), Hunt (1977, 1979), and Salisbury et al. (1987): "The visible and near infrared (VNIR) wavelength region provides some information for the presence of vegetation, iron oxides (hematite, goethite, jarosite) and rare earth elements. The shortwave infrared (SWIR) wavelength region assesses bearing minerals (clays, phyllosilicates, …) and the thermal infrared (TIR) wavelength region permits to distinguish silicates and carbonates".

The study area (Dahab basin, Sinai) is located within an arid climatic belt that crosses northern Africa. Rocks dominate the appearance of the landscape which is sparsely covered by desert vegetation. The goal of this study is to work out details of a classification approach for improving the existing geologic map in the Dahab basin. It further intends to point a way forward for updating the geological maps of the southern Sinai Peninsula as well as for the geologically similar regions in Egypt's eastern desert and in the western part of Saudi Arabia. A second reason for generating the geological map in this study is to provide an updated basis for hydrogeological investigations of the Dahab basin.

Related work

ASTER data have been successfully used in geological mapping since early 2000. In comparison to Landsat TM data, ASTER data has the advantage of combining wide spectral coverage and high spatial resolution in the visible and infrared regions which makes it attractive for geological mapping (e.g. Hewson et al. 2001, Bedell 2001). Yamaguchi & Naito (2003) studied spectral index images for lithological mapping. Index images are found by a linear transformation of reflectance values of the five ASTER SWIR bands. The idea of this transformation is to direct the transformation axes to the spectral pattern of the target minerals. The calculated indices are named according to the minerals: alunite, kaolinite, calcite and montmorillonite. An advantage of this approach is that the transformation coefficients are not scene dependent. A simulated ASTER dataset was used to prove the usefulness of the spectral index images.

Some studies used band ratios and the spectral unmixing technique applied to SWIR, TIR and emissivity data of the TIR bands to improve the mapping of the sedimentary, metasedimentary and volcanic areas. For example, CUDAHY & HEWSON (2002) distinguished some minerals in epithermal, porphyry and skarn groups by using the band ratio technique. ABDEEN et al. (2001) used ASTER band ratios (4/7, 4/1, 2/3*4/3) and (4/7, 3/4, 2/1) for mapping ophiolites, metasediment, volcaniclastics, and granitoid lithologic units of the Allaqi suture in the southeastern part of eastern desert of Egypt.

GAD & KUSKY (2007) recommended to use band ratios (4/7, 4/6, 4/10) for mapping the granite and metamorphic belt of the Wadi Kid area of Sinai and for mapping metamorphic rocks in the Arabian Nubian Shield and other arid areas. REDA et al. (2010) used band ratios (2+4)/3, (5+7)/6, (7+9)/8 to discriminate between different ophiolitic and granitic rocks in the central eastern desert of Egypt. MADANI & EMAM (2009) investigated the band ratio composite (8/5, 5/4, 7/8) to differentiate between alkali granites (younger granitoids), granodiorites and quartz diorite (old granitoids) in the Wadi el-Hudi area, which is located in southeastern desert of Egypt. QARI et al. (2008) utilized the (6/8, 4/8, and 11/14) band ratio image to discriminate the basement rocks in the Arafat area of the Western Arabian Shield, Saudi Arabia, and created a 1:100,000 geological map. Inspired by the work of MADANI & EMAM (2009) as well as QARI et al. (2008) a new band ratio combination (8/5, 4/8, 11/14) is proposed in this study. Details for this new band ratio stacking are given in sections 4 and 5.

WALDHOFF et al. (2008) used Hyperion data and hyperspectral analysis techniques like the minimum noise fraction transformation (MNF), the pixel purity index technique (PPI) and the spectral angle mapper classification (SAM) for geological mapping and compared it with ASTER data classification results. They concluded that the ASTER data allowed a more detailed classification of the surface composition of the study area. This is particularly noteworthy because Hyperion excels by its 220 contiguous spectral bands which cover the spectral range from 400 nm to 2.5 µm at a ground resolution of 30 m. ROWAN & MARS (2003) used in situ measurements of spectral reflectance curves for calibrating the VNIR bands of ASTER. Lithological mapping was carried out by selecting and introducing image spectra of various lithological groups into classification produces. HEWSON et al. (2001) studied a regolith and alteration area in Australia and showed how to improve an existing 1:100,000 geological map with ASTER data. For their investigations they relied on previous surveys using airborne HyMap recordings of visible, shortwave and thermal IR bands and spectral measurements collected in field campaigns. They concluded that their experimental results indicate "that ASTER could discriminate mineral groups not achievable from Landsat TM, though more precise mineral species mapping is not possible".

In the next section background information about the location of the study area and the available ASTER data is given. The geology of the study area is outlined in section 3. In the following two sections 4 and 5 the approach pursued in this research for creating a lithological map of the Dahab basin is presented and processing of the ASTER data using band ratios and supervised classification is discussed in the context of the achieved results. Section 6 summarizes an accuracy assessment of the generated map and presents conclusions of the achievements.

2 Location of the Study Area and ASTER Data

The Dahab basin is considered as one of the major hydrographic basins along the Gulf of Aqaba. The basin is located in the southeastern part of the Sinai Peninsula and is bounded by Latitude 28° 22' 43.4" and 28° 52' 18.5" N and Longitude 33° 55' 46.9" and 34° 31' 28.8" E. It occupies an area of about 2080 km^2 (Fig. 1, left). It is bordered by the Gulf of Aqaba to the east, Gebel Gidid, Gebel Sheikh El-Arab and Gebel Ferani to the south, Gebel Um Alawi, Gebel Um Loz and Gebel Hamami to the west, and Gebel Bradi and Gebel Gunah to the north (Fig. 1, right). It includes some main wadis in South Sinai as Wadi Nasab, Wadi Ramthy, Wadi Saal, Wadi Zaghraa, and Wadi El Ghaieb.

ASTER data are offered at various processing levels. Level-1A data are reconstructed, unprocessed instrument data at full resolution which consist of the image data, radiometric

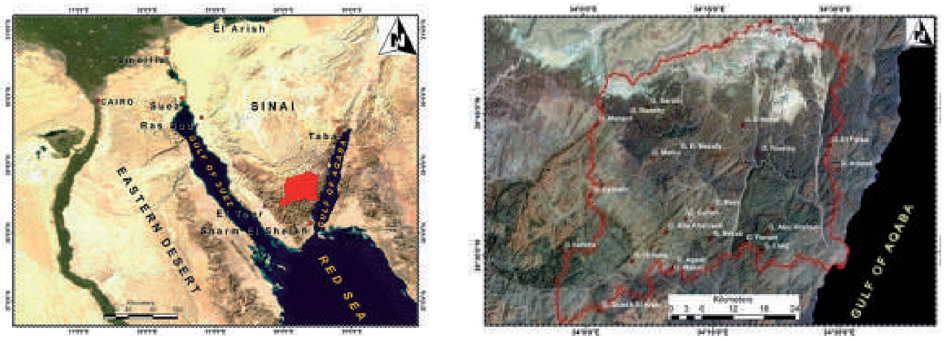


Fig. 1: Left: Location map of Wadi Dahab; right: Basin catchment area.

and geometric coefficients and other auxiliary data without applying the coefficients to the image data to maintain the original data values. The Level-1B data are generated applying these coefficients for radiometric calibration and geometric resampling. The scene used in this research is AST3A1 – 15 bands – 2006 which is a Level-3A data product. This so-called Terrain Correction Image includes Level-1B image data projected in UTM zone 36 (WGS 84) which are orthorectified using a DEM. The scene covers an area of 60 by 60 km^2 which apart from a small area in southeast encloses the Dahab basin completely.

The proposed procedure for lithological mapping of the study area using ASTER data (sections 4 and 5) mainly relies on the ASTER bands 4, 5, 6, 7, 8, 11 and 14. The spectral ranges of the 30 m resolution SWIR bands are: 1.600–1.700 µm (band 4), 2.145–2.185 µm (band 5), 2.185–2.225 µm (band 6), 2.235–2.285 µm (band 7), and 2.295–2.365 µm (band 8). The TIR bands 11 and 14 record emitted radiation in the wavelength ranges 8.475–8.825 µm and 10.95–11.65 µm, respectively. The spatial resolution of the TIR bands is at 90 m ground pixel size.

3 Geology of the Dahab Basin

The geology of the Dahab basin is discussed by many authors such as Hume (1906), Said (1962), Soliman (1986), El Shafei et al. (1992), Kora & Genedi (1995), Zalata et al. (1997), El Masry et al. (2003), and Hassen et al. (2007). The study area includes many rock types which can be subdivided according to their age into basement rocks and phanerozoic rocks. In this section the published knowledge on the geology of the basin is summarized with respect to the occurring rock types and its forming minerals.

The reflectance spectra of minerals are well known and catalogued, e.g. in the USGS Digital Spectral Library (Clark et al. 2007). The fact that rocks are a complex mixture of materials limits the direct utilization of those spectra for remote sensing. The use of the spectra is further lowered by the fairly broad bandwidth and the small number of spectral bands of ASTER. The challenge for the remote sensing approach is to analyse the reflectance of the mineral mix recorded by the ASTER bands. For example, the spectral characteristics of different rocks in the thermal infrared shows a direct dependency on the silica (quartz) contents (Kobayashi et al. 2010), so that the ASTER TIR bands will reveal quartz-related information about the rocks. How to reveal the rock specific information and how to exploit it within a classification scheme for lithological mapping will be discussed in sections 4 and 5.

3.1 *Basement rocks*

The EGSMA map (Fig. 2) shows that basement rocks are the prevalent rocks in the basin area, in particular in the southern and the central part of the basin. Basement rocks can be differentiated into igneous rocks, which cover more than 70% of the study area, and metamorphic rocks which mainly can be found at

Wadi Feirani, Wadi Saal, and some parts of Wadi Zaghraa (Fig. 2). The igneous rock comprises a variety of granitic rocks and younger gabbros. These rocks are intersected with acidic and basic dykes. El Masry et al. (2003) indicate the presence of ring-dyke at the Gebel Laig area with younger granitoid rocks.

Granitic rocks

Granitic rocks have wide areal extension at Wadi Nasab, Wadi Ramthy, Wadi Dahab and Wadi El Ghaieb. Hassen et al. (2004) and El Masry et al. (2003) divided the granitic rocks in the study area into late kinematic and past kinematic rocks which is substantially equivalent to older granitoid and younger granitoid rocks. The two groups of older and younger granitoids (Tab. 1) include the EGSMA differentiation of the granitic rocks into monzogranite, alkali granite, granodiorite and quartz diorite.

The mineral composition of younger granitic rocks is quartz and K-feldspar, plagioclase, hornblende and biotite, zircon, apatite, sercite (kaolinite) and opaque minerals. The older granitoids are mainly composed of plagioclase feldspar, K-feldspar, microperthite, hornblende, biotite, and quartz. Zircon, sphene, apatite and opaque minerals are among the accessory components. Weathering of hornblende, biotite and plagioclase leads to form clay mineral (El Masry et al. 2003).

Tab. 1: Granitic rock types in the Dahab basin area.

Types of granites	Occurring rocks	Forming minerals
older granitoid	granodiorite, quartz diorite, diorite	quartz, plagioclase, hornblende, pyroxene
younger granitoid	alkali granite, monzogranite	quartz, K-feldspar, plagioclase, biotite

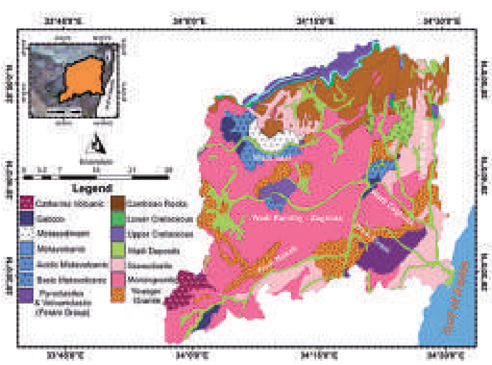


Fig. 2: Geological map of the Dahab basin (adopted from EGSMA 1994).

Metamorphic rocks

Metamorphic rocks in the study area are distributed along Wadi Saal, Wadi Ramthy and Wadi Zaghraa. Soliman (1986) and Hassen et al. (2004, 2007) mapped the metamorphic rocks at Wadi Saal and Wadi Zaghraa-Ramthy. They differentiated these rocks into metasediment, basic metavolcanic, acidic metavolcanic and metagabbro. The metamorphic belts are intruded by syn (older) and late (younger) granitoids and gabbroic rocks. Metasedimentary rocks consist mainly of phyllite, metasiltstone, metaconglomerate and volcanogenic sediment, whereas the metavolcanics include a wide variety of metamorphosed rock types such as andesite, dacite and rhyolite associated with minor basaltic bodies (Hassen et al. 2007).

3.2 *Phanerozoic Rocks*

Phanerozoic rocks cover the northern part of study area. Their outcrops expose mainly at the scarp of Gebel El Gounah and the northern Part of Wadi El Ghaieb, as well as minor exposures at Wadi Saal and Wadi Genah (Fig. 2). They are represented by a number of geological formations arranged from the oldest to the youngest as: 1) Cambrian rocks consist of laminated sandstone with intercalations of clay and ferruginous bands. They unconformable overlay the basement rocks, are coarse-to-medium-grained, weakly indurated to friable, and include kaolin matrix. 2) Lower cretaceous rocks consisting of grey and violet coloured pebbly and granular sandstone intercalated with kaolin in the upper part and impregnated with iron oxides at the top parts of the sequences. 3) Upper cretaceous rocks have yellow beds of fossiliferous sandstones, dolostones, limestones, marls, and glauconitic shales with pelecypod moulds, echinoids, and trace fossils of horizontal burrowings. The top of this formation constitutes a thick carbonate sequence of limestone, marl and dolomitic limestone, with thin inter-beds of silty claystone and yellowish-orange, fine-grained sandstone (Kora & Genedi 1995). Upper cretaceous rocks are mainly exposed at Wadi El Ghaib and Gebel El Gounah Scarp.

Mineralogically, the lower cretaceous rocks are mainly composed of quartz and kaolinite with a minor amount of calcite. This helps to distinguish them from upper cretaceous rocks in which the amount of calcite is high. A high degree of similarity exists in the mineral composition of cambrian rocks and the clastic part of lower cretaceous rocks especially with respect to quartz and kaolinite minerals. But the cambrian rocks are more ferriginated than lower cretaceous rocks, which allows differentiation between them.

4 Methodology

Lithological mapping may be carried out on the computer screen by human interpretation of the images. This is a promising way in particular if the human operator is very experienced. To increase the degree of automation within the mapping process the tools of image classification can be employed. The human operator is still a key factor for the mapping success as he will be involved in selecting proper training areas for supervised classification by taking advantage of existing maps or field visits. The digitized training regions are used to determine statistical parameters for classification. For the well-known maximum likelihood classification these are the mean vectors and covariance matrices for each training class. In maximum likelihood classification, all pixels are evaluated and assigned to the class of highest probability. Maximum likelihood classification of the entire study area in a one step process has been found to be not optimal. The mineral compositions the sedimentary rocks with sandstone of cambrian rocks and of lower cretaceous rocks are similar to the mineral compositions of granite rocks. Therefore, the analysis of the northern part of study area, which includes sedimentary rocks, is separated from the analysis of the southern part where the basement rocks are forming the main rock component.

The proposed overall process flow for creating a lithological map of the Dahab basin is shown in Fig. 3. The basic idea of this process flow is to use the prior knowledge of the existing EGSMA map to guide classification of the rocks. With this knowledge supervised classification will be specifically applied to the input image data for a certain area. According to the EGSMA map the phanerozoic rocks mainly cover the northern part and basement rocks dominate in the central and southern part of the Dahab basin. For the primary differentiation of phanerozoic rocks and basement rocks our process follows the proposal of Abdeen et al. (2001). Spatial separation between northern part and central part is done by manual digitization using ASTER band combination 7-3-1.

In each of the two regions band ratio images are used as input for supervised classification of different rock types. For this purpose some band ratios which have been successfully used by other researcher are used. In addition, a new ASTER band ratio stacking (8/5, 4/8, 11/14) is introduced. The training areas for the maximum likelihood classification are selected on the basis of existing geologic maps together with supporting field visits. As a part of this study, field visits at 23 locations of the study area have been undertaken. However, the idea to use the field mapped data as reference was rejected because of the small sample size. Therefore, the geologic maps are further used as references for evaluating the accuracy of the classification result.

Band ratioing has been widely used for lithological mapping due to its proven ability to produce distinct grey tones of imaged materials in certain ratios. A band ratio is created by dividing the digital number (DN) of one band by the corresponding DN of another band for each pixel (Drury 1987). The majority of fractional values are between zero and two or three. Thus, for visibility reasons the ratios are often rescaled to produce ratio images with higher contrast. Another well-known effect of ratioing is the reduction of the impact of shadow in the ratio images. Which band ratio is particularly suitable for enhancing a certain rock type or mineral depends on the dominance of the mineral in the reflected data. Spectral signatures give useful hints to decide about the bands used for ratioing. Combinations of three band ratio images can be visualized as colour composites. Features or minerals show up in distinct colours in these stacked ratio images. The question of which band ratios or band ratio stackings enhances the visibility of a particular rock type is analysed and discussed extensively in the next section.

The processing flow (Fig. 3) points out the different band ratio images used for maximum likelihood classification. In the northern part of the basin the mapped classes are cambrian rocks, upper cretaceous and lower cretaceous rocks, granodiorite and wadi deposits. In the central and southern part of the basin there are the metamorphic rocks with metavolcanic, basic metavolcanic, acidic metavolcanic, metasediment, phyllite and metagabbro. Wadi deposits are also taken into account in particular for comparison reasons with the reference map. For the igneous rocks as the other major group in the central and southern part of the study area the rock type classes granodiorite, alkali granite, monzogranite are mapped by image classification. Ring dykes are visually recognizable in the image. Supported by field visits they have been digitized interactively. In classifying this group of igneous rocks the metamorphic Feirani metavolcaniclastic rocks are added for comparison reasons with the reference map.

A post classification smoothing of the classification results is carried out by majority filtering and with the suppression of very small areas. The results of the lithological mapping process are the data found by vectorisation of the post processed classification maps.

The classification accuracy assessment is the last step in the overall processing flow (Fig. 3). For assessing the derived map quantitatively the error matrix method is used which compares the classification map against a reference. Ideally a representative sample of field mapped ground truth data are used as reference data. Due to the lack of suitable ground truth data the existing geologic maps are used. The EGSMA map is used as reference in the northern part of the basin. The more detailed map of Hassen et al. (2007) is used as reference in the areas of Wadi Saal and the Wadis Zaghraa and Ramthy. The map of El Masry (2003) is taken as reference in the areas of Wadi Nasab, Wadi El Ghaieb und Wadi Feirani. The limitations of this accuracy analysis are obvious; the error matrix provides information about how well the classification map and the existing maps coincide.

