AUTOMATIC DSM GENERATION FROM LINEAR ARRAY IMAGERY DATA

Zhang Li, Armin Gruen

Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology Zurich
ETH–Hoenggerberg; CH-8093 Zurich, Switzerland
Tel.: +41-1-633 31 57, Fax: +41-1-633 11 01
E-mail: <zhangl><agruen>@geod.baug.ethz.ch

Commission III, WG III/2

KEY WORDS: Linear Array Imagery, Image Matching, DSM

ABSTRACT:
CCD linear array sensors are widely used to acquire panchromatic and multispectral imagery for photogrammetric and remote sensing applications. The processing of this kind of images provides a challenge for algorithmic redesign and this opens the possibility to reconsider and improve many photogrammetric processing components. In addition, the basic capabilities of image matching techniques have so far not been fully utilized yet. This paper presents a matching procedure for automatic DSM generation from linear array imagery data. It can provide dense, precise and reliable results. The method uses a coarse-to-fine hierarchical solution with an effective combination of several image matching algorithms and automatic quality control. The DSMs are generated by combination of matching results of feature points, grid points and edges. Finally, a modified multi-photo geometrically constrained (MPGC) matching algorithm is employed to achieve sub-pixel accuracy for all the matched features with multi-image or multi-strip image data.
The proposed approach in this paper has been applied to different areas with varying textures and terrain types. The accuracy tests are based on the comparison between the high quality DEMs / DSMs derived from airborne Laser Scanner or manual measurements and the automatic extracted DSMs. Results with STARIMAGER, IKONOS and SPOT5 HRS images are reported. We demonstrate with these experiments that our approach leads to good results.

1. INTRODUCTION

In recent years, CCD linear array sensors are widely used to acquire panchromatic and multispectral imagery in pushbroom mode for photogrammetric and remote sensing applications. Linear scanners are carried on aircraft (e.g. ADS40), helicopter (e.g. STARIMAGER) or spacecraft (e.g. IKONOS) and allow for photogrammetric mapping at different scales.
Spaceborne optical sensors like SPOT, IKONOS, and QuickBird provide not only for high-resolution (0.6 – 5.0 m) and multi-spectral data, but also for the capability of stereo mapping. The related sensors are all using linear array CCD technology for image acquisition and are equipped with high quality orbit position and attitude determination devices like GPS and IMU systems.
Progress in the development of airborne linear array imaging system has also been made in the last decade. These systems use the three-line-scanner concept and provide for high resolution (0.5 – 0.03 m) panchromatic and multispectral image data with triplet overlap and along-track base direction. In the year 2000, Starlabo Corporation, Tokyo designed a new airborne digital imaging system, the Three-Line-Scanner (TLS) system (now called STARIMAGER (SI)), jointly with the Institute of Industrial Science, University of Tokyo (Murai, Matsumoto, 2000). The first generation camera STARIMAGER-100 (SI-100) contains three parallel one-dimensional CCD focal plane arrays, with 10200 pixels of 7µm each. Starlabo is currently developing a new generation camera system SI-200. This comes with an improved lens system and with 10 CCD arrays on the focal plane (3 × 3 work in RGB mode, 1 CCD array works in infrared mode). Each CCD array consists of 14 404 pixels at 5µm size. The system produces seamless high-resolution images (3 - 10 cm footprint on the ground) with three viewing directions (forward, nadir and backward). For the SI sensor and imaging parameters see Gruen, Zhang, 2002.
The processing of this kind of images provides a challenge for algorithmic redesign and this opens the possibility to reconsider and improve many photogrammetric processing components, like image enhancement, multi-channel color processing, triangulation, orthophoto and DEM generation and object extraction. We have recently developed a full suite of new algorithms and software system for the precision processing of this kind of data.
In this paper, we put particular emphasis on the automatic generation of DSMs. Originally we developed a matching approach and the related software "SI-Matcher" for multi-image processing of the very high-resolution SI images (Gruen, Zhang, 2003). Now this matching procedure has been extended and has the ability to process other linear array images as well. We will briefly report about the basic considerations for our procedure. Then we will address the key algorithms. We will give experimental results from the processing of SI, IKONOS and SPOT5 HRS images.

