Stereo matching of terrestrial digital photographs – an alternative for the generation of high-resolution DEMs in situations of poor data availability?

Martin MERGILI

Abstract

The paper evaluates the applicability of terrestrial digital photographs for stereo matching in order to generate high-resolution DEMs in areas with extremely pronounced relief. Images of the study slopes, located along a transect between Mendoza (Argentina) and Valparaíso (Chile), were recorded from the opposite side of the valley using a digital SLR camera. Stereo matching was performed using standard remote sensing software. The reference plane had to be tilted in order to enable a proper rectification. The resulting raster maps were adjusted in order to enhance the consistency with the reference datasets. The method, though connected to some limitations concerning the accuracy of the results, proved successful and may be a valuable tool for small-scale studies in areas with limited data availability, particularly in developing countries.

1 Background

High-resolution digital elevation models (DEMs) constitute an essential tool for a large variety of applications. Therefore an array of methods for generating such datasets has been developed in the previous years and decades (compare Tab. 1). Some of the most common approaches are:

(1) Airborne laser scanning (ACKERMANN 1999; PFEIFER 2006) comes up with spatial resolutions well below 1 m. In highly developed countries airborne laser scanning data are available for an increasing portion of the land surface. Terrestrial laser scanning is used more and more for obtaining high resolution DEMs of small areas.

(2) A common way for generating high resolution DEMs of large areas is stereo matching of high resolution satellite imagery like IKONOS (Zhang et al. 2002).

(3) For small-scale studies, matching of stereo aerial photos has widely been applied, replacing the traditional stereoscopic viewing (Gruen & Baltsavias 1986; Fig. 1). The method, being the same as for satellite imagery, is supported by standard remote sensing software. It makes use of the parallax between two images recorded from different positions as well as of the texture of the images in order to compute the position of each pixel in a reference coordinate system, based on camera geometry, control points, and the exterior orientation of the camera (optional). Würländer et al. (2004) provide a comparison of this method with laser scanning techniques.

Tab. 1: Some common methods for the generation of high-resolution DEMs.

Method	**Advantages**	**Disadvantages**
Airborne laser scanning	high accuracy, identification of vertical structures	expensive; cost-efficiently only for large areas
Terrestrial laser scanning	high accuracy, identification of vertical structures	specialized equipment and personnel
Stereo matching of high-resolution satellite imagery	cover of large areas	high cost for data
Stereo matching of aerial imagery	low cost for data, good data availability for many areas	labour-extensive for large areas, limited by the image quality, moderate accuracy

Use of the methods (1) and (2) usually implies considerable costs for data acquisition, largely constraining its application to projects with high budget. In addition, processing of data derived from laser scanning is still connected to specialized data processing tools that are now starting to be included into the standard GIS software tools like ArcGIS, GRASS (RUTZINGER et al. 2006), or PCI Geomatica.

For small-scale projects, stereo-matching of aerial images works fine in many areas, but nevertheless bears some limitations as well. Particularly in peripheral regions of developing countries imagery is not always available. Images may be of poor quality, and not all land surface types are suitable for image matching. Particularly patches with poor texture and/or particularly high or low albedo, like snow or ice, may cause serious problems, leading to large holes in the DEM.

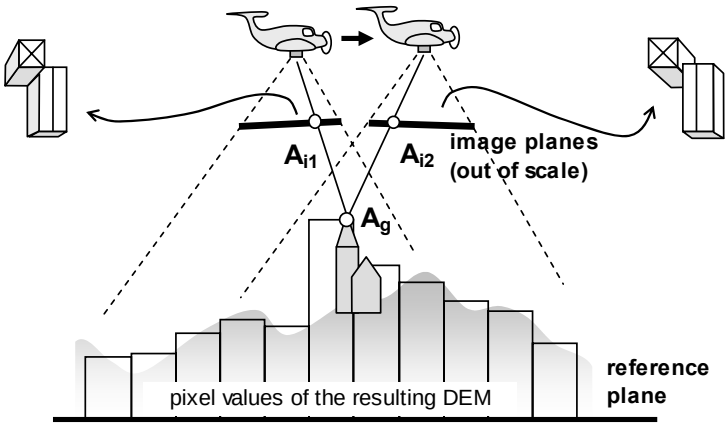

Fig. 1: Stereo matching of aerial images. Graphics modified from the PCI Geomatica Help. A_{i1} and A_{i2} represent the image coordinates of the clock tower's top, A_g its ground coordinates.