5 ASTER Data Analysis, Classification Results and Discussion

The usefulness of Landsat ETM band combination 7-4-2 for geological mapping in arid regions and the far-reaching consistency of this band combination to ASTER band combination 7-3-1 are pointed out by Abdeen et al. (2001). Fig. 4 left shows ASTER band combination 7-3-1 in which the metamorphic rocks appear in greenish and reddish colour, the granitic rocks as light yellow to light brown, the sedimentary rocks as white colour, and the wadi deposits as light grey colour. A manual mapping result of the wadi deposits is shown in Fig. 4 right. This vector layer of the wadi deposits will be used as overlay in other figures to simplify visual orientation.

5.1 *Central and Southern Part of the Dahab Basin: Igneous Rocks*

A new band ratio stacking with the ratio images (8/5, 4/8, 11/14) is used for differentiation between younger granitoids and older granitoids in the central and southern part of the study area. For a better understanding of this stacking the band ratio images are discussed in the context of the spectral signature of the dominant minerals.

Band ratio 4/8: Alkali granite appears dark, monzogranite grey, older granitoids show light grey to bright colour (Fig. 5A). The light colour of older granitoids is due to alteration products of hornblende and plagioclase into chlorite and clay minerals, whereas the presence of biotite and K-feldspar minerals in alkali granites (El Masry et al. 2003) produces a dark colour in band ratio 4/8. The dark col-

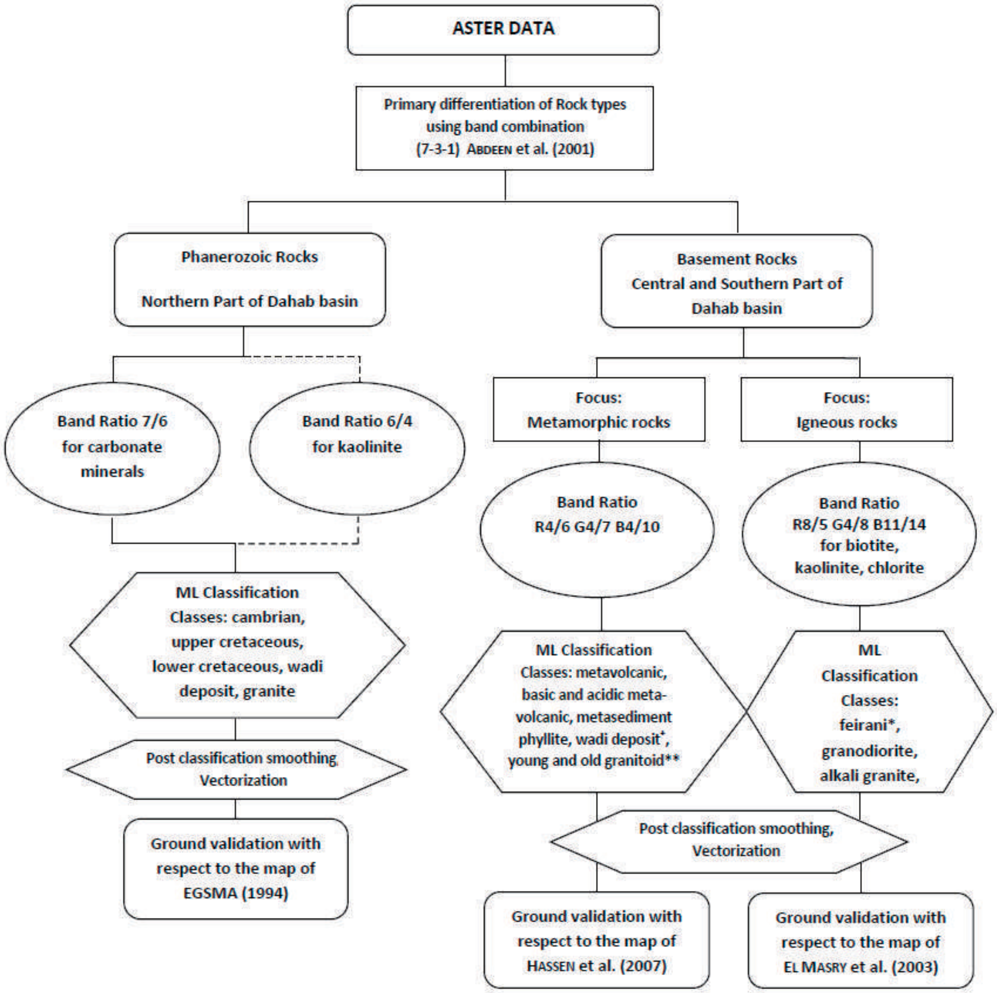


Fig. 3: Overall process flow for lithological mapping of the Dahab basin area (* is a metamorphic rock, ** is an igneous rock, + are not introduced into ML classification).

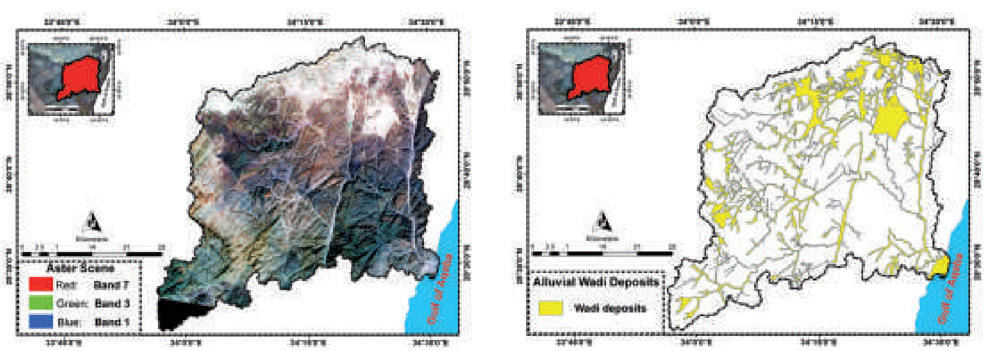
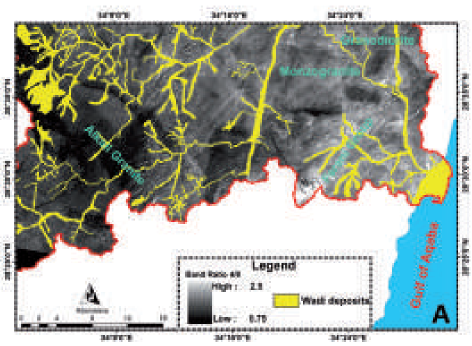


Fig. 4: Left: band combination 7-3-1 of Dahab basin area; right: map of the wadi deposits.

our is a consequence of the lower reflectance in band 4 (Figs. 5 B and C).

Band ratio 8/5: Younger granitoids appear as light grey and older granitoids as grey colour (Fig. 6A). In band ratio image 8/5, Feirani metavolcaniclastic rocks show dark grey colour. This is due to the presence of biotite and K-feldspar in addition to the alteration products of hornblende and plagioclase into clay minerals. Rocks rich in feldspar commonly weather to kaolinite. Fig. 6B shows an absorption feature of kaolinite near band 5 thus the high 8/5 band ratio values indicate younger granitoids. On the other hand, dark grey and grey colours of old granitoids and metamorphic rocks are due to the absorption property of chlorite (Fig. 5C), which leads to low 8/5 band ratio values.

Band ratio 11/14: Younger granitoids appear as dark colour, older granitoids show light grey to bright colour (Fig. 7). The dark colour of younger granitoids may be interpreted by their high ability to reflect the sun radiation on their light coloured surfaces, and hence they represent cooler surfaces which appear dark in the band image. In contrast, older granitoids



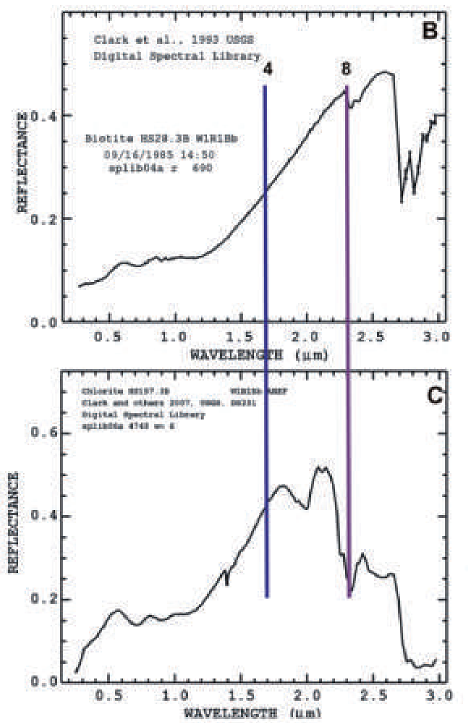


Fig. 5: A: band ratio 4/8 image; B: spectral signature of biotite; C: spectral signature of chlorite (Clark et al. 2007).

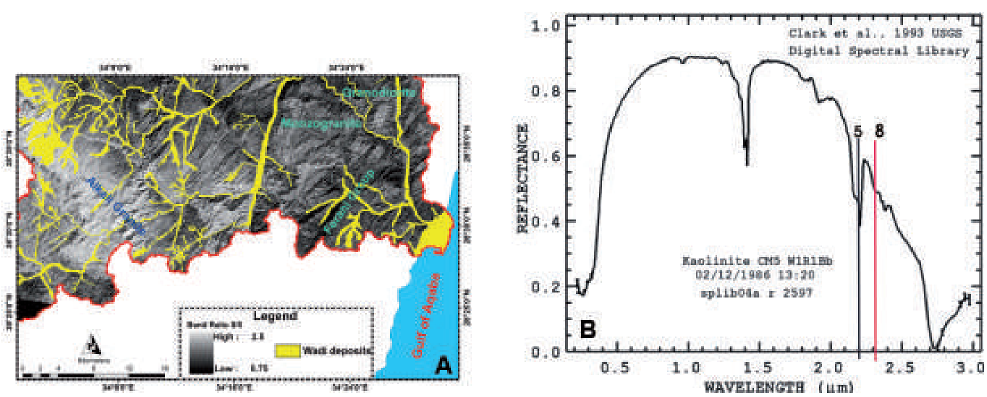


Fig. 6: A: band ratio 8/5 image; B: spectral signature of kaolinite (Clark et al. 2007).

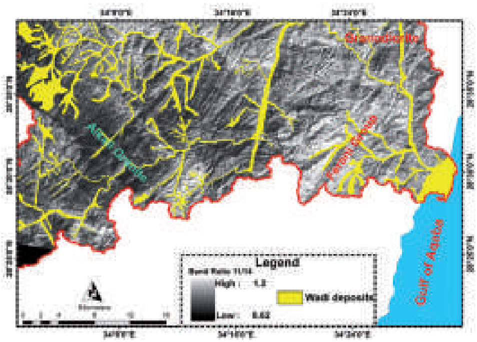


Fig. 7: Band ratio 11/14 image.

absorb more sun radiation and get warmer than other rock types and hence appear brighter in the band image.

In experiments with different band ratio stackings, we found that the colour composite (8/5, 4/8, 11/14) reveals subtle differences between the younger and older granitoids. Fig. 8A shows alkali granites that appear in red colour, monzogranites and acidic volcanic which show up in pink colour. The Feirani group appears as light green whereas older granitoid looks green and purple. The visual comparison of this new stacking with the ratio stack image (8/5, 5/4, 7/8) used by Medani & Emam (2009) in the El Hudi area of southeastern desert in Egypt shows its strength with respect to the discrimination of younger granitoids. Alkali granite and monzogranite rocks are much better discriminated in the new stacking. But with respect to older granitoids (granodiorite and Qz diorite) the Medani & Emam (2009) stacking seems to be a bit more favourable.

The colour composite of the single band ratios is used for defining the training areas of the four rock classes alkali granites, monzogranites, Ferani metavolcaniclastic, and old granitoids. For each class several (minimum 4) training areas are selected to get representative samples. In Fig. 10B the maximum likelihood classification result of the four rock classes is shown together with the wadi deposit layer. Alkali granites are colourized in red, monzogranites in pink, Ferani metavolcaniclastic in light green, old granitoids in dark green and wadi deposits in yellow.

For the accuracy investigation of the classification result the geologic map of El Masry et al. (2003) is used as reference. This local area map was generated based on detailed field work and covers the Gebel Feirani area. Fig. 2 shows the location of the Gebel Feirani area at the downstream part of the Dahab basin. In addition to alkali granites, monzogranites, and Feirani metavolcaniclastic, the El Masry map includes granodiorite, Qz-diorite, and tonalite. The latter two cannot be distinguished from granodiorite by image classification. Therefore, the prevalent granodiorite is introduced as a class on its own. Ring-dyke is interactively mapped because it is visually recognizable by its dyke-like body (Fig. 8). It has the same composition as alkali granites thus belongs to this class in the classification map.

The quantitative comparison between the results of the classification map and the reference map is summarized by the error matrix in Tab. 2. Reference data are selected from the reference map for the raster locations defined by the classification map. The procedure for

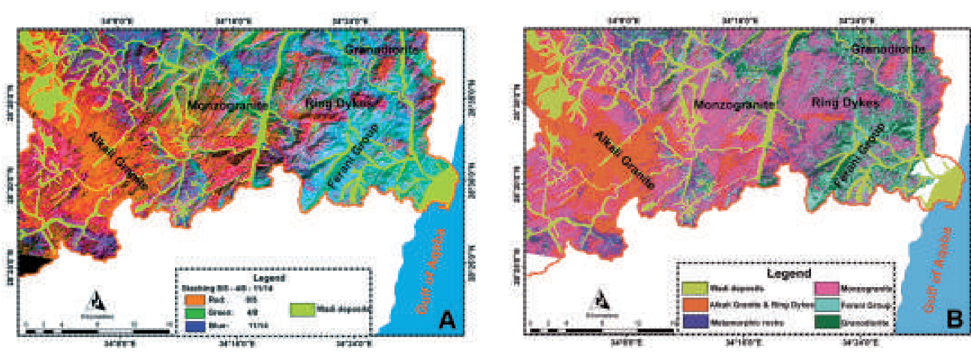


Fig. 8: A: stacked band ratio image (8/5, 4/8, 11/14); B: ML Classification using the stacked band ratio image (8/5, 4/8, 11/14).

Tab. 2: Accuracy evaluation of the classification map for igneous rock types.

		Class types from the reference map					
Class types of the classification map	Pixels	**Ferani**	**Grano**	**Alkali**	**Monzo**	***Total***	**User's Accuracy**
	Feirani	393	47	0	5	445	88%
	Granodiorite	81	392	8	3	484	81%
	Alkali Granite	0	5	775	18	798	97%
	Monzogranite	1	13	177	413	604	68%
	Total	475	457	960	439	2330	
Producer's Accuracy		83%	86%	81%	94%		**85%**

selecting samples in each category follows the stratified random sampling strategy. For each class in the classification map 4 to 5 samples are taken as reference data.

The results indicate an overall accuracy of 85%. Apart from the user's accuracy of 68% for monzogranite the user's and producer's accuracies of the other classes are all above 80%. The interfering contacts between monzogranite and alkali granite lead to a fairly high misclassification of monzogranite with the consequence of a low user's accuracy for this class. Similar is the situation for the Ferani metavolcanics and the granodiorite but with less significant consequences for the user's accuracy.

5.2 *Metamorphic Rocks*

The band ratio stacking (4/6, 4/7, 4/10) shows the metavolcanic (metatuffa) rocks in greenish yellow and the basic metavolcanic rocks in reddish brown colour. The acidic metavolcanic rocks show up by light violet colours, the metasediment phyllite rocks in light green, the younger granitoid rocks have dark blue colour and the light blue colour refers to old granitoid rocks (Figs. 9 A and B).

Training areas have been defined in the colour composite (4/6, 4/7, 4/10) by following the same procedure described already in section 5.1. The classification is carried out for seven classes of rock units. Fig. 10 shows the classification results for Wadi Saal (Fig. 10A) and Wadi Ramthy (Fig. 10B). Metavolcanics is colourized in blue, basic metavolcanics in purple, acidic metavolcanics in light green, metasediments (phyillite) in dark green, younger granitoid in red, older granitoid in pink and metasedimentary rocks (in Wadi Ramthy only) in dark green. Wadi deposits have been also introduced into classification and show up in yellow in the classification map.

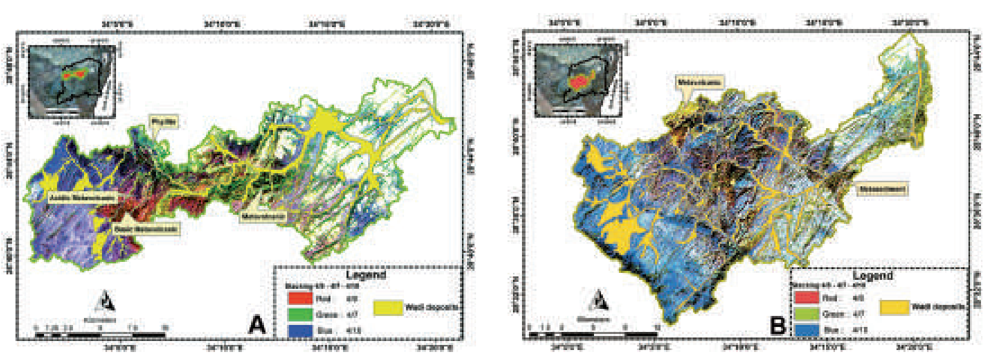


Fig. 9: Stacked band ratio (4/6, 4/7, 4/10) images; A: Wadi Saal area; B: Wadi Ramthy area.

For the accuracy investigation the classification maps are compared to the geologic map created by Hassen et al. (2007) based on field observations. This geologic map covers the areas of the Wadis Saal, Zaghraa and Ramthy. The error matrix found by stratified random sampling is listed in Tab. 3.

The overall accuracy of 83% confirms the good matching between the classification map and Hassen's reference map for both wadis. The error matrix (Tab. 3) has quite some similarity to the error matrix found for the classification of the igneous rocks (Tab. 2). The producer's accuracy is fairly high between 79% and 96% for all classes. The user's accuracy of 66% for metasedimentary rocks suffers from misclassification with metavolcanics and basic metavolcanics rocks. The small outcrops of acidic metavolcanics, metasediments, wadi deposits, and older granitoids produce a lot of uncertainty with respect to the other classes which results in user's accuracies between 66% and 76% for these classes.