2. MATCHING CONSIDERATIONS

Automatic DEM/DSM generation through image matching has gained much attention in the past years. A wide variety of approaches have been developed, and automatic DEM generation packages are in the meanwhile commercially available on several digital photogrammetric workstations. Although the algorithms and the matching strategies used may differ from each other, the accuracy performance and the problems encountered are very similar in the major systems and the performance of commercial image matchers does by far not live up to the standards set by manual measurements (Gruen et al., 2000). The main problems in DEM/DSM generation are encountered with

(a) Little or no texture
(b) Distinct object discontinuities
(c) Local object patch is no planar face
(d) Repetitive objects
(e) Occlusions
(f) Moving objects, incl. shadows
(g) Multi-layered and transparent objects
(h) Radiometric artifacts like specular reflections and others
(i) Reduction from DSM to DEM

The degree to which these problems will influence the matching results is imagescale-dependent. A DSM derived from 5 m pixelsize SPOT5 HRS images or 1 m pixelsize IKONOS images will be relatively better than one derived from 5 cm pixelsize SI images. To extract DSMs from very high-resolution aerial images, we should take into account the occlusions, the surface discontinuities such as man-made objects and trees, large areas with little or even no texture, repetitive patterns, etc.
On the other hand, linear array imagery provides for new characteristics and possibilities for image matching:

- It has the ability to provide 16 bit (effectively 12 bit) images, which should reduce the number of mismatches even in dark shadow areas.

● It has the ability to provide multiple images with multiple channels. Thus it enables the multi-image matching approach, which leads to a reduction of problems caused by occlusions, multiple solutions, surface discontinuities and results in higher measurement accuracy through the intersection of more than two image rays.
● It has the ability to provide for relatively precise orientation elements that can be used to enforce geometric constraints and restrict the search space along quasi-epipolar lines.
●The nearly parallel projection in along-track direction causes less occlusion on the nadir-viewing images.

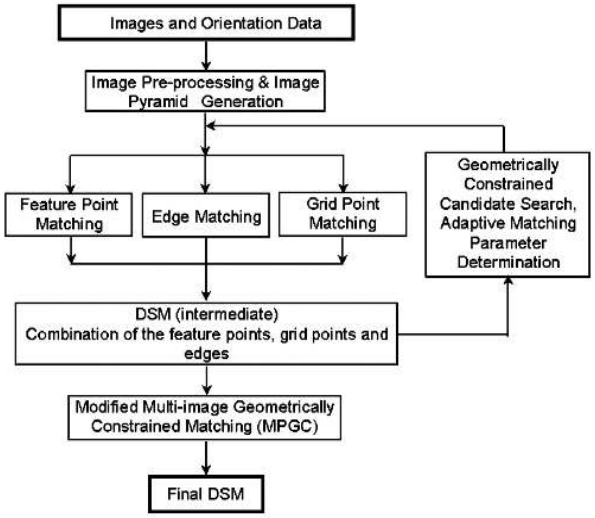

Figure 1: Workflow of our image matching procedure

Among the known matching techniques and algorithms, area-based (ABM) and feature-based matching (FBM) are the two main ones applied to automatic DSM generation in general. ABM and FBM have both advantages and disadvantages with respect to the problems presented above. The key to successful matching is an appropriate matching strategy, making use of all available and explicit knowledge concerning sensor model, network structure and image content.
Our matching approach is a hybrid method that combines ABM and FBM. It aims to generate DSMs by attacking the problems (a)-(f) mentioned above. It uses a coarse-to-fine hierarchical solution with a combination of several image matching algorithms and automatic quality control. Figure 1 shows the workflow of our matching procedure. After the image pre-processing and production of the image pyramid, the matches of three kinds of features, i.e. feature points, grid points and edges on the original images are finally found progressively starting from the low-density features on the images with low resolution. A triangular irregular network (TIN) based DSM is constructed from the matched features on each level of the pyramid, which in turn is used in the subsequent pyramid level for the approximations and adaptive computation of the matching parameters. Finally the modified MPGC matching is used to achieve more precise matches for all the matched features and identify some false matches. In the MPGC procedure, multiple strip image data can be introduced and combined. More details of our approach will be provided in the next paragraph.