In contrast to aerial imagery, hardly any attempts to use terrestrial imagery for the extraction of high-resolution DEMs have become public up to now. This appears somewhat astonishing as

(1) the mathematical background and the methodology are almost the same as for aerial imagery,

(2) in areas with pronounced relief it may be easy to access spots with a good overview of the area of interest, particularly in the case of small-scale studies,

(3) the spatial resolution and the image quality may be much better than in the case of aerial imagery, and finally

(4) in other fields, similar methods have been successfully applied for representing the geometry of buildings, or even of human faces (GRUEN & BALTSAVIAS 1989).

The purpose of the present study is to attack this gap by evaluating the possibilities as well as the limitations connected to the generation of high-resolution DEMs by stereo-matching of terrestrial photographs.
The experiment is part of a project dealing with the prediction of debris flows based on a deterministic model, therefore requiring a DEM at an appropriate resolution (≤5 m). The study includes six slopes along a transect across the Andes between Mendoza (Argentina) and Valparaíso (Chile), following an important road corridor connecting the two countries (Fig. 2).

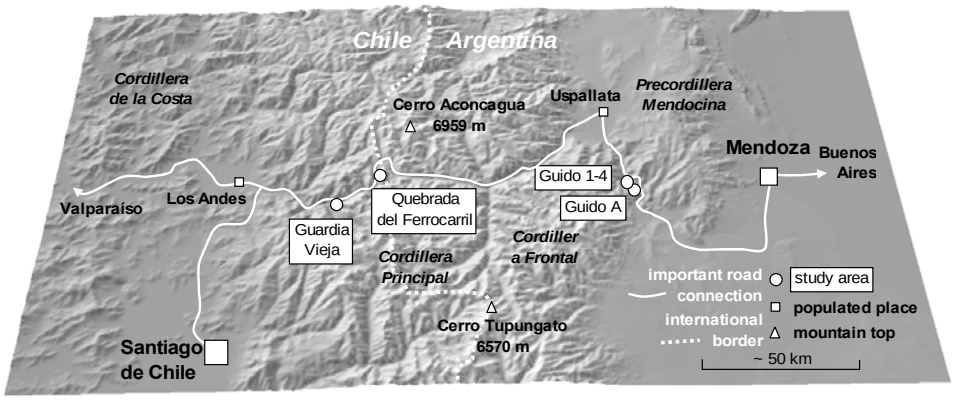

Fig. 2: Location of the study areas. Shaded relief map calculated from SRTM data.

2 Data

The stereo images were recorded using a standard digital SLR camera (Konica Minolta Dynax 5D) with a chip size of 23.5 x 15.7 mm. A focal length of 18 mm was chosen for all images, the distortion parameters of the lens were not known. The recording points were

chosen in a way that the distances between the stereo images would be sufficient for ensuring a reasonable stereo effect, but not too much in order to prevent troubles with automatic pixel matching. The positions were approximated using a standard field GPS device with an accuracy of ≤10 m. The distance to the target slopes was of a magnitude of some 100s of meters. Figure 3 illustrates one of the stereo pairs used.

Fig. 3: Example of a stereo pair of terrestrial images recorded from the slope opposite to the study area Guido 3. Source: Mergili, 10/2006.

Orthophotos and a 30 m resolution DEM were used as ground references. The orthophotos were generated from aerial imagery provided by the provincial government of Mendoza (scale 1:20,000), and by the aero-photogrammetric service of Chile (1:60,000), respectively. The 30 m DEM was purchased as final product derived from ASTER satellite imagery. Additionally, GPS control points were recorded in the study areas. Due to their unbalanced patterns (caused by the inaccessibility of parts of the slopes) and their relatively low vertical accuracy it was decided not to use them as GCPs.

3 Methods

The workflow of the study, represented in Fig. 4, may be divided into the following steps:

(1) Collecting GCPs from orthophotos and from the ASTER DEM.

(2) Definition of an appropriate reference plane. In contrast to stereo matching based on aerial imagery it was not possible just to use a simple projected coordinate system like UTM, for the reason that the rectification algorithm applied in the OrthoEngine of PCI Geomatica, which was used for the DEM extraction, is not suitable for dealing with extreme variations of the z coordinate (elevation) compared to the image size, like appearing in the terrestrial photographs. Therefore a slope-parallel reference plane had to be introduced. It was defined by a rotation α and an inclination β, related to the UTM system (Fig. 5). α and β were determined individually for each study area by minimizing the root mean square (RMS) values of the z-coordinates of the respective GCPs.