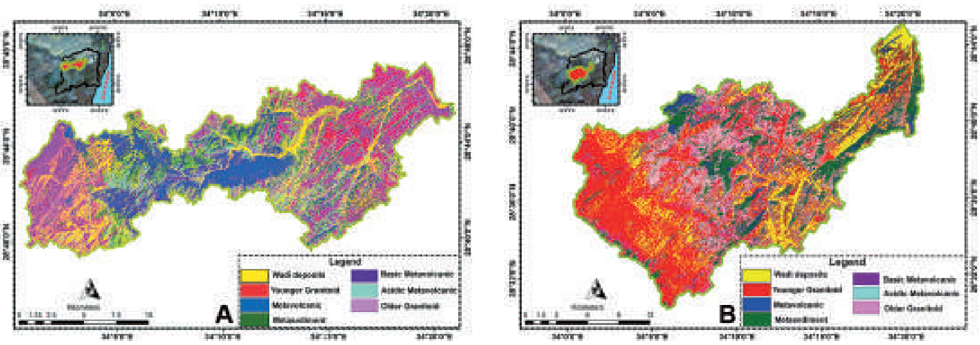


Fig. 10: ML classification of stacked band ratio (4/7, 4/6, 4/10) images; A: Wadi Saal; B: Wadi Ramthy.

Tab. 3: Accuracy evaluation of the classification map for metamorphic rock types.

		Class types from reference source								
	Pixels	**met_vol**	**met_ba**	**met_ac**	**met_sed**	**wd**	**young**	**old**	***Total***	**User's Accuracy**
Class types of the classification map	**Metavolcanic**	281	23	1	0	0	0	1	306	92%
	Basic metavolcanic	49	181	0	2	3	0	0	235	77%
	Acidic metavolcanic	3	11	56	0	1	3	0	74	76%
	Metasediment	8	11	1	41	1	0	0	62	66%
	Wadi deposits	5	2	1	3	28	1	0	40	70%
	Younger granitoid	0	0	4	0	0	73	0	77	95%
	Older granitoid	0	0	2	0	0	5	22	29	76%
	Total	346	228	65	46	33	82	23	823	
Producer's Accuracy		81%	79%	86%	89%	85%	89%	96%		83%

5.3 *Northern Part of the Dahab Basin: Sedimentary Rocks*

Band ratio 7/6 can be used to differentiate sedimentary rocks with respect to carbonate minerals, in particular calcite and aragonite which are the main components of limestone (upper cretaceous) rocks. The absorption features of calcite near band 7 (Fig. 11B) together with its high reflectance in band 7 produces the dark appearance of the cretaceous rocks in Fig. 11A. The cambrian rocks appear as bright colour, whereas the granites appear as darker grey tone.

Alternatively band ratio 6/4 can be used to discriminate between different classes of sedimentary rocks. Quartz with kaolin (altered to clay) together with iron oxide are the main minerals of cambrian rocks which lead to dark colours in this band ratio because of the absorption feature of kaolinite near band 6 and the high reflectance in band 4. Upper cretaceous rocks show up in grey and granites appear in bright tones in band ratio 6/4 (Fig. 12). A visual comparison of the 7/6 versus the 6/4 band ratio indicates that the discriminative efficiency of 7/6 ratio is higher than that of the 6/4 ratio. Therefore, the analysis focused in analysing band ratio 7/6. A joint use of both bands would have been possible but was not pursued in this analysis.

Supervised classification including post classification smoothing and vectorization is carried out by taking band ratio 7/6 as input image. In the classification map, upper cretaceous rocks is colourized in red and lower cretaceous rocks in light green. For granites violet and for cambrian rocks blue is used. Wadi deposits appear yellow in the classification map (Fig. 13).

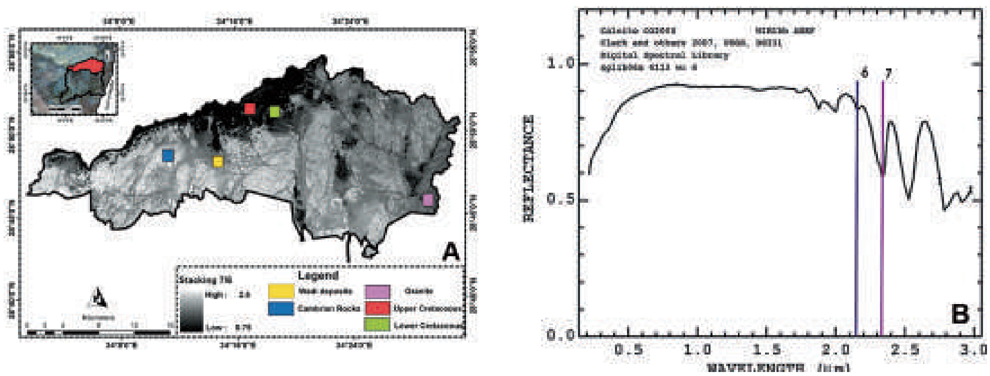


Fig. 11: A: band ratio 7/6 image of the northern part of Dahab basin area; B: spectral signature of calcite (Clark et al. 2007).

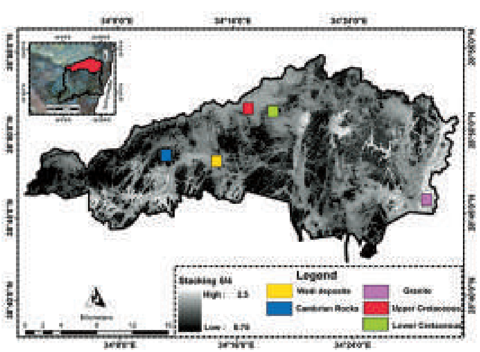


Fig. 12: Band ratio 6/4 image of the northern part of the Dahab basin.

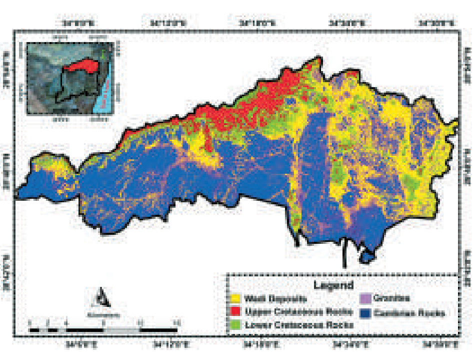


Fig. 13: Supervised classification of band ratio 7/6 image of the northern part of Dahab basin.

Tab. 4: Accuracy evaluation of the classification map for phanerozoic rock types.

		Class types from reference source						
	Pixels	Camb.	Up. Cret.	wd	Low. Cret.	Gran.	***Total***	**User's Accuracy**
Class types of the classification map	Cambrian rocks	191	0	0	0	23	214	89%
	Upper Cretaceous	0	97	0	6	0	103	94%
	Wadi deposits	0	0	994	2	13	1009	99%
	Lower Cretaceous	0	12	1	131	3	147	89%
	Granites	1	0	36	2	111	150	74%
	Total	192	109	1031	141	150	**1623**	
Producer's Accuracy		99%	89%	96%	93%	74%		**94%**

For the accuracy investigation the geological map provided by EGSMA (1994) is used as a reference. The overall accuracy is quite high (94%, Tab. 4). User's and producer's accuracies of 89% to 99% indicate a high agreement of the classification map and the reference. The only exception are the granitic rocks with a user's and a producer's accuracy of 74%. Granitic rocks and alluvial wadi deposits are in contact with each other which is probably the main reason for this lower accuracy.

The final geological map of the Dahab basin created according to the proposed process flow (Fig. 3) is shown in Fig. 14. It comprises 19 classes of phanerozoic, metamorphic and igneous rocks. Apart from the ring dykes all other classes have been created by supervised image classification followed by post classification smoothing and vectorization of the raster data.

6 Conclusion

The investigations on the suitability of ASTER data for lithological mapping have led to the development of an overall processing strategy for mapping granitic rocks, metamorphic rocks and sedimentary rock in Wadi Dahab basin. The goal to develop a classification based approach in which only minor interactive digitization is included is fully achieved. Interactive digitization was used to separate the northern and the central and southern part of the basin as well as for the mapping of the ring dykes. The study shows the suitability of maximum likelihood classification taking various band ratios and band ratio stackings into account. Through classification different types of granitoid rocks (monzogranites, alkali granites, granodiorites), metamorphic rocks (metasediments and metavolcanics) and phanerozoic rocks (cambrian, lower cretaceous, upper cretaceous rocks and loose wadi deposits) have been differentiated. For checking the quality of the classification map an accuracy assessment is carried out. For this purpose an error matrix is determined which compares the classification map with respect to existing geologic maps.

As a part of the development a new band ratio stacking (8/5R, 4/8G and 11/14B) is proposed for differentiation of younger granitoids (monzogranites and alkali granites) and older granitoids (granodiorite) in the Dahab area. Alkali granite and monzogranite rocks are very well discriminated by the new stacking. With respect to older granitoids (granodiorite, Qz diorite and tonalite) Medani & Emam (2009) stacking seems to be still more favourable. For the discrimination of different metamorphic rock types our procedure follows the proposal of Gad & Kusky (2007) by applying the band ratio stacking (4/6R, 4/7G and 4/10B). The phanerozoic rocks in the northern part of study are diffentiated using the (7/6) band ratio.

The accuracy investigations of the classification results are carried out with respect to the geological reference maps published by

Hassen et al. (2004), El Masry et al. (2003) and EGSMA (1994). The calculated error matrices indicate overall accuracies between 83% and 94%. Altogether this underlines that the created lithological map fits reasonably well to the existing maps which have been mainly created by field work. The created lithological map adds some rock units to the general geological map of EGSMA. Added are ring-dykes at Wadi El Ghaieb, acidic metavolcanics, basic metavolcanics and metasediments at the Wadis Saal, Zaghraa and Ramthy and metasediments at Wadi Ramthy. Alluvial wadi deposits are also included in the final lithological map as shown in Fig. 14.

References

Abdeen, M.M., Allison, T.K., Abdelsalam, M.G. & Stern, R.J., 2001: Application of ASTER band-ratio images for geological mapping in arid regions; the Neoproterozoic Allaqi Suture, Egypt. – Geological Society of America **3** (3): 289.

Bedell, R.L., 2001: Geological mapping with ASTER satellite: new global satellite data that is a significant leap in remote sensing geologic and alteration mapping. – Special Publication Geological Society of Nevada **33:** 329–334.

Clark, R.N., Swayze, G.A., Wise, R., Livo, K.E., Hoefen, T.M., Kokaly, R.F. & Sutley, S.J., 2007: USGS Digital Spectral Library splib06a. – U.S. Geological Survey, Data Series **231**.

Cudahy, T. & Hewson, R., 2002: ASTER Geological Case Histories: Porphyry-Skarn-Epithermal, Iron Oxide Cu-Au and Broken Hill Pb-Zn-Ag. – Communication in the Workshop "Mapping the Earth with ASTER", London.

Drury, S.A., 1987: Remote sensing of geologic structures in temperate agriculture terrains. – Geological Magazine **123:** 113–121.

EGSMA, 1994: Egyptian Geological Survey and Mining Authority. – Geologic map of Sinai, Arab Republic of Egypt. Sheet No. 1, scale 1:250,000.

El Masry, N.N., Hassen, I.S. & Hegazi, A.M., 2003: A Newly recognized example of a Late Precambrian subsurface cauldron subsidence intrusion in Southern Sinai: Jabal Laiq ring dyke. – **5th** International Conference on the Geology of the Middle East: 549–558.

El Shafei, M. & Kusky, T.M., 2003: Structural and tectonic evolution of the Neoproterozoic Feiran–Solaf metamorphic belt, Sinai Peninsula: implication for the closure of Mozambique Ocean. – Precambrian Research **123:** 269–293.

El Shafei, M.K., Khawasik, S.M. & El Ghawaby, M.A., 1992: Deformational styles in the tectonites of Wadi Sa`al area, South Sinai. – **3rd** Conference Geological Sinai Development: 1–8, Ismailia, Egypt.

Farina, P., Catani, F., Colombo, D., Fumagalli, A., Kukavicic, M., Marks, F. & Moretti, S., 2005: Remote sensing: a tool for landslide investiga-

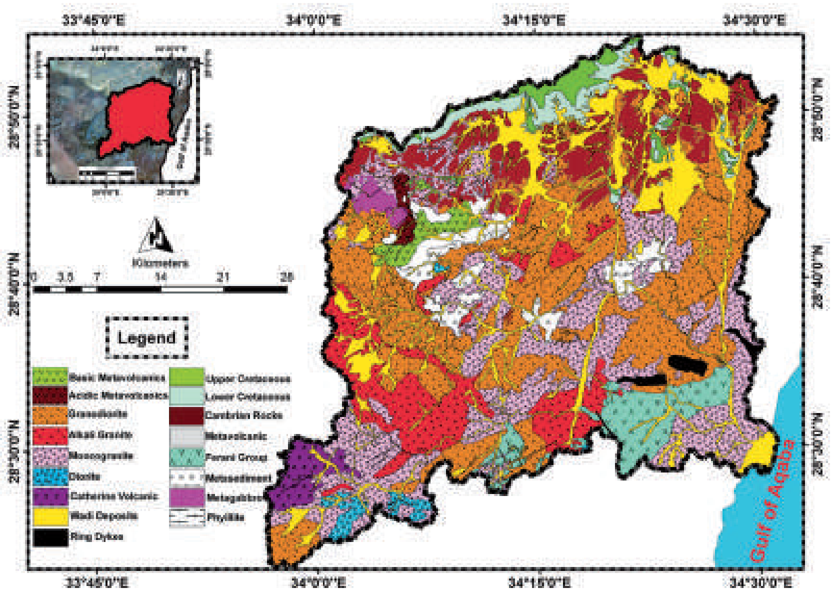


Fig. 14: Final lithological map of Dahab basin area.

tions at a basin scale. – Geophysical Research Abstracts **7:** 10157–10168.

Gad, S. & Kusky, T.M., 2007: ASTER spectral ratioing for lithological mapping in the Arabian-Nubian shield, the Neoproterozoic Wadi Kid area, Sinai, Egypt. – Gondwana Research **11** (3): 326–335.

Gomez, C., Delacourt, C., Allemand, P., Ledru, P. & Wackerle, R., 2005: Using ASTER remote sensing data set for geological mapping in Namibia. – Physics and Chemistry of the Earth **30:** 97–108.

Hassen, I.S., Ibrahim, S.K. & Elemer, P.M., 2004: Evolution and origin of the metavolcanics at Wadi Saâl area, south Sinai, Egypt. – Annals Geological Survey Egypt **27:** 61–78.

Hassen, I.S., El Shafei, M.K. & Stüwe, K., 2007: Late Proterozoic crustal evolution in the Arabian-Nubian Shield of Wadi Zaghra tectonites, South Sinai, Egypt. – Annals Geological Survey Egypt **V.XXIX:** 77–93.

Hewson, R.D., Cudahy, T.J. & Huntington, J.F., 2001: Geologic and alteration mapping at Mt Fitton, South Australia, using ASTER satellite-borne data. – International Geosciences and Remote Sensing Symposium **2:** 724–726.

Hume, W.F., 1906: The Topography and Geology of the Peninsula of Sinai (south-eastern portion). – 280 p., Egyptian Survey Department, Cairo.

Hunt, G.R., 1977: Spectral signatures of particulate minerals in the visible and near infrared. – Geophysics **42:** 501–513.

Hunt, G.R., 1979: Near infrared (1.3–2.4 µm) spectra of alteration minerals – Potential for use in remote sensing. – Geophysics **44:** 1974–1986.

Kobayashi, C., Orihashi, Y., Hiarata, D., Naranjo, J.A., Kobayashi, M. & Anma, R., 2010: Compositional variations revealed by ASTER image analysis of the Viedma Volcano, southern Andes Volcanic Zone. – Andean Geology **37** (2): 433–441.

Kora, M. & Genedi, A., 1995: Lithostratigraphy and facies development of Upper Cretaceous carbonates in East Central Sinai, Egypt. – Facies **32:** 223–236.

Knipling, E.B., 1970: Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. – Remote Sensing of Environment **1:** 155–159.

Madani, A. & Emam, A.A., 2009: SWIR ASTER band ratios for lithological mapping and mineral exploration: A case study from El Hudi area, southeastern desert, Egypt. – Arabian Journal of Geosciences **4:** 45–52.

Qari, M., Madani, A., Matsah, M. & Hamimi, Z., 2008: Utilization of Aster and Landsat data in geologic mapping of basement rocks of Arafat Area, Saudi Arabia. – Arabian Journal for Science and Engineering **33** (1C): 99–116.

Reda, A., Kusky, T. & Ghulam, A., 2010: Lithological mapping in the Central Eastern Desert of Egypt using ASTER data. – Journal of African Earth Science **56:** 75–82.

Rowan, L.C. & Mars, J.C., 2003: Lithologic mapping in the Mountain Pass, California area using Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data. – Remote Sensing of Environment **84:** 350–366.

Said, R., 1962: The Geology of Egypt. – 377 p., Elsevier, Amsterdam.

Salisbury, J.W., Walter, L.S. & Vergo, N., 1987: Mid-infrared (2.1–25 µm) Spectra of Minerals. – United States Geological Survey, Open File Report **1:** 87–263, Reston.

Soliman, F.A. 1986: Geology of Wadi Sa'al area with special emphasis of metamorphism and tectonics, Central Sinai of Egypt. – 240 p., Ph.D. thesis, Faculty of Science, Suez Canal University, Ismailia, Egypt.

Waldhoff, G., Bubenzer, O., Bolten, A., Koppe, W. & Bareth, G., 2008: Spectral analysis of ASTER, Hyperion and QuickBird data for geomorphological and geological research in Egypt (Dakhla Oasis, Western Desert). – XXI ISPRS Congress, WG **VIII/12:** 1201–1206, Beijing.

Yamaguchi, Y. & Naito, C., 2003: Spectral indices for lithologic discrimination and mapping by using the ASTER SWIR bands. – International Journal of Remote Sensing **24** (22): 4311–4323.

Zalata, A.A., El Metwally, A.A., El Aassy, I.E. & El Sayed, A.A., 1997: Evolution and Geochemistry of the Basement rocks of west Dahab area, Southeastern Sinai. – Third Conference on Geochemistry, Alexandria University: 1–15.