3. THE MATCHING APPROACH

3.1 Image Preprocessing

In order to reduce the effects of radiometric problems like strong bright and dark regions and to optimise the images for subsequent feature extraction and image matching, a pre-processing method, which combines an edge-preserving smoothing filter and the Wallis filter. First, the edge preserving smoothing filter proposed by Saint-Marc et al., 1991 was applied to reduce the noise, while sharpening edges and preserving even fine detail such as corners and line endpoints. Next, the Wallis filter, which strongly enhances already existing texture patterns, is applied. The examples of Figure 2 indicates that even in shadow and "homogeneous" areas the image content is enhanced.
The image pyramid starts from the original resolution images. Each pyramid level is generated by multiplying a generation kernel and reduces the resolution by factor 3. The pyramid level number is a pre-defined value that is either a user-input or can be determined according to the height range of the imaging area.

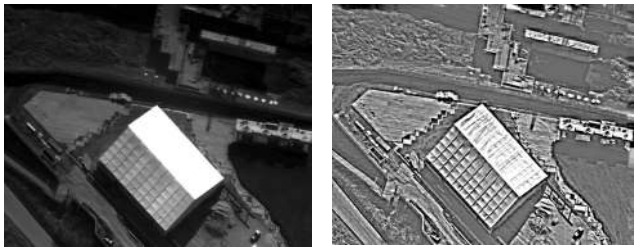

Figure 2: SI image before and after pre-processing

3.2 Feature Point Matching

We use the Foerstner interest operator to extract well-defined feature points that are suitable for image matching. Firstly the reference image is divided into small image patches (the nadir viewing SI or satellite image is selected as reference). Only one feature point will be extracted in each image patch. The density of the feature points can be controlled by the size of the patches.
The determination of the correspondences of the given points on the search images is carried out using the geometrically constrained cross-correlation method (see Gruen, Zhang, 2003). The matching candidates are computed by cross-correlation with a set of adaptively determined parameters like the image window size w_s, the threshold of the correlation coefficient c_t and the search distance. The approximate DSM that is derived from the higher-level of the image pyramid is used to estimate these parameters.
We incorporate the method proposed by Kanade & Okutomi, 1994 to select an optimal window size by evaluating the image content and the disparity within the matching window. As a result, in flat areas with small image intensity variations, the window size w_s increases and in areas of large terrain elevation variations it decreases. The threshold of the correlation coefficient c_t should also vary according to the roughness of the terrain. We set a larger value in flat areas and smaller value in rough terrain areas. The search window size depends on the terrain elevation variation in a small neighborhood of the given point and on the image geometry. In relatively flat areas the size of the search window decreases and vice versa.
By adaptive selection of these parameters, we can both reduce the processing time and the probability for multiple matching candidates. The number of matching candidates can be further reduced by introducing a third image. For every candidate, its position on the third image can be predicted using the image orientation elements. If the correlation coefficient between the reference and the third image is lower than the threshold, this matching candidate will be dropped. However, we cannot completely avoid the ambiguity problem due to reasons like repetitive patterns. Our procedure takes n ($\leq$ 5) matching candidates with correlation coefficient values above the threshold c_t.
As a result, for each feature point on the reference image several matching candidates can be computed. The correct match is determined by analysing the following quality measures sequentially:
a) The correct match should have a clear and sharp maximal correlation coefficient. If there are more than one candidates and the value of the first correlation coefficient peak is more than two times of that of the secondary peak, the candidate that has the largest correlation coefficient value will be considered the correct match.
b) Using the same matching parameters, the feature point can be back-matched from the search images to the reference image. If the difference between this two-way matching is less than one pixel, the candidate is assumed to be the correct match.
c) Under the condition that the feature point appears on more than two images, the residuals of the photogrammetric forward intersection should be less than 2-3 times the standard deviations of image coordinates of the triangulation adjustment.