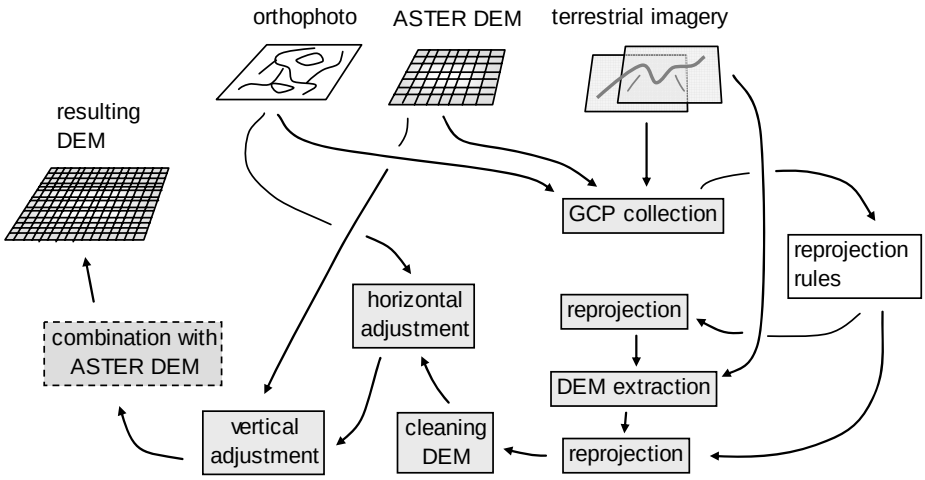

Fig. 4: Workflow of the study.

(3) Recalculating the GCPs to the new reference plane. The transformation was given by

$$x' = x\cos\alpha - y\sin\alpha \qquad \text{Eq. (1),}$$

$$y' = \left(y\cos\alpha + x\sin\alpha\right)/\cos\beta + \Delta z\sin\beta \qquad \text{Eq. (2), and}$$

$$z' = \Delta z\cos\beta \qquad \text{Eq. (3),}$$

where x, y, and z are the coordinates in the UTM reference plane, but with P_0 as origin (compare Fig. 5), x', y', and z' are the coordinates in the new reference system, α is the rotation angle, and β stands for the inclination angle (compare Fig. 5). Δz is the elevation difference between the new reference plane and the elevation of P_0 in the UTM system, given by

$$\Delta z = \left(y\cos\alpha + x\sin\alpha\right)\tan\beta \qquad \text{Eq. (4).}$$

(4) Image matching based on textural features and the parallax between the images, as mentioned above and illustrated in Fig. 1. This step is a standard functionality of PCI Geomatica and will therefore not be discussed in detail in this place.

(5) Recomputation of the DEM to the UTM reference system. This was done by converting the result of (4) into a lattice, and applying Eq. (1) to (4) to the coordinates, inverting α and β. The resulting irregularly spaced cluster of points was recalculated to a raster map, using a natural neighbours algorithm. Artificial tops and sinks within the DEM were then removed manually (referred to as "cleaning DEM" in Fig. 4).

(6) Horizontal adjustment. A geometrical correction was necessary in order to synchronize the new DEM datasets with the underlying orthophoto and the ASTER DEM. Distinct topographic features clearly recognizable in both the new DEM and the orthophoto were used as control points.

(7) Vertical adjustment. The raster was subtracted from the ASTER DEM in order to determine large-scale distortions in the stereo DEM. A correction raster, as an extremely smoothed representation of the difference between the datasets, was interpolated and applied for the adjustment of the stereo DEM. Thus the small-scale patterns were conserved while fitting the large-scale patterns to the ASTER DEM, removing distortions and allowing for a proper combination of the two datasets.