Addresses of the Authors:

Adel Omran and Prof. Dr. Volker Hochschild, Institute of Geoscience, Tübingen University, Rümelinstr. 19–23, D-72070 Tübingen, e-mail: adel-fouad-abdou.omran@student.uni-tuebingen.de, volker.hochschild@uni-tuebingen.de

Prof. Dr. Michael Hahn, Faculty of Geomatics, Computer Science and Mathematics, University of Applied Sciences Stuttgart, Schellingstraße 24, D-70174 Stuttgart, e-mail: michael.hahn@hft-stuttgart.de

Associ. Prof. Dr. Ahmed El-Rayes and Prof. Dr. Mohamed Geriesh, Faculty of Science, Suez Canal University, University, Ismailia, e-mail: ahmed_elrais@science.suez.edu.eg, helmi_geriesh@yahoo.com

Manuskript eingereicht: März 2012
Angenommen: August 2012

PFG 2012 / 6, 0727–0740
Stuttgart, December 2012

Report

Automatic 3D Object Reconstruction from Multiple Images for Architectural, Cultural Heritage and Archaeological Applications Using Open-Source Software and Web Services

Thomas P. Kersten & Maren Lindstaedt, Hamburg

Keywords: 3D, automation, comparison, image matching, modelling, point cloud, reconstruction

Summary: Constant improvements in the performance of internet and computer technologies combined with rapid advancements in computer vision algorithms now make it possible to efficiently and flexibly reconstruct the 3D geometry of objects. Objects of different sizes can be modelled using image sequences from commercial digital cameras that are processed by web services and freely available software packages, forming low-cost systems for numerous applications (restoration, historical care of monuments, visualization, analysis of the state of construction and the damage, etc.). In this contribution various cultural objects (historical buildings, statues/figures, archaeological finds, etc.) have been reconstructed in order to investigate the potential of this technology which enables the automatic generation of 3D point clouds or surface models (as 3D polygons) with photo-realistic texture from image data. These so-called low-cost systems represent an efficient alternative to expensive terrestrial laser scanning systems for the as-built documentation of 3D objects in architecture, cultural heritage and archaeology. The accuracy of the automatically generated 3D models is assessed by comparison with results from terrestrial laser scanning.

Zusammenfassung: *Automatische 3D-Objektrekonstruktion aus digitalen Bilddaten für Anwendungen in Architektur, Denkmalpflege und Archäologie durch open-source Software und Webservices.* Durch die stetig zunehmende Leistungsfähigkeit des Internets und der Computertechnologie sowie der raschen Weiterentwicklung von Computer Vision Algorithmen ist es heute möglich, die 3D-Geometrie von Objekten unterschiedlicher Größe mit handelsüblichen digitalen Kameras als Low-Cost-Systeme für zahlreiche Anwendungen (Restaurierung, historische Denkmalpflege, Visualisierung, Analyse des Bauzustandes und der Beschädigung, etc.) effizient und flexibel in Bildsequenzen zu erfassen. Anhand von diversen Kulturobjekten (historische Gebäude, Statuen/Figuren, archäologische Fundstücke, etc.) wird in diesem Beitrag das Potential von Webservices und frei verfügbaren Softwarepaketen aufgezeigt, mit denen 3D-Punktwolken oder Oberflächenmodelle (als 3D-Polygone) mit fotorealistischer Textur automatisch aus Bilddaten erzeugt werden. Diese so genannten Low-Cost-Systeme stellen heute für die As-Built-Dokumentation von 3D-Objekten in Architektur, Denkmalpflege und Archäologie eine effiziente Alternative zu teuren terrestrischen Laserscanningsystemen dar. Die Genauigkeit der automatisch erzeugten 3D-Modelle wird durch den Vergleich mit Ergebnissen des terrestrischen Laserscannings aufgezeigt.

1 Introduction

As state-of-the-art geodetic measuring methods photogrammetric multi-image techniques and, increasingly, terrestrial laser scanning, as a standalone system or in combination with other methods, are used for precise 3D data acquisition of complex objects. Requirements for the generation of 3D models are often very high with respect to level-of-detail, complete-

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0152 1432-8364/12/0152 $ 4.00

ness, reliability, accuracy (geometrical and visual quality), efficiency, data volume, costs and operational aspects, but the priority order depends upon the object to be recorded.

However, in recent years, real alternatives to classical systems and methods are presented by the large number of digital cameras on the market, which can be efficiently and successfully used as passive low-cost sensors when combined with appropriate algorithms such as structure-from-motion (SfM) and/or dense image matching for different 3D applications (object reconstruction, navigation, mapping, tracking, recognition, gaming, etc.). Due to the very low costs and current approval for open-source methods such systems (sensors in combination with appropriate algorithms) are very popular in many application fields. Nevertheless, the metrological aspect should not be neglected, if these systems are to be acknowledged as serious measuring and modelling procedures. Therefore, clear statements about the accuracy potential and efficiency of such systems must be empirically investigated through appropriate testing. In this context the 3D modelling results must be also analysed and compared with respect to reference data.

Practical examples of image-based modelling have been reported by Kersten et al. (2004), Kersten (2006) and Remondino & Menna (2008), while Remondino et al. (2008) generate comparable results for the 3D documentation of cultural monuments using image-based and range-based procedures in comparison. Barazzetti et al. (2009) present the combined use of photogrammetric and computer vision procedures for automatic and exact 3D modelling of terrestrial objects. They also show that similar results can be achieved with image-based and range-based recording systems. Barthelsen et al. (2012) present an approach for detailed and precise automatic dense 3D reconstruction of urban scenes using possibly unordered image sets from consumer cameras on a small unmanned aerial system.

In this paper the potential of web services and freely available software packages is investigated on the basis of practical examples where 3D point clouds or surface models (as 3D polygons) with photo-realistic texture are automatically derived from image data. After a brief introduction to the applied software in section 2 the entire workflow for the image-based low-cost 3D reconstruction procedure (section 3) is outlined. Practical results and 3D comparisons with reference data are summarized in section 4.

2 Applied Software

For investigation of the automatic generation of 3D point clouds and 3D surface models from image data the following software packages and/or web services were used: Bundler/PMVS2 and VisualSFM (open-source software), Microsoft Photosynth (web service), and Autodesk Photofly and/or 123D Catch Beta (web service).

2.1 *Bundler/PMVS2*

Bundler (Snavely et al. 2008) and PMVS2 (patch based multi view stereo software, Furukawa & Ponce 2010) were developed at the University of Washington in Seattle (USA) in C and C++ under the GNU General Public License as freely available software. Bundler works as a structure-from-motion (SfM) procedure for arbitrarily arranged imagery and

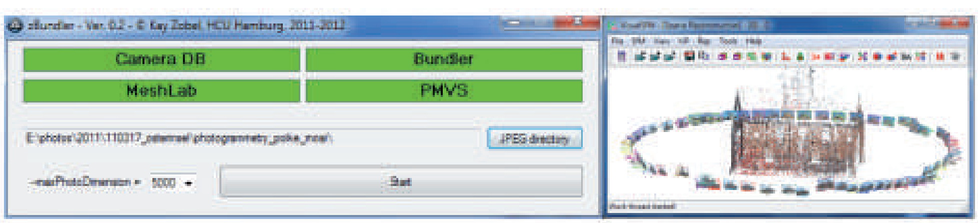


Fig. 1: Left: HCU graphical user interface (GUI) for the automatic workflow of Bundler/PMVS2; right: GUI for accelerated data processing with VisualSFM.

was developed for Microsoft's Photo Tourism Project (Snavely et al. 2006). Feature extraction in the images is performed by the SIFT (scale invariant feature transform) algorithm from Lowe (2004). The software supports camera calibration data (focal length f from EXIF data, two radial distortion parameters k1 and k2), image orientations and a thin 3D point cloud (scene geometry) as results for any image blocks using a modified bundle block adjustment from Lourakis & Argyros (2004). The results of Bundler are used in PMVS2 in order to generate a denser point cloud of non-moving objects by dense image matching. As well as the 3D coordinate each point additionally receives the colour value of the object taken from the images.

For use at HCU Hamburg a graphical user interface (Fig. 1, left) was developed which provides automation in the workflow between necessary software elements. As an example, after input of the images Bundler and PMVS2 are executed automatically and the result is finally presented in MeshLab. MeshLab is an open-source, portable, and extensible system for the processing and editing of unstructured 3D triangular meshes developed at the Visual Computing Lab, which is an institute of the National Research Council of Italy in Pisa (Cignoni et al. 2008).

2.2 *VisualSFM*

VisualSFM is a GUI application of multicore-accelerated SfM, which was developed at the University of Washington (Wu 2007). The software is a re-implementation of the SfM system of the Photo Tourism Project and it includes improvements by integrating both SIFT on the graphics processing unit (SiftGPU) and multicore bundle adjustment (Wu 2011). The camera parameters are defined as follows. The focal lengths (in pixels) of the camera are automatically calculated according to EXIF data. The principal point for each image is assumed to be at image centre except when using a single fixed calibration and the software uses only one radial distortion value. Dense reconstruction can also be performed through VisualSFM using PMVS/CMVS (patch or cluster based multi view stereo software, Furukawa & Ponce 2010).

2.3 *Microsoft Photosynth*

Photosynth has been developed for the Photo Tourism Project in co-operation between Microsoft Live Labs and the University of Washington (Snavely et al. 2006). The program Bundler forms the basis for automatic generation of 3D point clouds by free data processing using an external web service. For the use of Microsoft Photosynth a program for uploading the photos to a server and a Windows Live ID, e.g. email address, is initially necessary (Photosynth 2012). Depending upon the number of photos a result can be viewed in some seconds or a few minutes later online in all usual browsers; however no access to the data is possible. Only using the external program SynthExport from Hausner (2010), the computed 3D points and the camera parameters can be exported. However, the results correspond to those from Bundler. Nevertheless, Pomaska (2009) demonstrated how a photorealistic 3D model with true scale can be generated from imagery based on a low level polygon mesh and UV texture mapping using Photosynth for point cloud generation.

2.4 *Autodesk Photofly/123D Catch Beta*

In summer 2010 Autodesk introduced the project Photofly, a free web service, which allows the user to derive a meshed 3D model automatically from at least five overlapping photos of the recorded object (Abate et al. 2011). The basis of Photofly is the program smart3D-capture from the French company acute3D in Nice. Photofly uses algorithms from computer vision and photogrammetry and the performance of cloud computing in order to provide 3D models efficiently from 2D imagery. The fundamental algorithms of this software are described by Courchay et al. (2010), while detailed information about the algorithms used are not available from Autodesk.

Photofly uses the "Autodesk Photo Scene Editor", which must be installed on the user's computer as a communication platform between users and server. The very clear user interface of this software transfers the selected photos to the server. Depending upon the object complexity a 3D model will be provided for download within a short time period, i.e. usually in some minutes. This can be processed further within the software. Important functions in the software include selecting parts of the triangle meshes, navigation options, selecting of points and the definition of a reference distance for the absolute scaling of the model. Individual photos can be inserted into the model after the initial modelling by measurement of identical points (tie points). For the generation of 3D models three quality levels, mobile, standard and maximum (optimal result), are available. The results can be exported to different formats, e.g. OBJ or LAS. In November 2011 Photofly was replaced by 123D Catch Beta after the company acute3D has presented its software smart3Dcapture to the public in October 2011. Detailed information about 123D Catch and related tutorials are available under 123D (2012).

3 Workflow

The general workflow for image-based 3D reconstruction using low-cost systems is illustrated in Fig. 2 to document the degree of automation of the individual procedures which is symbolised in colour (red = manual, yellow = interactive and green = automatic).

For photogrammetric object recording multiple photos are taken of the object from different positions, whereby coverage of common object parts should be available from at least three but preferably five photographs from different camera positions. After import of the images into the respective processing software the parameters for camera calibration (interior orientation) and (exterior) image orientations are automatically computed. The subsequent generation of 3D point clouds or 3D surface models is also carried out in full automatic mode. Only for the 3D transformation of the point cloud or the meshed model into a superordinate coordinate system must the user measure control points interactively. The derived 3D model is automatically textured using the original image data so that video sequences, e.g. in 123D Catch Beta, can be generated using these models. If a CAD model needs to be constructed from a transformed and geo-referenced, coloured 3D point cloud, manual processing has to be carried out in a CAD program, e.g. AutoCAD. This 3D CAD model can later be manually textured in visualisation software, e.g. 3D studio, Cinema4D, Maya, etc., using the digital photographs in order to produce visualisations and/or video sequences.

4 Results and 3D Comparisons

In this section the results of the applied software packages Microsoft Photosynth, Bundler/PMVS2, VisualSFM and Autodesk Photofly and/or 123D Catch Beta, respectively,

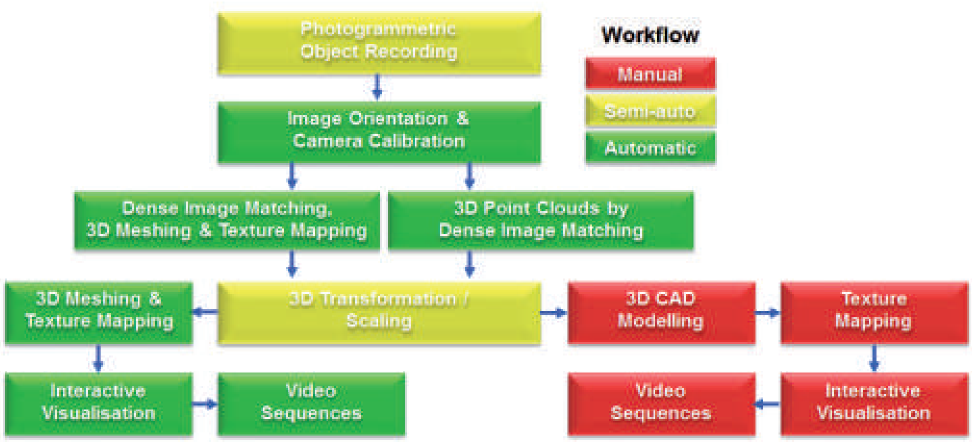


Fig. 2: Workflow for image-based low-cost 3D object reconstruction procedures.

Tab. 1: Statistics on different generated 3D objects using image-based systems, Pho. = photos.

Object	Camera/lens (mm)	Pixel	#Pho.	Software/service	# Points	# Triangle
Town house	Nikon D90 / 20	4288x2848	19	Photosynth	20,237	-
Town house	Nikon D90 / 20	4288x2848	19	Photofly	272,350	515,442
Town house	Nikon D90 / 20	4288x2848	19	Bundler/PMVS2	1,016,874	895,986
Fire House	Nikon D90 / 18	4288x2848	66	123D Catch	176,919	352,091
Fire House	Nikon D90 / 18	4288x2848	66	Bundler/PMVS2	1,541,924	-
Fire House	Nikon D90 / 18	4288x2848	66	VisualSFM	1,167,906	-
Zwinger	Nikon D90 / 28	4288x2848	15	Photosynth	18,553	-
Zwinger	Nikon D90 / 28	4288x2848	15	Photofly	155,697	285,669
Zwinger	Nikon D90 / 28	4288x2848	15	Bundler/PMVS2	917,965	-
Lion	Nikon D90 / 20	4288x2848	39	123D Catch	344,679	686,285
Lion	Nikon D90 / 20	4288x2848	39	Bundler/PMVS2	1,373,712	2,669,244
Moai Poike	Nikon D70 / 35	3008x2000	27	123D Catch	85,092	169,131
Moai Poike	Nikon D70 / 35	3008x2000	27	Bundler/PMVS2	629,644	-
Stone	Optio X, D70/80	diverse	48	123D Catch	-	291,613
Stone	Optio X, D70/80	diverse	48	Bundler/PMVS2	214,940	-
Pottery	Nikon D40 / 34	3008x2000	80	Bundler/PMVS2	-	323,402
Fragment	Nikon D90 / 24	4288x2848	58	VisualSFM	-	1,135,284

using the standard parameter values of each software package are presented for applications in architecture, cultural heritage and archaeology (Tab. 1), whereby individually generated datasets were compared with reference data from Zoller + Fröhlich's IMAGER 5006h and IMAGER 5010 terrestrial laser scanners (Tab. 2).

4.1 *Applications in Architecture*

Fig. 3 shows the results for the front façade of the Old Segeberger Town House (Bürgerhaus) – one of the oldest existing buildings in Schleswig-Holstein dated from the year 1539 with the front façade from 1606 – which are generated by Photosynth, Bundler/PMVS2 and Photofly. The 19 photographs used were acquired with a Nikon D90 (Nikkor 20 mm lens) from different positions as part of a student project at the HCU Hamburg. The result from Photosynth (20,237 points) is not suitable for façade modelling, while Photofly with 272,350 points and 515,442 triangles supplied the most attractive visual result. However, at the sides of the front of the façade and at the roof edges so-called virtual points, which do not exist in reality (Fig. 3, right), were meshed with Photofly. With Bundler/PMVS2 a dense point cloud of 1,016,874 points was generated (Fig. 3, centre right).

Fig. 3: Front façade of the old Town House (Bürgerhaus) in Bad Segeberg (19 photos from Nikon D90 with 20 mm lens); left: original photo; centre left: point cloud from Photosynth; centre right: Bundler/PMVS2; right: 3D meshing from Photofly.

The 3D meshing for the data from Bundler/PMVS2 of the front façade was carried out using Geo-magic Studio 12 and 2012. Geomagic Studio transforms 3D point clouds from range- or image-based systems and polygon meshes into accurate, usable 3D digital models for advanced product design, reverse engineering, custom manufacturing, CAD and analysis (Geomagic 2012). The meshed 3D front façade of Bundler/PMVS2 (from 5 mm grid spacing) and Photofly is visualized in Fig. 4 in comparison with the reference data of the IMAGER 5006h. According to the technical specifications of the system manufacturer the range noise of this scanner (1 sigma) for unfiltered raw data at 25 m distance is 2.6 mm for a reflectivity of 10% (black), 1.5 mm for a reflectivity of 20% (dark grey), and 0.7 mm for a reflectivity of 100% (white) using a scanning rate of 127,000 pixel/s (high power mode). Some areas at the black timber framework could not be measured with Bundler/PMVS2 or with the scanner thus small gaps are visible (Fig. 4, left). The triangle meshing of Photofly shows a noisy front with distinctive artefacts at the edges (Fig. 4, centre) while gaps were simply closed.

The absolute scaling of the data from Bundler/PMVS2 and Photofly was performed using measurements of two photogrammetric control points on the façade and the distance between them. Prior to 3D comparison both datasets were registered to the laser scanning data by an iterative closest point algorithm (ICP) (Besl & McKay 1992) in Geomagic. In the 3D comparison of the meshes from image-

Fig. 4: 3D meshing of the front façade from the Town House in Bad Segeberg; left: Bundler/PMVS2; centre: Photofly; right: IMAGER 5006h.

Fig. 5: Registration to the reference data of IMAGER 5006h (green < 3 mm); left: Bundler/PMVS2; right: Photofly.

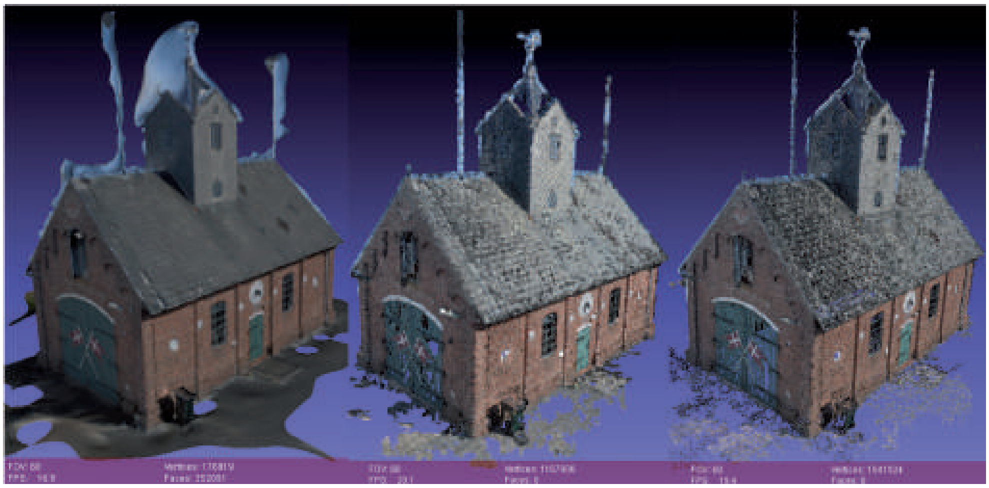


Fig. 6: Old Fire Brigade House located in the village Kirkeby of the Danish island Rømø; left: Meshed 3D model from 123D Catch Beta; centre: point cloud from VisualSFM; right: Bundler/PMVS2.

and range-based systems Bundler/PMVS2 clearly shows a better result than Photofly (Fig. 5) because most differences are smaller than 3 mm (see green areas). Due to the significantly higher number of 3D points obtained a geometrical better result was expected for Bundler/PMVS2 than for Photofly which is also documented in Tab. 2 through the lower number for the average deviation and the standard deviation.

In order to compare computation times an image block of 66 photographs with the Nikon D90 (Nikkor 18 mm lens) of the Old Fire Brigade House located in the village Kirkeby of the Danish island Rømø (Fig. 6) was computed with Bundler/PMVS2 and VisualSFM. The notebook used for processing consisted of an Intel Core i7 CPU Q740 processor with 1.73 GHz, an internal memory of 16 GB RAM running on the operating system Windows 7 Enterprise (64 bit), and a NVIDIA graphic card GeForce GT 445M. Computation time for Bundler/PMVS2 was 24 hours supplying 1,541,924 points, while the computing time of VisualSFM was only 42 minutes, which corresponds to a factor of 33 times faster than Bundler. However, approximately 374,000 points (24%) fewer were measured and these were particularly evident in the gaps within the shadow and roof areas (Fig. 6). Similar results (factor 24 faster, 16% fewer points) were obtained with an image block of 44 images (Nikon D90, 20 mm lens) from the former Swedish toll house in the old harbour of the city of Wismar (Kersten et al. 2012). In a further test the meshed model of 123D Catch Beta was computed in 16 min using the web service. This result looked visually better than the other two but the edges are smoothed and there are some virtual (geometrically incorrect) points at the antennas and on the top roof areas.

4.2 *Applications in Cultural Heritage*

As an example of applications in cultural heritage documentation the results of a figure from the Zwinger palace in Dresden are illustrated in Fig. 7. The figure was acquired from two different heights in 15 photographs using a Nikon D90 with a Nikkor zoom lens (focal length 28 mm). Photofly could generate a visually-appealing, near-complete mesh with 285,669 triangles from 155,697 points (Fig. 7, left), while with Bundler/PMVS2, despite measuring 917,965 points, gaps were visible (Fig. 7, centre) and in Photosynth a very small point density was measured with only 18,553 points (Fig. 7, right). Unfortunately, no reference data was available for this figure meaning that no geometrical analysis could be carried out.

The second example for applications in cultural heritage documentation is the 3D comparison of different results for the lion figure at the entrance of the imperial church (Kaiserdom) in Königlutter as illustrated in Figs. 9 and 10. The lion was acquired by image and range-based methods, i.e. in 39 photographs from different heights around the object using a Nikon D90 with a Nikkor lens (focal length 20 mm) and from two close scan stations using the laser scanner IMAGER 5006h. 123D Catch Beta could generate a visually appealing, almost complete meshed model with 686,285 triangles from 344,679 points (Fig. 8, top left), while with Bundler/PMVS2 despite 2,433,077 points, computed in 9 hours and 51 minutes (15 min per photo), some gaps are visible (Fig. 9, bottom left). The two registered clouds of the laser scanner which do not represent full coverage of the lion, were reduced to 660,000 points in total with a grid spacing of 2 mm due to the huge data volumes, which yielded a meshed model of 1,314,603 triangles. The 3D comparison of the meshed models from Bundler/PMVS2, also reduced to 2 mm grid spacing, and from 123D Catch Beta with the reference data was carried out in Geomagic after ICP registration of each model to the reference data. The results in Figs. 9 and 10 indicate deviations of ±2 mm to the reference for the most parts of the meshed models of the figure, while the average deviations are in the range of ±20 mm (Tab. 2) due to higher deviations at the wall and on the ground. The 3D comparison of 123D Catch Beta vs. IMAGER 5006h shows higher deviations especially in areas with higher curvature and edges due to smoothing effects, while the comparison of Bundler/PMVS2 vs. IMAGER 5006h demonstrates some systematic deviations on the front part of the lion.

4.3 *Applications in Archaeology*

Photofly and/or 123D Catch Beta could generate 3D models that are quite attractive visually when the archaeological objects are small and relatively rotund. A moai at the volcano crater Poike on Easter Island (Chile) was acquired in 27 surrounding photos with a Nikon D70 (Nikkor zoom lens with 35 mm). The result from 123D Catch Beta provided an almost-complete, textured 3D model (Fig. 10) with 169,131 triangles (from 85,092 points) while with Bundler/PMVS2 despite measurement of nearly 630,000 points, some gaps at the neck and at the bottom of the figure were visible.

Fig. 7: Figure in the Zwinger palace of Dresden; left: Autodesk Photofly; centre: Bundler/PMVS2; right: Microsoft Photosynth.

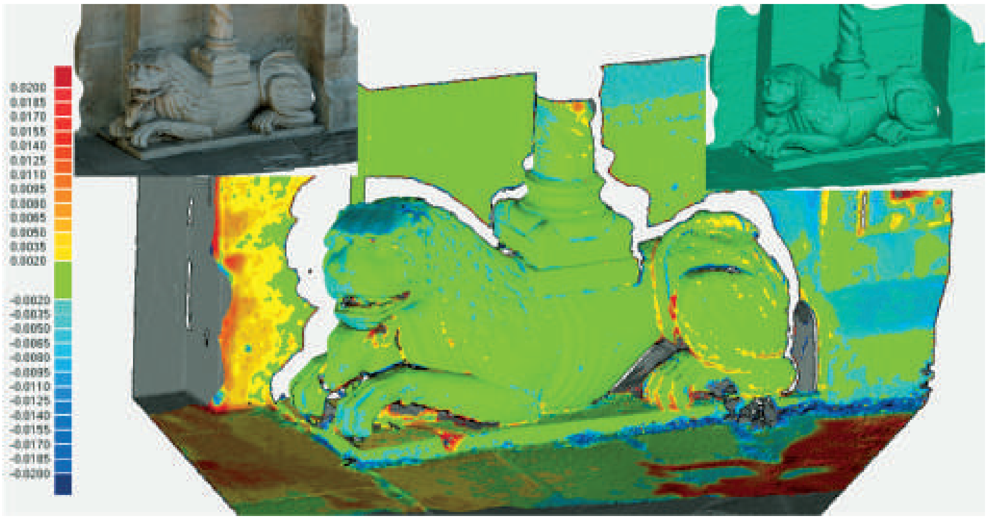


Fig. 8: 3D comparison of the lion figure at the imperial church in Königslutter between the 3D models from 123D Catch Beta and IMAGER 5006; top left: 3D model with texture; top right: without texture from 123D Catch Beta (green < 2 mm).

Fig. 9: 3D comparison of the lion figure at the imperial church in Königslutter between the 3D models from Bundler/PMVS2 and IMAGER 5006; bottom left: point cloud from Bundler/PMVS2; bottom centre: IMAGER 5006h; bottom right: meshed 3D model from the IMAGER 5006h (green < 2 mm).

A geometrical accuracy analysis of this data could not be carried out due to the absence of reference data.

A 3D model of the rear of a moai eye, carved from Obsidian stone, could only be successfully generated using 123D Catch Beta, because both laser scanner IMAGER 5006i and Bundler/PMVS2 could only measure very noisy point clouds with each including some gaps (Fig. 11). For measurements

Fig. 10: Small moai (height 0.7 m) at Poike on Easter Island (Chile); left: measured points; centre left: 3D meshing; centre right: textured 3D model in each case from 123D Catch Beta; right: dense coloured point cloud measured with Bundler/PMVS2 identifying some gaps (white ellipses).

Fig. 11: Black Obsidian stone as the rear of a moai eye at Easter Island; left: point cloud from IMAGER 5006i; centre: from Bundler/PMVS2; right: meshed 3D model from 123D Catch Beta.

with 123D Catch Beta and Bundler/PMVS2 image data from the following cameras were used: a) 27 photos from the Pentax Optio X (2560 x 1920 pixel) with 10 mm focal length, b) ten photos from a Nikon D70 (3008 x 2000 pixel) with a focal length of approx. 40 mm, and c) eleven photos from a Nikon D80 (3872 x 2592 pixel) with a focal length of 50 mm. The textured 3D model of the moai eye generated with 123D Catch Beta consists of 291,613 triangles (Fig. 11, right), while in total only 214,940 points were measured with Bundler/PMVS2 and only 95,820 points were scanned with the IMAGER 5006i. Although 3D comparison with reference data could not be conducted for accuracy analysis, subjective comparison of the provided 3D model suggests that it corresponds very well to the original. The model could not be scaled as no metric information was available.

For the next two examples (pottery/ceramic and architectural fragment from Yeha/Ethiopia) reference data were available. The pottery (ceramic(s)) was acquired by camera and laser scanner in order to compare image-based and range-based object recording. The pottery (Fig. 12) was photographed in the upper part with 50 photos and in the lower part with 30 photos using a Nikon D40 (Nikkor zoom lens, focal length 34 mm). With the terrestrial laser scanner IMAGER 5006h nine scans were required to achieve complete recording of the object. The triangle meshing of the pottery (approx. 160,590 triangles) was computed using the registered scans in Geomagic. The image data was processed with Bundler/PMVS2 using a reduced image resolution of 2400 pixels, i.e. a point cloud was separately generated for the upper and lower object part (Fig. 12). Using the measuring scale (10 cm), which was additionally placed into object space for the photographs, the scaling of the two datasets could be carried out. Firstly, the upper and lower 3D point clouds were registered using ICP algorithm. Secondly the 3D model was generated using the common point cloud with 0.5 mm grid spacing. Finally, a 3D comparison between model and reference was computed in Geomagic. This showed that differences larger than 1 mm only occurred at the boundary regions and in those areas with large curvatures (Fig. 12). These differences can be explained by the higher resolution of the imagery compared to the more smoothed scan data. The average deviation between image-based and range-based 3D model is 1 mm, while the standard deviation is 1.4 mm (Tab. 2).

Another archaeological finds from the excavation in Yeha (Ethiopia) was an architectural fragment which was photographed in two image blocks using a Nikon D90 (Nikkor zoom lens 24 mm): 33 photos (back side) and 25 photos (front side). A small scale bar was placed into object space for image acquisition thus facilitating scaling of the two datasets during processing. The image data was processed with VisualSFM using the full image resolution of the photographs for each image block producing two point clouds (595,933 points in total with grid spacing of 1 mm). After registration of the two point clouds with ICP algorithm a combined meshed 3D model (approximately 1,135,284 triangles in total) was computed (Fig. 13 left in each case).

As a reference, the object was scanned with the IMAGER 5010 from nine scan stations. According to the technical specifications of the system manufacturer the IMAGER 5010 has improved range noise behaviour. Range noise of this scanner (1 sigma) for unfiltered

Fig. 12: Left: Photograph of the pottery from Yeha (Ethiopia) with the measuring stick for object scaling; centre left: composing the pottery from Bundler/PMVS2; centre right: meshed and smoothed 3D model of the pottery from IMAGER 5006h; right: result of 3D comparison in Geomagic Studio 12 (green < 1 mm).

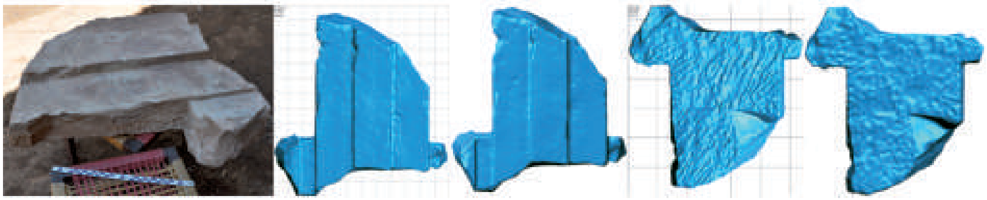


Fig. 13: Architectural fragment (archaeological find) in Yeha (Ethiopia): from left to right: photograph with scale bar; meshed 3D model of front (VisualSFM); meshed 3D model of front (IMAGER 5010); rear (VisualSFM); rear (IMAGER 5010).

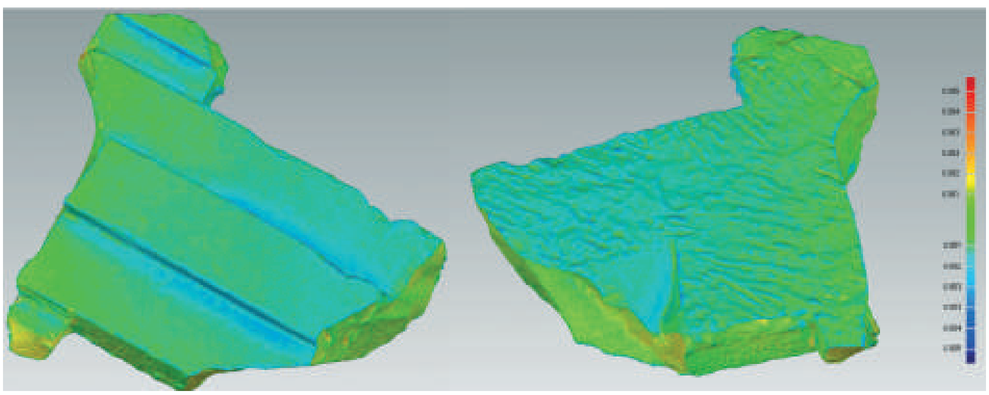


Fig. 14: Deviation between 3D model of the architectural fragment to reference data of the IMAGER 5010 illustrated in Geomagic Studio 2012 (green < 1 mm).

raw data at 10 m distance is 0.5 mm for a reflectivity of 14% (black), 0.4 mm for a reflectivity of 37% (dark grey), and 0.3 mm for a reflectivity of 80% (white). The registered point cloud (2,072,781 points) was filtered and reduced in Geomagic to a regular point distance of 2 mm, resulting in a mesh of approximately 1,120,580 triangles (Fig. 13). First the scan data was slightly smoothed using a noise filter. After ICP registration of the meshed model from image-based data with the meshed model from laser scanning a 3D comparison was performed in Geomagic producing the results of the front and rear sides of the fragment presented in Fig. 14. Differences from the image-based data to the reference, which are below 1 mm in the average (Tab. 2), are depicted in green. This comparison demonstrated that the dataset corresponds most closely to the result from laser scanning in the range of ±1 mm. Since the image data obviously supplied more details with a higher resolution than the smoothed scanning data in this case study (Fig. 14), these comparisons should primarily

Tab. 2: Comparison of image-based measurement (Geomagic Studio) and terrestrial laser scanning measurement (reference data); dev. = deviation, av. = average, std. = standard; (mm).

Object	Software/Criterion	# meshes	Max. dev.	Av. dev.+	Av. dev.-	Std. dev.
Town house	Photofly	504,449	530.0	36.6	-17.9	77.1
Town house	Bundler/PMVS2	895,986	517.0	7.5	-4.2	18.0
Lion	123D Catch Beta	686,285	153.0	18.7	-18.4	38.3
Lion	Bundler/PMVS2	2,669,244	153.0	14.9	-13.2	32.2
Pottery	Bundler/PMVS2	323,402	8.0	0.9	-1.0	1.4
Fragment	VisualSFM	1,135,284	−4.7	0.6	-0.8	0.8

detect gross errors in the models, which obviously did not arise here. A fringe projection (structured light) system could supply more precise reference data with higher resolution, but such a system was not available in Yeha/Ethiopia at that time. However, an average deviation of better than 0.3 mm in the 3D comparison between a 3D model of a similar fragment generated from images and scanned with a fringe projection system Breuckmann OptoTOP-HE confirmed the good results of these image-based systems (Kersten & Lindstaedt 2012).

5 Conclusion and Outlook

In this contribution economical, image-based recording and modelling procedures, which are able to generate precise and detailed 3D surface models from terrestrial photographs for applications in architecture, cultural heritage and archaeology, were presented. The quality of results from image-based systems depends on the number of images used, the image resolution, photo scale, illumination conditions and the parameter settings of the software applied. The results from some software packages, e.g. Bundler/PMVS2, using images from SLR digital cameras and standard parameter settings in the software are comparable with results from expensive terrestrial laser scanners. Object recording with cameras is simple, very fast, very flexible and economical. The entire procedure is to a large extent automated and works even without targets as control points. Objects can be scaled through the inclusion of a single scale bar in object space. The results presented here show that the open-source software Bundler/PMVS2, VisualSFM and the web service Photofly and/or 123D Catch Beta can generate equivalent 3D models when compared with terrestrial laser scanners, although the exact results depend upon the size and shape of the objects. However, the reliability of the image-based systems requires improvement because some of the achieved results were geometrically unusable. Depending upon object material and lighting conditions, noisy point clouds were sometimes produced with Bundler/PMVS2. Photofly/123D Catch Beta showed results that were visually very attractive (smoothed – whereby small holes are automatically filled by the software) when used on small and roundish objects. These results could fulfill the requirements of many users for visualisation in the Internet. However, the results from Photosynth are not useful for 3D modelling since the point density is too small. Nevertheless, further investigations with other more precise recording procedures such as fringe projection (structured light) or close-range scanners are necessary for small objects in order to be able to provide verified statements about the geometrical quality. Furthermore, CAD modelling with image-based 3D point clouds such as presented by Kersten et al. (2004), Kersten (2006) and Kersten (2007) should be investigated in the future to evaluate the potential of this automatically generated data.