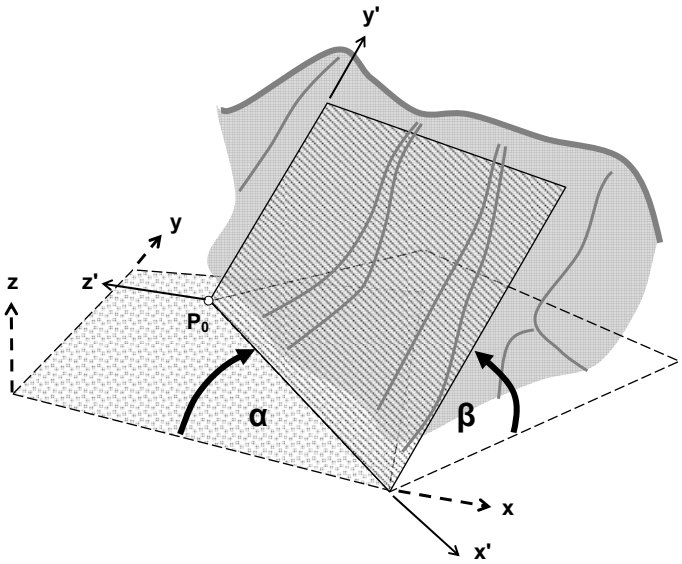

Fig. 5: Reference plane introduced for rectification and DEM generation.

4 Results

Image correlation worked well for some of the study areas, for some others holes of considerable sizes – primarily caused by snow patches and cloudy conditions at the time of the recording – led to a lower quality level. The results will be discussed in more detail for the study areas Guido A and Guido 3 (compare Fig. 2). The stereo DEMs for these areas showed no holes in the central areas of the DEMs, and the matching scores were above 80 % over the majority of the surfaces. A spatial resolution of 2 m was chosen for Guido A, 1 m for Guido 3. A few artificial peaks and sinks (1 % of the surface of Guido A, 3.5 % of the surface of Guido 3) had to be removed in order to derive visually clean DEMs for both study areas (Fig. 6).

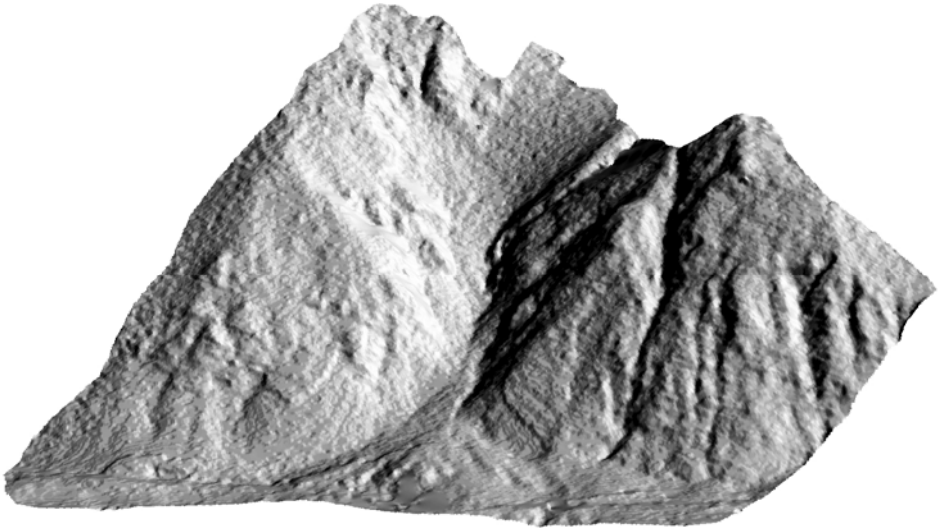

Fig. 6: DEM (resolution 2m) derived from stereo images for the study area Guido A, after removing peaks and holes. The international road is well recognizable at the lower edge of the DEM. Its length represented in the dataset is about 800 m, the elevation difference 700 m.

A horizontal adjustment was required for both stereo DEMs in order to bring them in cover with the orthophoto and the ASTER DEM. For Guido A, the overlay with the ASTER DEM showed considerable vertical deviations between the two datasets. More than 25 % of the pixels were more than 40 m off, only about 35 % less than 20 m (Fig. 7). For Guido 3, the datasets corresponded better, with more than 50 % of the pixels being less than 20 m off.

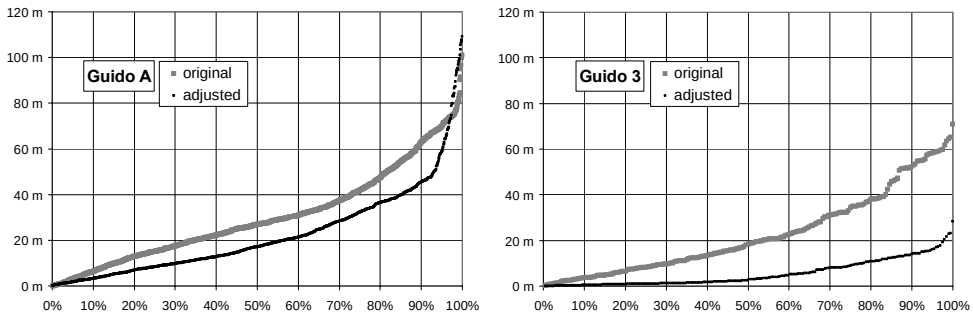

Fig. 7: Deviation between the stereo DEM and the ASTER DEM for the study areas Guido A (left) and Guido 3 (right), before and after the vertical adjustment. The absolute deviation between the two datasets, plotted along the ordinate, falls below the indicated value at the corresponding percentage of pixels (abscissa).