These investigations have shown that Bundler/PMVS2 was more efficient than Photofly/123D Catch Beta for larger objects although the computation times were substantially longer. The significantly shorter processing time with the web service implies that Autodesk makes plentifully computing resources available for this service. Due to the optional use of these web services the resources of the user's own computer could be preserved, and a time-saving is achieved relative to the computation on a local PC. However, the availability of a fast Internet connection is not guaranteed everywhere in the world, particularly at the locations of many archaeological expeditions, meaning that the operation of open-source software on a user's own PC has an advantage. Furthermore, issues relating to data privacy should not be neglected when using web services.

Due to the sensible combination of computer vision algorithms and photogrammetric procedures the workflow from object recording, through 3D modelling to visualization has become increasingly automated without significantly neglecting geometrical accuracy. These combined procedures are only at the beginning of their development, since the speed of such algorithms can be significantly increased by future implementation of the software on graphics processor units (GPU) and because the mutual integration of both procedures can still be substantially optimized. A large number of algorithms for pixel-based matching in stereo or multi-view photographs with different performance potential from the field of computer vision are also available (Scharstein & Szeliski 2009) and have yet to be applied to this field. However, the level of automation is so high that many solutions are black-boxes with poor repeatability and low reliability. Remondino et al. (2012) present an investigation of automated image orientation packages in order to clarify potentialities and performances (including for camera calibration) when dealing with large and complex datasets. Their results demonstrate that, in case of complex and long sequences, SfM methods suffer of reliability and repeatability.

The dominant market position for the last 10 years of airborne and terrestrial laser scanning as tools for extensive data acquisition is being challenged by efficient photogrammetric procedures supported by computer vision algorithms and improved computer technology.

Acknowledgements

The authors would like to gratefully acknowledge the support of Kay Zobel (HafenCity University Hamburg) for developing the customised graphical user interface of Bundler/PMVS2 and MeshLab (Fig. 1), which provided an automatic workflow for the software.

References

123d, 2012: http://www.123dapp.com/catch/ (3.9.2012).

Abate, D., Furini, G., Migliori, S. & Pierattini, S., 2011: Project Photofly: New 3D Modelling Online WEB Service. – Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **38** (5/W16), http://www.isprs.org/proceedings/XXXVIII/5-W16/pdf/abate_etal.pdf (4.9.2012).

Barazzetti, L., Remondino, F. & Scaioni, M., 2009: Combined use of photogrammetric and computer vision techniques for fully automated and accurate 3D modeling of terrestrial objects. – SPIE Optics+Photonics, 7447, 2–3 August, San Diego, CA, USA.

Barthelsen, J., Mayer, H., Hirschmüller, H., Kuhn, A. & Michelini, M., 2012: Orientation and Dense Reconstruction from Unordered Wide Baseline Image Sets. – PFG **2012** (4): 421–432.

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

Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F. & Ranzuglia, G., 2008: MeshLab: an open-source mesh processing tool. – Scarano, V., De Chiara, R. & Erra, U. (eds.): Eurographics Italian Chapter Conference, The Eurographics Association: 129–136.

Courchay, J., Pons, J.-P., Monasse, P. & Keriven, R., 2010: Dense and Accurate Spatio-temporal Multi-view Stereovision. – Computer Vision – ACCV 2009, Lecture Notes in Computer Science **5995:** 11–22.

Furukawa, Y. & Ponce, J., 2010: Accurate, Dense, and Robust Multi-View Stereopsis. – IEEE Transactions on Pattern Analysis and Machine Intelligence **32** (8): 1362–1376.

Geomagic, 2012: http://www.geomagic.com (3.9.2012).

Hausner, C., 2010: SynthExport. http://synthexport.codeplex.com/ (4.1.2012).

KERSTEN, TH., 2006: Combination and Comparison of Digital Photogrammetry and Terrestrial Laser Scanning for the Generation of Virtual Models in Cultural Heritage Applications. – IOANNIDES, M., ARNOLD, D., NICCOLUCCI, F. & MANIA, K. (eds.): 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage, VAST **2006:** 207–214.

KERSTEN, TH., 2007: Virtual Reality Model of the Northern Sluice of the Ancient Dam in Marib/Yemen by Combination of Digital Photogrammetry and Terrestrial Laser Scanning for Archaeological Applications. – International Journal of Architectural Computing, Special Focus on Cultural Heritage **02** (05): 339–354, Multi-science.

KERSTEN, TH., ACEVEDO PARDO, C. & LINDSTAEDT, M., 2004: 3D Acquisition, Modelling and Visualization of north German Castles by Digital Architectural Photogrammetry. – Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **35** (5): 126–132.

KERSTEN, TH. & LINDSTAEDT, M., 2012: Image-Based Low Cost Systems for Automatic 3D Recording and Modelling of Archaeological Finds and Objects. – EuroMed 2012 – Int. Conference on Cultural Heritage. – IOANNIDES, M., FRITSCH, D., LEISSNER, J., DAVIES, R., REMONDINO, F. & CAFFO, R. (eds.). – Lecture Notes in Computer Science (LNCS) **7616**: 1–10, Springer.

KERSTEN, TH., LINDSTAEDT, M., MECHELKE, K. & ZOBEL, K., 2012: Automatische 3D-Objektrekonstruktion aus unstrukturierten digitalen Bilddaten für Anwendungen in Architektur, Denkmalpflege und Archäologie. – SEYFERT, E. (Hrsg.): Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V. **21:** 137–148, DGPF, Potsdam, CD-ROM.

LOURAKIS, M.I.A. & ARGYROS, A.A., 2004: Design and Implementation of a Generic Sparse Bundle Adjustment Software Package Based on the Levenberg-Marquardt Algorithm. – Institute of Computer Science, FORTH-ICS, Tech. Rep. **340**, Heraklion, Greece, http://www.ics.forth.gr/~lourakis/sba (4.9.2012).

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

PHOTOSYNTH, 2012: http://photosynth.net/ (4.9.2012).

POMASKA, G., 2009: Utilization of Photosynth Point Clouds for 3D Object Reconstruction. – **22nd** CIPA Symposium, Kyoto, Japan, http://cipa.icomos.org/fileadmin/template/doc/KYOTO/34.pdf (4.9.2012).

REMONDINO, F. & MENNA, F., 2008: Image-based surface measurement for close-range heritage documentation. – Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **37** (B5-1): 199–206.

REMONDINO, F., EL-HAKIM, S.F., GRÜN, A. & ZHANG, L., 2008: Turning images into 3-D models. – IEEE Signal Processing Magazine **25** (4): 55–65.

REMONDINO, F., DEL PIZZO, S., KERSTEN, TH. & TROISI, S., 2012: Low-cost and Open-Source Solutions for Automated Image Orientation – a Critical Overview. – EuroMed 2012 – Int. Conference on Cultural Heritage. – IOANNIDES, M., FRITSCH, D., LEISSNER, J., DAVIES, R., REMONDINO, F. & CAFFO, R. (eds.). – Lecture Notes in Computer Science (LNCS) **7616**: 40–54, Springer.

SCHARSTEIN, D. & SZELISKI, R., 2009: Middlebury Stereo Vision Page. http://vision.middlebury.edu/stereo/ (7.1.2012).

SNAVELY, N., SEITZ, S.M. & SZELISKI, R., 2006: Photo tourism: exploring photo collections in 3D. – ACM Transactions on Graphics **25** (3): 835–846.

SNAVELY, N., SEITZ, S.M. & SZELISKI, R., 2008: Modeling the World from Internet Photo Collections. – International Journal of Computer Vision **80** (2): 189–210.

WU, C., 2007: SiftGPU: A GPU implementation of Scale Invariant Feature Transform (SIFT). http://cs.unc.edu/~ccwu/siftgpu (4.9.2012).

WU, C., 2011: VisualSFM: A Visual Structure from Motion System. http://www.cs.washington.edu/homes/ccwu/vsfm/ (4.9.2012).

Address of the authors:

THOMAS P. KERSTEN and MAREN LINDSTAEDT, HafenCity Universität Hamburg, Labor für Photogrammetrie & Laserscanning, Hebebrandstraße 1, D-22297 Hamburg, Tel.: +49-40-42827-5343, Fax: +49-40-42827-5399, e-mail: {thomas.kersten}{maren.lindstaedt}@hcu-hamburg.de

Manuskript eingereicht: April 2012
Angenommen: August 2012

Berichte von Veranstaltungen

22. ISPRS-Kongress in Melbourne, Australien, 25.8.–1.9.2012

Wie wir alle wissen, war 2012 wieder das Jahr des ISPRS-Kongresses, diesmal des zweiundzwanzigsten.

Diese Veranstaltung ist weltweit das wichtigste Forum für Wissenschaft und Technik der Photogrammetrie und fast ebenbürtig auch der Fernerkundung. Neben dem Fachprogramm war die Entwicklung und Pflege persönlicher Kontakte mehr denn je wichtig im Zeitalter der elektronischen Kommunikation.

Die Generalversammlung der ISPRS wählte Prof. Christian Heipke zum neuen Generalsekretär der ISPRS. Die DGPF gratuliert ihm ganz herzlich zu diesem Erfolg und wünscht ihm in den kommenden vier Jahren immer eine glückliche Hand in der Gestaltung der ISPRS.

Der neue ISPRS-Vorstand sieht wie folgt aus:

Präsident: Chen Jun, China
Generalsekretär: Christian Heipke, Deutschland
1. Vizepräsident: Orhan Altan, Türkei
2. Vizepräsidentin: Maguerite Madden, USA
Schatzmeister: Jon Mills, Großbritannien
Kongressdirektorin 2016: Lena Halounova, Tschechische Republik

Der Kongress war durch folgende Zahlen charakterisiert:

Registrierte Teilnehmer: 1941
Teilnehmende Länder: 74
Die vier größten Ländergruppen: Australien (500), China (320), Deutschland (156) und Japan (114)
Tagesbesucher: 71

Begutachtete Artikel: 292
Artikel für die ISPRS-Archives: 787
Eingeladene Artikel: 30
Plenarvorträge: 11
Kurzbeiträge: 382

22. ISPRS-Kongress: Thomas Kolbe, Christian Heipke (ISPRS-Generalsekretär) und Uwe Stilla in Melbourne, Australien.

DOI: 10.1127/1432-8364/2012/0153 1432-8364/12/0153 $ 4.00

ISPRS-Vorstand: Chen Jun, Christian Heipke, Orhan Altan, Lena Halounova, Marguerite Madden, Jon Mills.

Teilnehmer an Workshops und Tutorials: 95
Teilnehmer an Fachexkursionen: 65

Beim Kongress gab es eine Reihe von offiziellen Ehrungen der ISPRS. Als neue Ehrenmitglieder wurden Ian Dowman (Großbritannien) und Deren Li (China) gewählt. Von den 28 weiteren Ehrungen entfielen sechs auf deutsche Wissenschaftler:

Dieter Fritsch, Stuttgart:
Fellowship-Preis, verliehen für hervorragende Dienste für die ISPRS

Christian Heipke, Hannover:
erstmals vergebener Frederick J. Doyle Preis, verliehen für Beispiel gebende Leistungen und Erfolge bei der Entwicklung von Photogrammetrie, Fernerkundung und Geoinformatik

Jan-Henrik Haunert, Würzburg, vorher Hannover:
Otto von Gruber Preis, verliehen für eine hervorragende Veröffentlichung eines Nachwuchswissenschaftlers, hier Dissertation zur kombinatorischen Optimierung für die Generalisierung von Geodaten

Hans-Gerd Maas, Dresden:
Karl Kraus Medaille, verliehen für ein gutes Lehrbuch, hier Airborne and Terrestrial Laser Scanning, zusammen mit George Vosselman

Uwe Sörgel und Franz Rottensteiner, beide Hannover:
President's Citations, verliehen für Engagement und Erfolge bei der Leitung von ISPRS-Arbeitsgruppen

In Melbourne hat Elsevier als Verlagshaus des "ISPRS Journal of Photogrammetry and Remote Sensing" erstmals Auszeichnungen für "excellent reviewers" vergeben, und zwar für den Zeitraum der letzten acht Jahre. Mit diesem Preis wird die harte, freiwillige Arbeit im Hintergrund geehrt, ohne welche gute Wissenschaft nicht möglich wäre. Unter den sieben Ausgezeichneten sind drei aus Deutschland: Johannes Engels, Stuttgart, Uwe Weidner, Karlsruhe, und Helmut Mayer, Neubiberg.

Der 23. Kongress wird 2016 in Prag, Tschechische Republik, stattfinden.

Wolfgang Kresse, Neubrandenburg und
Stefan Hinz, Karlsruhe

RSSS12 – Remote Sensing Summer School 2012, 19.–20. Juli 2012 in München

Das diesjährige *IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2012)* fand vom 22. bis zum 27. Juli im internationalen Konferenzcenter in München statt und bot den mehr als 2700 Teilnehmern aus 68 Ländern über 2500 Präsentationen. In der Vorbereitungsphase zu der nun nach 30 Jahren wieder in München organisierten Veranstaltung kam im Präsidium der *Geoscience and Remote Sensing Society (GRSS)* der Wunsch auf, neben den im Vorfeld zum Symposium angebotenen Tutorials erstmals auch eine Sommerschule für Studierende und Doktoranden zu veranstalten. Diese Aufgabe hat Prof. Uwe Stilla (Senior Member IEEE) mit seinem Team von der Technischen Universität München übernommen.

Begrüßung der Teilnehmer durch Uwe Stilla.

Für die Auftaktveranstaltung wurde von Uwe Stilla das Akronym RSSS12 (Remote Sensing Summer School 2012) eingeführt und sowohl generell über die Webseite der IGARSS 2012, verschiedene Veranstaltungskalender und die Internetplattform LinkedIn, als auch direkt durch Anschreiben von Instituten und studentischen Vertretungen der Universitäten beworben. Aufgrund der finanziellen Unterstützung der GRSS für Nachwuchsförderung konnte die Teilnahmegebühr für die Veranstaltung inklusive des gebundenen Vorlesungsmanuskripts mit 330 Seiten, der Tagesverpflegung und zwei Abendveranstaltungen für 50 € angeboten werden.

Kurz nach Freischaltung der Webseite zur Anmeldung waren 85% der verfügbaren Plätze vergeben. Die letzten 15% der Plätze wur-

Teilnehmer der RSSS12 vor dem Haupteingang der Technischen Universität München.

Ice-Breaker an der TUM, rechts: Präsident der GRSS, Jón Atli Benediktsson, mit Vortragenden.

„Networking" im Münchner Hofbräuhaus.

den unter allen folgenden Anmeldungen verlost. Insgesamt wurden von den 120 Bewerbungen aus 25 Ländern 70 Studierende aus 20 Ländern eingeschrieben.

Eröffnet wurde die RSSS12 mit der Begrüßung durch Uwe Stilla, der einen Überblick über Veranstaltungsablauf und Vorlesungen gab und die Vortragenden vorstellte. Die zweitägige Veranstaltung gliederte sich in acht Vorlesungen mit jeweils zwei Vorlesungen am Vor- und Nachmittag und zwei geselligen Abendveranstaltungen. Als Vortragende wurden international bekannte Experten eingeladen, die insbesondere die Themenbereiche Hyperspektral- und Radarfernerkundung vertieften.

Der erste Vortrag *„Methods and Systems for Forest Applications"* von David Goodenough (University of Victoria, Canada) befasste sich mit der Integration von Multi-Sensor-Daten und historischen Informationen in Geoinformationssystemen. Lorenzo Bruzzone (University of Trento, Italien) erläuterte in seinem Vortrag *„Challenges in the analysis of multitemporal remote sensing images"* die aktuellen Forschungsentwicklungen im Bereich Änderungsdetektion aus Bilddaten der neuesten Satellitenmissionen. Das Thema Segmentierung von Hyperspektraldaten und ihre Weiterverarbeitung mit geometrischen und statistischen Methoden war Schwerpunkt des Vortrags *„Hyperspectral Image and Signal Processing"* von Jocelyn Chanussot (Grenoble Institute of Technology, Frankreich) und Antonio Plaza (University of Extremadura, Spanien). Melba Crawford (Purdue University, USA) führte schließlich das Thema Hyperspektraldaten in ihrem Vortrag *„Advanced Methods for Classification of Hyperspectral Data"* fort und ging insbesondere auf überwachte und semi-überwachte Verfahren ein. Einen Fokus auf urbane Szenen setzte Paolo Gamba (University of Pavia, Italien) mit seinem Vortrag *„Multispectral Urban Remote Sensing and Data Fusion"*. Uwe Stilla (Technische Universität München) konzentrierte sich in seinem Thema *„SAR Urban Remote Sensing"* auf den Einfluss hochaufgelöster SAR Daten auf die Analyse urbaner Szenen und die speziellen Eigenschaften urbaner Objekte in SAR-Abbildungen. Richard Bamler (DLR) erweiterte das Thema SAR um *„SAR Interferometry and Tomography"* und erläuterte die Funktionsweise und den Einfluss verschiedener Konfigurationen der Interferometrie. Um *„Morphological Profiles and Attribute Profiles"* ging es schließlich im Vortrag von Prashanth Marpu (Masdar Institute of Science and Technology, Vereinigte Arabische Emirate) und Mauro Dalla Mura (University of Trento, Italien). Sie erläuterten den Einsatz von morphologischen und attributierten Profilen bei Klassifikation, Segmentierung und Objektextraktion in der Fernerkundung.

Vor der Ice-Breaker Party, die am ersten Abend in den Räumen der Photogrammetrie und Fernerkundung der TUM stattfand, begrüßte der GRSS-Präsident Jón Atli Benediktsson die Teilnehmer und dankte dem TUM-Team für die Organisation der Veranstaltung. Auf Grund der positiven Resonanz beabsichtigt die GRSS zukünftig jährlich im Vorfeld der IGARSS eine Summer School abzuhalten. Die Teilnehmer nutzten bis spät in die Nacht intensiv die Möglichkeit für einen

Austausch untereinander und für ein Gespräch mit den Referenten.

Am zweiten Abend ergab sich beim gemeinsamen Essen im traditionellen Festsaal des Münchner Hofbräuhauses erneut die Möglichkeit das Netzwerk mit anderen Studierenden und Doktoranden der Fernerkundung weiter auszubauen und die bereits geknüpften persönlichen Kontakte zu vertiefen. Begeistert wurde auch das bayrische Bühnenprogramm aufgenommen. Viele der ausländischen Teilnehmer sahen zum ersten Mal erstaunt, dass Peitschen als Begleitinstrumente zum Akkordeon eingesetzt werden können.

Zusammenfassend kann auf eine erfolgreiche Veranstaltung zurückgeblickt werden, was auch die nach der RSSS12 durchgeführte Email-Umfrage bei den Teilnehmern bestätigt. Interessant ist dabei auch zu erwähnen, dass ein bemerkenswerter Anteil der ausländischen Teilnehmer nur wegen der RSSS12 angereist war, ohne die nachfolgende IGARSS zu besuchen. Sicherlich würde ein ähnliches Konzept mit ausgewählten Experten der Fernerkundung als nationale oder internationale Veranstaltung im Rahmen der Nachwuchsförderung der DGPF auf Zuspruch stoßen.

Das Vortragsprogramm, die Kurzfassungen der Vorlesungen und eine Fotogalerie finden sich unter http://www.igarss2012.tum.de.

Ludwig Hoegner, München

Hochschulnachrichten

Technische Universität München

Herr Dipl.-Ing. Jens Leitloff promovierte am 26.9.2011 an der Fakultät für Bauingenieur- und Vermessungswesen (Fachgebiet Photogrammetrie und Fernerkundung) der Technischen Universität München mit der Arbeit *„Detektion von Fahrzeugen in optischen Satellitenbildern“* zum Dr.-Ing.

1. Gutachter: Prof. Dr.-Ing. Uwe Stilla, Technische Universität München (TUM),
2. Gutachter: Prof. Dr.-Ing. Stefan Hinz, Karlsruher Institut für Technologie (KIT)

Kurzfassung:
In den vergangenen 40 Jahren ist das Verkehrsaufkommen durch Privatpersonen um mehr als das 3,5-fache gestiegen und der Straßengüterverkehr nahm um ca. 800% zu. In urbanen Gebieten verursacht der Verkehr bis zu 75% der Feinstaubbelastung. Daher besteht ein erhöhter Bedarf an intelligenterer Verkehrsüberwachung und -beeinflussung, welchen zunächst eine bessere Modellierung vorausgeht. Die Erfassung der hierfür benötigten Daten erfolgt meist durch terrestrische stationäre Sensoren wie Induktionsschleifen und Brückenkameras. Da diese Systeme nur einen Teil des Straßennetzes abdecken, wird im Rahmen der Arbeit ein neuer Ansatz zur Erfassung verkehrsrelevanter Parameter aus optischen Satellitenbildern entwickelt. Die flächendeckend gewonnenen Daten können als wichtige Ergänzung für vorhandene Verkehrsmodelle genutzt werden.

Um eine möglichst geringe Fehlerrate bei der Fahrzeugerkennung zu erreichen, werden Daten über das Straßennetz aus vorhandenen Geoinformationen genutzt. Diese dienen hauptsächlich zum Ausschluss nicht relevanter Bildregionen. Da Fahrzeuge in Satellitenbildern keine Merkmale wie Front- oder Heckscheiben zeigen, erweisen sich etablierte Verfahren zur Fahrzeugerkennung aus höher aufgelösten Daten als nicht anwendbar.

In dieser Arbeit wird ein neuer Ansatz entwickelt, der sowohl isolierte als auch in Gruppen auftretende Fahrzeuge berücksichtigt. Es wird ein zweistufiges Klassifikationsverfahren zur Detektion isolierter Fahrzeuge vorgeschlagen. Dieses erreicht eine hohe Zuverlässigkeit. Hypothesen für Fahrzeugreihen werden durch differentialgeometrische Linienextraktion gebildet. Zur Erkennung von Einzelfahrzeugen innerhalb der Fahrzeugreihen werden speziell auf die Problemstellung angepasste robuste Schätzverfahren eingehend untersucht. Nach Zusammenführung der Ergebnisse der Mustererkennung und der robusten Parameterschätzung ist es möglich, eine quantitative Abschätzung über das Verkehrsaufkommen einzelner Straßen zu geben.

Erstmalig wird für Einzelfahrzeuge automatisch der Bewegungsstatus geprüft. Hierfür wird die zeitliche Differenz zwischen den Aufnahmen der einzelnen Spektralkanäle genutzt. Die Initialisierung der verwendeten Methoden erfordert kein Expertenwissen. Alle zur Detektion und Bewegungsschätzung verwendeten Parameter werden aus manuell erfassten Trainingsdaten abgeleitet oder durch statistische Testverfahren gewonnen.

Das Verfahren wurde mit QuickBird Bildern eines städtischen Bereichs evaluiert. Während die Vollständigkeit der Ergebnisse bei 65% liegt, erreicht der Ansatz eine hohe Zuverlässigkeit von durchschnittlich 95%. Für einige Fahrzeuge konnte ein signifikanter Versatz zwischen den Aufnahmen der Spektralkanäle berechnet werden. Die daraus abgeleitete Geschwindigkeit ist für städtische Gebiete plausibel.

Die Dissertation ist verfügbar unter: www.pf.bv.tum.de/pub/2011/leitloff_phd11_dis.pdf

Technische Universität München

Herr Dipl.-Math. techn. Marcus Hebel promovierte am 18.9.2012 an der Fakultät für Bauingenieur- und Vermessungswesen (Fachgebiet Photogrammetrie und Fernerkundung) der Technischen Universität München mit der Arbeit *„Änderungsdetektion in urbanen Gebieten durch objektbasierte Analyse und schritthaltenden Vergleich von Multi-Aspekt ALS-Daten"* zum Dr. rer. nat.

1. Gutachter: Prof. Dr.-Ing. Uwe Stilla, Technische Universität München (TUM),
2. Gutachter: Prof. Dr.-Ing. Rüdiger Westermann, Technische Universität München (TUM),
3. Gutachter: Prof. Dr. rer. nat. Maurus Tacke, Fraunhofer IOSB.

Kurzfassung:
Zur Automatisierung der Änderungsdetektion in urbanen Gebieten werden typischerweise Fernerkundungsdaten zeitversetzt aus der Vogelperspektive aufgenommen und analysiert. Dabei werden besondere Anforderungen an die eingesetzte Sensorik und die Methodik zur Datenauswertung gestellt, wenn ein sofortiges Vorliegen der Ergebnisse unerlässlich ist. Solche Randbedingungen bestehen z.B. bei der Unterstützung von Hubschrauberpiloten im Rahmen von Überwachungsaufgaben oder Rettungseinsätzen.

Zum Monitoring städtischer Gebiete wird in dieser Arbeit eine Änderungsdetektion mit Daten hubschraubergetragener Laserscanner vorgeschlagen (engl. Airborne Laser Scanning, ALS). Eine Besonderheit des hierzu verfolgten Ansatzes ist die Betrachtung von Multi-Aspekt ALS-Daten, die sich durch die Verwendung eines in Schrägsicht vorausblickenden Laserscanners ergeben. Diese Sensorkonfiguration ist im Hinblick auf die oben genannten Aufgaben erforderlich und ermöglicht außerdem eine für ALS sonst untypische Erfassung von Fassadenflächen.

Im ersten Teil der Arbeit wird eine Methodik vorgestellt, durch die sowohl eine Kalibrierung des Sensorsystems als auch eine Zusammenführung der Multi-Aspekt ALS-Daten eines urbanen Gebiets erzielt werden. Die dazu beschriebene Vorgehensweise ist insbesondere auch für die hier untersuchte Schrägsicht des Laserscanners geeignet. Im Zuge einer objektbasierten Analyse der einzelnen Punktwolken werden planare Flächenstücke mit Hilfe eines Segmentierungsverfahrens identifiziert, das ein Flächenwachstumsverfahren mit einem RANSAC-Schätzverfahren kombiniert. Anschließend werden homologe Flächenstücke anhand geometrischer Attribute ausfindig gemacht. Mit Hilfe einer neuartigen Methode können Planaritätsbedingungen für diese Zuordnungen in lineare Gleichungssysteme überführt werden, durch deren Lösung sich einerseits die Boresight-Kalibrierung des ALS-Systems und andererseits die Angleichung der ALS-Datensätze durchführen lassen.

Der zweite Teil der Arbeit behandelt die während einer erneuten Befliegung stattfindende Änderungsdetektion, die auf den zuvor vereinheitlichten Referenzdaten aufbaut. Es wird eine neue Herangehensweise für den schritthaltenden Vergleich von ALS-Daten vorgestellt. Anstelle eines Vergleichs von Punktwolken werden dabei 3D-Raumbereiche entlang der Ausbreitungswege der Laserpulse bezüglich der Zustände *leer*, *belegt* und *unbe-*

stimmt bewertet. Das dazu vorgeschlagene Vorgehen basiert auf der Wissensrepräsentation und Informationsfusion entsprechend der Dempster-Shafer Evidenztheorie, wobei Änderungen als Konflikte in der Raumbelegung erkennbar werden. Zusätzlich werden Objektmerkmale ausgewertet, um Änderungsereignisse verschiedenen Kategorien zuzuordnen.

Im dritten Teil der Arbeit werden die durchgeführten Experimente zur Systemkalibrierung, Datenregistrierung und Änderungsdetektion vorgestellt und diskutiert. Die Kalibrierung des verwendeten ALS-Experimentalsystems konnte bei allen erfassten urbanen Gebieten zuverlässig vorgenommen werden. Das erarbeitete Verfahren zur schritthaltenden Änderungsdetektion wurde ebenfalls erfolgreich anhand realer Multi-Aspekt ALS-Daten validiert, wofür ein ausgewähltes urbanes Gebiet im Abstand eines Jahres jeweils entlang mehrerer Flugrichtungen erfasst wurde.

Die Dissertation ist verfügbar unter: www.pf.bv.tum.de/pub/2012/hebel_phd12_dis.pdf

Wettbewerb: Ableitung von 3D-Strukturlinien aus 3D-Punktwolken

Für vielfältige Aufgaben der Vermessungsverwaltungen der Länder und der Wasser- und Schifffahrtsverwaltung des Bundes werden 3D-Strukturlinien, z.B. 3D-Bruch- und Geländekanten oder 3D-Kantenlinien, als ein Basisprodukt zur Bereitstellung nutzerorientierter geotopographischer Produkte benötigt. So werden diese 3D-Strukturlinien u. a. benötigt zur

- Vervollständigung des Grunddatenbestandes im Sinne der GeoInfoDok unter Beachtung des AAA-Standards,
- topographischen Modellierung im ATKIS Basis-DLM,
- kartographischen Ausgestaltung der ATKIS DTK-Reihe,
- optimierten Ableitung von DOP,
- Modellierung von DGM-W und
- hydronumerischen Modellierung.

In diesem Kontext ist nach wie vor eine automatisierte, auch zumindest weitestgehend automatisierte, praxisreife Lösung zur Ableitung von 3D-Strukturlinien aus 3D-Punktwolken hoch aufgelöster ALS-Daten oder Bildkorrelationsdaten nicht verfügbar. Eine adäquate Lösung wird aber aus fachlichen und wirtschaftlichen Gründen benötigt, da ressourcen- und zeitbedingt rein interaktive Erfassungen von 3D-Strukturlinien nicht realisierbar sind. Vertreter der Projektgruppe „ATKIS®-DGM“ der AdV führen eine Markterkundung zur Suche von interessierten und grundsätzlich geeigneten Institutionen durch, die alleine oder als Arbeitsgemeinschaft mit der Ausführung einer Fachstudie beauftragt werden können. Ebenso sollen die Kosten der Fachstudie abgeschätzt werden.

Um eine Problemlösung im Sinne der genannten Verwaltungen zu forcieren, ist ein 3-stufiges Vorgehen angedacht:

- Stufe 1: Markterkundung
- Stufe 2: Fachstudie
- Stufe 3: Realisierung einer praxisreifen Softwarelösung

Die Fachstudie soll konkrete Ansätze und Lösungen zur Verfügung stellen, um 3D-Strukturlinien aus 3D-Punktwolken praxisreif abzuleiten.

Zur Erstellung der Fachstudie werden dem Auftragnehmer exemplarische Daten, 3D-Punktwolken, zur Verfügung gestellt, an denen nachgewiesen werden soll, in welcher Quantität und Qualität 3D-Strukturlinien für die relevanten Produktklassen geländeklassenbezogen abgeleitet werden können. Das Vorgehen einschließlich aller Ansätze mit den Algorithmen und den Anforderungen an die Stufe 3 ist detailliert zu dokumentieren.

Interessierte Institutionen können weitergehende Informationen unter folgendem Kontakt anfordern. Eine Interessenbekundung zur Teilnahme am Wettbewerb ist bis Ende März 2013 zu stellen an:

Herbert Brockmann
Bundesanstalt für Gewässerkunde
Tel.: +49 (0) 261 13065214
e-mail: brockmann@bafg.de

Veranstaltungskalender

2012

6.–7. Dezember: **3D-NordOst 2012**, 15. Anwendungsbezogener Workshop zur Erfassung, Modellierung, Verarbeitung und Auswertung von 3D-Daten am GFaI in Berlin-Adlershof. 3d-nordost.de.

2013

16.–18. Januar: **Workshop on the Applications of Computer Vision (WACV 2013)** in **Cleanwater Beach, Tampa**, USA. cvl.cse.sc.edu/wacv2013/

13.–14. Februar: **Oldenburger 3D-Tage**. jade-hs.de/3dtage

17.–23. Februar: 17. **Internationale Geodätische Woche** in **Obergurgl**, Österreich. uibk.ac.at/vermessung/veranstaltung/obergurgl.html

27. Februar – 1. März: **DGPF Jahrestagung 2013 (Dreiländertagung)** in **Freiburg**. dgpf.de/neu/jahrestagung/informationen.htm

13.–15. März: **Geoinformatics 2013** in **Heidelberg**. geoinformatik2013.de

19.–20. März 2013: **Internationales 3D-Forum Lindau**, 3d-forum.li

24.–28. März: **ASPRS Annual Meeting** in **Baltimore**, USA. asprs.org/Conferences/Baltimore-2013

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

30. April – 2. Mai: 8th **International Symposium on Mobile Mapping Technology 2013** in **Tainan**, Taiwan. conf.ncku.edu.tw/mmt2013/wm02.htm

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

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

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

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

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

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

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

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

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

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

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

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

Jahresübersicht 2012

Vorstand der DGPF

Präsident

Prof. Dr. rer. nat. Thomas H. Kolbe
Technische Universität München, Institut für Geodäsie, GIS und Landmanagement, Lehrstuhl für Geoinformatik, Arcisstraße 21, 80333 München, Germany, Tel.: +49-89-289-23888, Fax: +49-89-289-22878, e-mail: thomas.kolbe@tum.de

Vizepräsident

Prof. Dr.-Ing. Uwe Stilla
Technische Universität München, Fachgebiet Photogrammetrie und Fernerkundung, Arcisstraße 21, 80333 München, Tel.: 089-289-22671, Fax: 089-289-23202, e-mail: stilla@tum.de

Sekretär

Prof. Dr.-Ing. Eberhard Gülch
Hochschule für Technik Stuttgart, Fakultät Vermessung, Mathematik und Informatik, Schellingstr. 24, 70174 Stuttgart, Tel.: 0711-8926-2610, Fax: 0711-8926-2556, e-mail: sekretaer@dgpf.de

Schatzmeister

Dr.-Ing. Herbert Krauss
Rodenkirchener Straße 47, 50997 Köln, Tel.: 02233-22514, Fax: 032 222-427178, e-mail: mh.krauss@t-online.de

Hauptschriftleiter

Prof. Dr.-Ing. Wolfgang Kresse
Hochschule Neubrandenburg, Fachbereich Landschaftsarchitektur, Geoinformatik, Geodäsie und Bauingenieurwesen, Brodaer Straße 2, 17033 Neubrandenburg, Tel.: 0395-5693-4106, Fax: 0395-5693-4999, e-mail: kresse@hs-nb.de

Beirat

Prof. Dr. Cornelia Glässer
Martin-Luther-Universität Halle-Wittenberg, Institut für Geographie, Von-Seckendorff-Platz 4, 06120 Halle, Tel.: 0345-55-260 20, Fax: 0345-55-271 68, e-mail: cornelia.glaesser@geo.uni-halle.de

Beirat

Dr. rer. nat. Klaus-Ulrich Komp
EFTAS Fernerkundung, Technologietransfer GmbH, Oststraße 2–18, 48145 Münster, Tel.: 0251-1330-70, Fax: 0251-1330-733, e-mail: klaus.komp@eftas.com

Beirat

Prof. Dr.-Ing. habil. Hans-Gerd Maas
Technische Universität Dresden, Professur für Photogrammetrie, Helmholtzstraße 10, 01062 Dresden, Tel.: 0351-463-32859, Fax: 0351-463-37266, e-mail: hans-gerd.maas@tu-dresden.de

Beirat

Prof. Dr.-Ing. Monika Sester
Leibniz Universität Hannover, Institut für Kartographie und Geoinformatik (ikg), Appelstraße 9a, 30167 Hannover, Tel.: 0511-762-3588, Fax: 0511-762-2780, e-mail: monika.sester@ikg.uni-hannover.de

Beirat

Dr.-Ing. Eckhardt Seyfert
Landesvermessung und Geobasisinformation Brandenburg, Heinrich-Mann-Allee 103, 14473 Potsdam, Tel.: 0331-8844-506, Fax: 0331-8844-126, e-mail: eckhardt.seyfert@geobasis-bb.de

Ehrenmitglieder der DGPF

Prof. Dr.-Ing. Friedrich Ackermann, Stuttgart
Prof. Dr.-Ing. Wolfgang Förstner, Bonn
Prof. Gerd Hildebrandt, Freiburg
Dr.-Ing. Otto Hofmann, Brunnthal
Prof. Dr.-Ing. Gottfried Konecny, Hannover
Direktor Fritz Erich Krause, Münster
Dipl.-Ing. Horst Schöler, Stadtsteinach
Prof. Dr.-Ing. Klaus Szangolies, Jena