Vertical adjustment of the stereo DEMs was performed for both areas, decreasing the deviation from the ASTER DEM (compare Fig. 7). However, maintaining the topographical structures was given priority to a perfect fit, considering the relatively large inaccuracies to be expected in the ASTER DEM itself (10s of meters). The offset error was only moderately reduced for the Guido A DEM, leading to slightly less than 60 % of the pixels being less than 20 m off. In contrast, the adjustment of the Guido 3 DEM led to a considerable reduction of the deviation, with about 75 % of the pixels less than 10 m off and more than 95 % of the pixels less than 20 m off.

A lot of emphasis was put on the synchronization of the connection lines of the stereo DEM with the ASTER DEM, in order to allow for a hydrologically, geomorphologically, and visually consistent combination of the two datasets. Fig. 8 illustrates the combined DEMs for the study areas Guido A and Guido 3, enabling an estimation of the hydrological input to the areas of interest. The hydrological consistency of the composite DEMs was tested using the standard parameters flow direction and flow accumulation. It proved successful for both of the study areas.

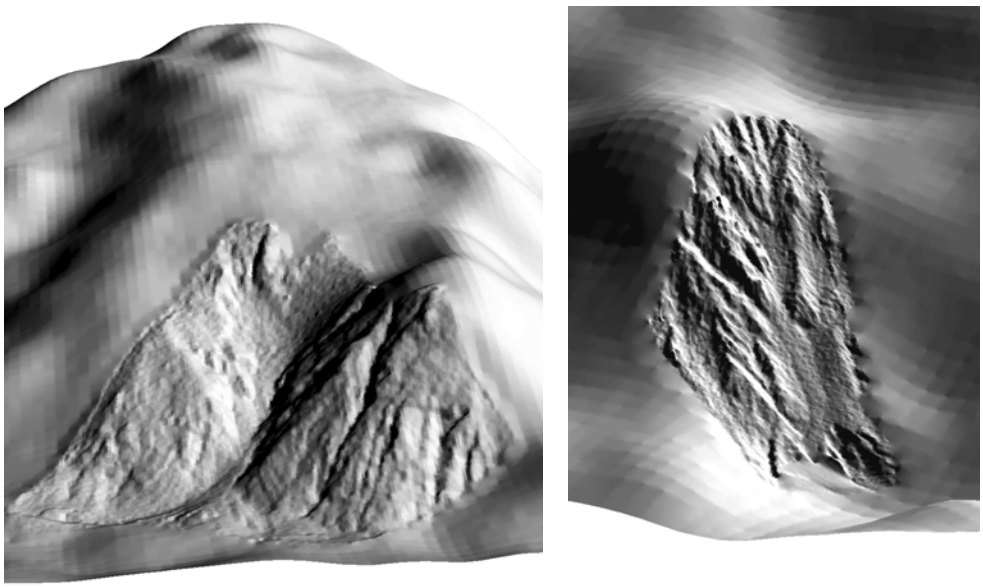

Fig. 8: Combined stereo and ASTER DEMs for the study areas Guido A (left side, resolution 5 m) and Guido 3 (right side, resolution 2 m, width of the stereo DEM about 200 m).

5 Discussion

The method presented above does not lead to high-precision results – it rather provides a more or less distorted image of the reality. Therefore it is required to know about the uncertainties and limitations, but also about the potential of DEMs derived from terrestrial stereo images.

The absolute accuracy of the resulting DEM product is limited by the accuracy of the GCPs used for rectification. The orthophotos from which the GCPs were collected had a sufficient resolution and accuracy, and its features were clearly identifiable in the terrestrial imagery. However, the elevation of the GCPs was a major factor of uncertainty as only an ASTER DEM with a resolution of 30 m was available over the entire surface of the study areas. Unfortunately it was not possible to use GPS records as GCPs, as (1) they were only available for accessible areas and mixing them with other GCPs would have introduced inconsistencies, (2) their horizontal and vertical accuracy was limited, and (3) it was difficult in the field to recognize appropriate GCPs.