In diesem Jahr verstarb Prof. Dr.-Ing. Heinz Draheim, Karlsbad

Arbeitskreise der DGPF

Aus- und Weiterbildung

Dr.-Ing. Görres Grenzdörffer
Universität Rostock, Agrar- und Umweltwissenschaftliche Fakultät, Bereich: Professur Geodäsie und Geoinformatik, Anschrift: Justus-v.-Liebig-Weg 6, 18051 Rostock, Tel.: 0381-498-3206, Fax: 0381-498-3202, e-mail: goerres.grenzdoerffer@uni-rostock.de

Bildanalyse und Bildverstehen

Dr.-Ing. Uwe Weidner
Karlsruher Institut für Technologie, Institut für Photogrammetrie und Fernerkundung, Englerstraße 7, 76131 Karlsruhe, Tel.: 0721-608-43945, Fax: 0721-608-48450, e-mail: uwe.weidner@kit.edu

Geoinformatik

Prof. Dr. rer. nat. habil. Jürgen Döllner
Hasso-Plattner-Institut für Softwaresystemtechnik, Computer Graphics Systems, Anschrift: Prof.-Dr.-Helmert-Straße 2-3, 14482 Potsdam, Tel.: 0331-5509-170, Fax: 0331-5509-172, e-mail: doellner@hpi.uni-potsdam.de

Hyperspektrale Fernerkundung

Dr. Andràs Jung
Institut für Lasertechnologien in der Medizin und Messtechnik, an der Universität Ulm, Helmholtzstr. 12, 89081 Ulm, Tel.: 0731-1429997, Fax: 0731-1429442, e-mail: info@andrasjung.de

Radarfernerkundung und Flugzeuglaserscanning

Prof. Dr.-Ing. Uwe Sörgel
Leibniz Universität Hannover, Institut für Photogrammetrie und GeoInformation, Nienburger Straße 1, 30167 Hannover, Tel.: 0511-762-2981, Fax: 0511-762-2483, e-mail: soergel@ipi.uni-hannover.de, www.ipi.uni-hannover.de/us.html

Auswertung von Fernerkundungsdaten

Dr. habil. Horst Weichelt
RapidEye AG, Molkenmarkt 30, 14776 Brandenburg an der Havel, Tel.: 03381-8904-311, Fax: 03381-8904-101, e-mail: dgpf-akfe@h-weichelt.de

Standardisierung und Qualitätssicherung

Dipl.-Ing. Sven Baltrusch
Landesamt für Innere Verwaltung Mecklenburg-Vorpommern, Amt für Geoinformation, Vermessung und Katasterwesen, Lübecker Straße 289, 19059 Schwerin, Tel.: 0385-588-56322, e-mail: sven.baltrusch@laiv-mv.de

Optische 3D-Messtechnik

Prof. Dipl.-Ing. Thomas Kersten
HafenCity Universität Hamburg, Labor für Photogrammetrie & Laserscanning, Hebebrandstraße 1, 22297 Hamburg, Tel.: 040-428-27 5343, Fax: 040-428-27 5359, e-mail: thomas.kersten@hcu-hamburg.de, www.hcu-hamburg.de/geomatik/kersten

Fernerkundung in der Geologie

Dr. Hans-Ulrich Wetzel
GeoForschungsZentrum Potsdam (GFZ), PB 1.5, Telegrafenberg, 14473 Potsdam, Tel.: 0331-288-1194, Fax: 0331-288-1192, e-mail: wetz@gfz-potsdam.de

Sensoren und Plattformen

Prof. Dr. Norbert Haala
Universität Stuttgart, Institut für Photogrammetrie, Geschwister-Scholl-Straße 24d, 70174 Stuttgart, Tel.: 0711-685-83383, Fax: 0711-685-83297, email: norbert.haala@ifp.uni-stuttgart.de

3D-Stadtmodelle

Dipl.-Ing. Bettina Petzold
Hessisches Landesamt für Bodenmanagement und Geoinformation, Schaperstraße 16, 65195 Wiesbaden, Tel.: 0611-535-5372, Fax: 0611-535-5351, e-mail: bettina.petzold@hvbg.hessen.de, www.hvbg.hessen.de

Dipl.-Ing. Ekkehard Matthias
Freie und Hansestadt Hamburg, Landesbetrieb Geoinformation und Vermessung, Sachsenkamp 4, 20097 Hamburg, Tel.: 040-428-26-5750, Fax: 040-428-26-5966, e-mail: ekkehard.matthias@gv.hamburg.de, www.geoinfo.hamburg.de

Berichterstatter für ISPRS und CIPA

Kommission I – Image Data Acquisition - Sensors and Platforms

Dr.-Ing. Franz Kurz
D-82230 Oberpfaffenhofen
e-mail: franz.kurz@dlr.de

Kommission II – Theory and Concepts of Spatial Information Science

Prof. Dr.-Ing. Monika Sester
D-30167 Hannover
e-mail: monika.sester@ikg.uni-hannover.de

Kommission III – Photogrammetric Computer Vision and Image Analysis

Prof. Dr.-Ing. Stefan Hinz
D-76128 Karlsruhe
e-mail: stefan.hinz@kit.edu

Kommission IV – Geodatabases and Digital Mapping

Dr.-Ing. Volker Walter
D-70174 Stuttgart
e-mail: volker.walter@ifp.uni-stuttgart.de

Kommission V – Close-Range Sensing: Analysis and Applications

Dr.-Ing. Danilo Schneider
D-01062 Dresden
e-mail: danilo.schneider@tu-dresden.de

Kommission VI – Education and Outreach

Dipl.-Inf. Gerhard König
D-10623 Berlin
e-mail: gerhard.koenig@tu-berlin.de

Kommission VII – Thematic Processing, Modeling and Analysis of Remotely Sensed Data

Dr.-Ing. Uwe Weidner
D-76128 Karlsruhe
e-mail: weidner@ipf.uni-karlsruhe.de

Kommission VIII – Remote Sensing Applications and Policies

Prof. Dr. Irmgard Niemeyer
D-52425 Jülich
e-mail: i.niemeyer@fz-juelich.de

CIPA – Internationales Komitee für Architekturphotogrammetrie

Prof. Dr.-Ing. Michael Scherer
D-44780 Bochum, e-mail: michael.scherer@ruhr-uni-bochum.de

Gutachter der PFG im Jahr 2012

Gute und unabhängige Gutachten garantieren die Qualität einer wissenschaftlichen Zeitschrift und sind absolut notwendig für ihre internationale Anerkennung und die Aufnahme in einen Citation Index. Deshalb bedanken sich die Schriftleiter der PFG bei den Gutachterinnen und Gutachtern des Jahres 2012 ganz herzlich. Sie haben neben dem Editorial Board wesentlich zur positiven Entwicklung der PFG beigetragen. Durch die Einwerbung von deutlich mehr Artikeln als in den Vorjahren ist auch die Last der Bearbeitung gestiegen, so dass viele zwei oder sogar mehr Gutachten beigesteuert haben. Nochmals ein ganz herzliches Dankeschön an:

- Hanno Ackermann, Hannover
- Pablo d'Angelo, Oberpfaffenhofen
- Michael Arens, Ettlingen
- Costas Armenakis, York, Ontario, Kanada
- Clemens Atzberger, Wien, Österreich
- Stefan Auer, München
- Timo Balz, Wuhan, China
- Susanne Becker, Stuttgart
- Jan Böhm, London, Großbritannien
- Patrick Erik Bradley, Karlsruhe
- Vincent Brandou, Rennes, Frankreich
- Claus Brenner, Hannover
- Christian Briese, Wien, Österreich
- Dimitri Bulatov, Ettlingen
- Florian Burkert, München
- Matthias Butenuth, München
- Peter Caccetta, Floreat, Australien
- Vittorio Casella, Pavia, Italien
- Nicolas Champion, Paris, Frankreich
- Christopher Conrad, Würzburg
- Michael Cramer, Stuttgart
- Mattia Crespi, Roma, Italien
- Tobias Dahinden, Hannover
- Martin Drauschke, Neubiberg
- Enrique Dunn, Chapel Hill, North Carolina, USA
- Markus Even, Ettlingen
- Daniel Eggert, Hannover
- Henri Eisenbeiss, Zürich, Schweiz
- Wolfgang Förstner, Bonn
- Andrew Frank, Wien, Österreich
- Clive Fraser, Melbourne, Australien
- Markus Gerke, Enschede, Niederlande
- Richard Gloaguen, Freiberg
- Jens Göpfert, Hannover
- Jose Gonzalez, Albacete, Spanien
- Michael Gösele, Darmstadt
- Carsten Götz, München
- Görres Grenzdörffer, Rostock
- Gerd Gröger, Bonn
- Anne Grote, Hannover
- Eberhard Gülch, Stuttgart
- Norbert Haala, Stuttgart
- Ayman Habib, Calgary, Canada
- Ronnie Hänsch, Berlin
- Jan-Henrik Haunert, Würzburg
- James Head, Providence, Rhode Island, USA
- Christian Heipke, Hannover
- Wieke Heldens, Oberpfaffenhofen
- Olaf Hellwich, Berlin
- Petra Helmholz, Hannover
- Tobias Hillmann, Neubrandenburg
- Stefan Hinz, Karlsruhe
- Heiko Hirschmüller, Oberpfaffenhofen
- Thorsten Hoberg, Hannover
- Volker Hochschild, Tübingen
- Joachim Höhle, Aalborg, Dänemark
- Angela Hof, Bochum
- Eija Honkavaara, Masala, Finnland
- Michelle Horta, São Carlo, Brasilien
- Hai Huang, Hannover
- Juha Hyyppä, Helsinki, Finnland
- Karsten Jacobsen, Hannover
- Josef Jansa, Wien, Österreich
- Jarek Jasiewicz, Poznań, Polen
- Boris Jutzi, Karlsruhe
- Helmut Kager, Wien, Österreich
- Martin Kappas, Göttingen
- Willi Karel, Wien, Österreich
- Viktor Kaufmann, Graz, Österreich
- Thomas Kersten, Hamburg
- Barend Köbben, Enschede, Niederlande
- Tomasz Kogut, Koszalin, Polen
- Thomas Krauss, Oberpfaffenhofen
- Jens Kremer, Kreuztal
- Peter Krzystek, München
- Werner Kuhn, Münster
- Richard Ladstädter, Graz, Österreich
- Tobia Lakes, Berlin

- Derek Lichti, Calgary, Kanada
- Ralf Ludwig, München
- Thomas Luhmann, Oldenburg
- Hans-Gerd Maas, Dresden
- Ioannis Manakos, Thessaloniki, Griechenland
- Lauri Markelin, Masala, Finnland
- Lucas Martinez, Castelldefels, Spanien
- Helmut Mayer, Neubiberg
- Chris McGlone, McLean, Virginia, USA
- Jochen Meidow, Ettlingen
- Werner Melle, Kahla
- Eckart Michaelsen, Ettlingen
- Ulrich Michel, Heidelberg
- Matthias Möller, Berlin
- Jeremy Morley, Nottingham, Großbritannien
- Daniel Muhle, Hannover
- Stephan Nebiker, Muttenz, Schweiz
- Andreas Nüchter, Bremen
- Francesca Pieralice, Roma, Italien
- Daniela Poli, Ispra, Italien
- Jorge Recio, Valencia, Spanien
- Camillo Ressl, Wien, Österreich
- Rudolf Richter, Oberpfaffenhofen
- Petri Rönnholm, Helsinki, Finnland
- Ribana Roscher, Bonn
- Uwe Schäffer, Hannover
- Jochen Schiewe, Hamburg
- Konrad Schindler, Zürich, Schweiz
- Benno Schmidt, Bochum
- Florian Schmidt, Karlsruhe
- Ralph Schmidt, Hannover
- Andreas Schmitt, Bonn
- Michael Schmitt, München
- Danilo Schneider, Dresden
- Maria v. Schönermark, Oberpfaffenhofen
- Simon Schuffert, Karlsruhe
- Alexander Schunert, Hannover
- Karl Segl, Potsdam
- Jie Shan, West Lafayette, Indiana, USA
- Beril Sirmacek, Delft, Niederlande
- Jan Skaloud, Lausanne, Schweiz
- Martin Smith, Nottingham, Großbritannien
- Uwe Sörgel, Hannover
- Hélène Sportouche, Paris, Frankreich
- Bahman Soheilian, Paris, Frankreich
- Gunho Sohn, Toronto, Kanada
- Uwe Stilla, München
- Christoph Stiller, Karlsruhe
- David Tarboton, Logan, Utah, USA
- Hannes Taubenböck, Oberpfaffenhofen
- Rüdiger Tauch, Potsdam
- Hans-Peter Thamm, Berlin
- Barbara Theillen-Willige, Stockach
- Antje Thiele, Karlsruhe
- Rene Wackrow, Loughborough, Großbritannien
- Oliver Wallgrün, University Park, Pennsylvania, USA
- Björn Waske, Bonn
- Jan Wegner, Zürich, Schweiz
- Uwe Weidner, Karlsruhe
- Martin Weinmann, Karlsruhe
- Thomas Weser, München
- Jochen Willneff, Weilheim
- Nils Wolf, Bochum
- Bernhard Wrobel, Meckenheim
- Michael Yang, Hannover
- Chunsun Zhang, Melbourne, Australien
- Yongjun Zhang, Wuhan, China
- Marcel Ziems, Hannover
- Francesco Zucca, Pavia, Italien

Korporative Mitglieder

Firmen

AEROWEST GmbH
AICON 3D Systems GmbH
aphos Leipzig AG
Becker GeoInfo GmbH
Bernhard Harzer Verlag GmbH
Blom Deutschland GmbH
Brockmann Consult GmbH
bsf swissphoto GmbH
Büro Immekus
CGI Systems GmbH
con terra GmbH
DB Netz AG
DELPHI IMM GmbH
Deutsches Bergbau-Museum
EFTAS Fernerkundung Technologietransfer GmbH
ESG Elektroniksystem- und Logistik-GmbH
Esri Deutschland GmbH
EUROPEAN SPACE IMAGING
Eurosense GmbH
fokus GmbH
g.on experience gmbh
GAF GmbH
GeoCart Herten GmbH
GeoContent GmbH
Geoinform. & Photogr. Engin. Dr. Kruck & Co. GbR
geoplana Ingenieurgesellschaft mbH
GEOSYSTEMS GmbH
GGS - Büro für Geotechnik, Geoinformatik, Service
Hansa Luftbild AG
IGI - Ingenieur-Gesellschaft für Interfaces mbH
ILV Ingenieurbüro für Luftbildauswertung und Vermessung
Imetric 3D GmbH
Infoterra GmbH
INVERS - Industrievermessung & Systeme
ITT Visual Information Solutions Germany
J. Linsinger ZT-GmbH
Jena-Optronik GmbH
Leica Geosystems GmbH
Luftbilddatenbank-Würzburg
Messbildstelle GmbH
Microsoft Photogrammetry
MILAN Geoservice GmbH
M.O.S.S. Computer Grafik Systeme GmbH
PHOENICS GmbH
PMS - Photo Mess Systeme AG
RIEGL Laser Measurement Systems GmbH
RWE Power AG, Geobasisdaten/Markscheidewesen
technet GmbH
TRIGIS Vermessung + Geoinformatik GmbH
Trimble Germany GmbH
trimetric 3D Service GmbH
Wichmann, VDE Verlag GmbH
Z/I Imaging Ltd.

Behörden

Amt für Geoinformationswesen der Bundeswehr
Bayerische Landesanstalt für Wald und Forstwirtschaft
Bundesamt für Kartographie und Geodäsie
Bundesmininisterium für Ernährung, Landwirtschaft und Verbraucherschutz
Hessisches LA für Bodenmanagement und Geoinformation
Innenministerium NRW, Gruppe Vermessungswesen
Institut für Umwelt- und Zukunftsforschung
LA für Geoinformation und Landentwicklung, BW
LA für Vermessung und Geoinformation, Bayern
LB Geoinformation und Vermessung, Hamburg
LB für Küstenschutz, Nationalpark und Meeresschutz, SH
Landesvermessung und Geobasisinformation Niedersachsen
Märkischer Kreis, Vermessungs- und Katasteramt
Regierungspräsident Tübingen, Abt. 8 Forstdirektion
Regionalverband Ruhr
Staatsbetrieb Sachsenforst Pirna
Stadt Bocholt, Fachbereich 31
Stadt Düsseldorf, Vermessungs- und Katasteramt
Stadt Köln, Amt für Liegenschaften, Vermessung und Kataster
Stadt Wuppertal, Vermessung, Katasteramt und Geodaten
Thüringer LA für Vermessung und Geoinformation

Hochschulen

BTU Cottbus, Lehrstuhl für Vermessungskunde
FH Frankfurt a.M., FB 1, Studiengang Geoinformation
FH Mainz, Institut für Raumbezogene Informations- und Messtechnik
Jade Hochschule, Institut für Angewandte Photogrammetrie und Geoinformatik
HCU HafenCity Universität Hamburg, Geomatik
HfT Stuttgart, Vermessung und Geoinformatik
HS Bochum, FB Vermessung und Geoinformatik
HS Karlsruhe, Fakultät für Geomatik
HTW Dresden, FB Vermessungswesen/Kartographie
LUH Hannover, Institut für Kartographie und Geoinformatik
LUH Hannover, Institut für Photogrammetrie und Geoinformation
MLU Halle, FG Geofernerkundung
Ruhr-Uni Bochum, Geographisches Institut
RWTH Aachen, Geodätisches Institut
TU Bergak. Freiberg, Institut für Markscheidewesen und Geodäsie
TU Berlin, Computer Vision & Remote Sensing
TU Berlin, Institut für Geodäsie und Geoinformationstechnik
TU Braunschweig, Institut für Geodäsie und Photogr.
TU Clausthal, Institut für Geotechnik und Markscheidewesen
TU Darmstadt, Institut für Photogrammetrie und Kartographie
TU Dresden, Institut für Photogrammetrie und Fernerkundung
TU München, FG Photogrammetrie und Fernerkundung
TU Wien, Institut für Photogrammetrie und Fernerkundung
Uni Bonn, Institut für Photogrammetrie
Uni Göttingen, Institut für Waldinventur und Waldwachstum
Uni Heidelberg, IWR Interdisziplinäres Zentrum für Wissenschaftliches Rechnen
Uni Kassel, FB Ökologische Agrarwissenschaften
Uni Kiel, Geographisches Institut
Uni Stuttgart, Institut für Photogrammetrie
Uni Würzburg, Geographisches Institut
Uni zu Köln, Geographisches Institut

PFG – Jahrgang 2012, Heft 1 – 6
Inhaltsverzeichnis Jahrgang 2012

Editorials

Originalbeiträge, alphabetisch nach Autoren

Beiträge aus Wissenschaft und Technik

Berichte und Mitteilungen