The situation would prove even more difficult in areas without orthophotos available. For such cases, panchromatic LANDSAT images (15 m resolution) or SPOT images would be an alternative. However, a low horizontal accuracy and potential problems with the identification of proper GCPs would be the consequence.

The accuracy of the resulting DEM products was not only limited by the accuracy of the GCPs, but also by the constraints of the rectification algorithm. All standard remote sensing software products, including the PCI Geomatica OrthoEngine, have been developed for the processing of satellite and aerial imagery, which both

(1) represent the earth surface more or less from straight above, with almost perpendicular x- and y- ground coordinates in the images, and an extremely short z-coordinate, and

(2) cover a horizontal surface exceeding the elevation by far over one scene.

Terrestrial digital photographs as used in the present study, in contrast, lack these geometric characteristics. The ground coordinates are arranged in an irregular pattern within the images, and the elevation variation over the image may be of the same magnitude as its horizontal extent. The definition of an appropriate, slope-parallel coordinate system can partly overcome this problem and make this method possible, but still, the $z/(x+y)$ - relationship is comparable to standard aerial imagery in extremely mountainous terrain, leading to a reduced level of accuracy in the resulting DEM. Thorough post-processing (horizontal and vertical adjustment) can remove part of the error, but inaccuracies additional to those caused by the reference data always remain.

6 Conclusions

The experiences from the experiment described in this paper allow for the following conclusions:

(1) Stereo-matching of terrestrial digital photographs constitutes a considerable potential for the generation of high-resolution DEMs and can be an extremely valuable tool, particularly for small-scale studies in areas where no data of higher accuracy is avail-

able, for example in peripheral areas of developing countries. However, the method is not universally applicable. The study area has to be visible from at least two accessible viewpoints for recording the images, and the images have to be recorded under clear weather conditions (no shadows of clouds).

(2) Due to the narrow field of view characterizing terrestrial images compared to aerial images, for many research purposes the DEM has to be combined with other high resolution DEMs, e.g. derived from aerial imagery, or with medium resolution DEMs like ASTER-derived datasets. The combination has to be done with caution in order to ensure the consistency of the dataset.

(3) The expectable errors in derived products like slope, curvature, or hydrological characteristics are hard to quantify, but taking the quality of the DEMs for the study areas Guido A and Guido 3 as reference, it is to be expected that for many research purposes – particularly if the DEM is used for modeling processes like landslides or runoff – the impacts of errors and inaccuracies in other datasets (substrate parameters, vegetation patterns, precipitation input) usually exceed the impact of the inaccuracies in the DEM. In contrast, the method is hardly applicable if high-precision datasets are required, for instance for the detection of slow movements.

References

ACKERMANN, F. (1999): *Airborne laser scanning - present status and future expectations*. ISPRS Journal of Photogrammetry & Remote Sensing 54: 64-67.

GRUEN, A. & BALTSAVIAS, E. (1987): *High-Precision Image Matching for Digital Terrain Model Generation*. Photogrammetria 42: 97-112.

GRUEN, A. & BALTSAVIAS, E. (1989): Automatic 3-D *Measurement of Human Faces with CCD-Cameras*. Paper presented at the Biostereometrics ’88, 14.-17. November, Basel. Proc. of SPIE 1030: 106-116.

PFEIFER, N. (2006): *Airborne Laser Scanning - Examples for operational Application in Hydrology*. Geophysical Research Abstracts 8, 10790.

RUTZINGER, M., HÖFLE, B., GEIST, T. & STÖTTER, J. (2006): *Object-based building detection based on airborne laser scanning data within GRASS GIS environment.* Proceedings UDMS: Urban data Management Symposium, Aalborg. CD-ROM.

WÜRLÄNDER, R., EDER, K. & GEIST, T. (2004): *High quality DEMs for glacier monitoring - image matching versus laser scanning*. International archives of Photogrammetry, Remote Sensing and spatial information science 35, B7: 753 - 758.

ZHANG, L., PATERSKI, M. & BALTSAVIAS, E. (2002): *Matching of Ikonos Stereo and Multitemporal GEO Images for DSM Generation*. Proc. Map Asia 2002, Asian Conference on GIS, GPS, Aerial Photography and Remote Sensing, 7.-9. August, Bangkok. CD-ROM.