The points that pass these tests will be indicated as reliable matches. The points that cannot pass these tests may have multiple solutions. This matching ambiguity will be solved by the following global image matching method through local smoothness constraints (see chapter 3.4).

3.3 Edge Matching

To reconstruct a DSM from very high-resolution images over urban areas we must take into account the problems caused by surface discontinuities, occlusions and the significant perspective projection distortion. Even with satellite images, line features are also important for capturing and modeling terrain features such as ridgelines and breaklines. Matching the edges is a possible solution to these problems. However we should consider the following problems:

- The edge on one image may break into more than one segment due to image noise, occlusions and the deficiencies of the feature extraction algorithms.
- The conjugate edges on different images may have quite different shapes due to the projection distortion.
- There may be many similar features in a search area.

The edge matching procedure presented here is based on the evaluation of the local geometric and photometric attributes of edges for the solution of disambiguities. The quasi-epipolar geometry and the DSM data derived from the higher-level of the image pyramid are used to provide for the matching candidates for each edgel. A figural continuity constraint satisfaction scheme (the disparity along an edge should change smoothly) and a shape matching approach are used to achieve the final results.

The well-known Canny operator is used to locate the intensity discontinuities. Then edgels are linked into free-form edges through a local processing that analyses the characteristics of these pixels in a small neighborhood. This approach is carried out independently on three images. Only edges above a minimum length (30 pixels for SI images and 15 pixels for satellite images) are considered for matching.

The edgels along the given edge are matched with the edges that are defined at the intersection points between the candidate edges and the correspondend epipolar curve within the search window on one of the search images. The search window can be determined by using the same method as in chapter 3.2. There may be several matching candidates within the search window. To solve this ambiguity problem we perform the following three steps sequentially:

a) Evaluation of the difference of the local edge orientation between the given edgel and its candidates. The local orientation for an edgel is the image intensity gradient computed modulo 2π. Candidates with differences above a threshold will be dropped. Considering the possible relief distortion, this constraint should not be too tight. For example, we use 40 degrees for SI images.
b) Evaluation of the normalized cross-correlation coefficient of the intensity values on each side of the edgel. Exclusion of those candidates that have a very low correlation coefficient, e.g. less than 0.5.
c) If more than two images are available, each candidate can be validated on the third image through the indicator used in step b).

After these three steps, the given edgel may still have more than one candidate. The problem will be further solved by using the figural continuity constraint through a relaxation method. This method examines the candidates by computing how much support they receive from their local neighborhood along the edge. We select the candidate that gains the highest support as the correct match for each edgel. For each edge, the edgels that have only one candidate will serve as "anchors" for this relaxation method. By linking the successfully matched edgels in the search images, we obtain the correspondent edge matc. Finally, we do a shape matching between the given edge and its correspondent edge through least squares adjustment. Only the edges with small shape matching errors will be kept.

Figure 3 shows an example of our edge matching.

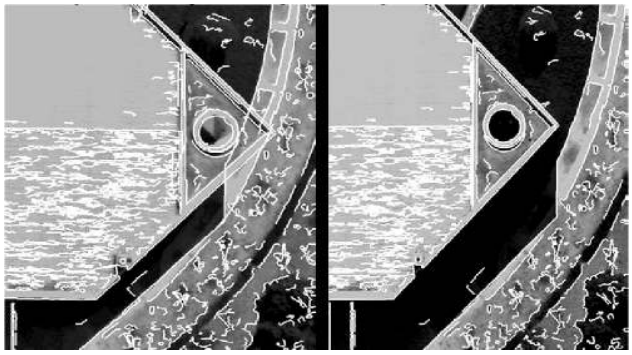

Figure 3: Edge matching with SI images (5 cm resolution)

3.4 Grid Point Matching

We use grid points to create uniformly distributed points over the whole images even in areas with very little or no texture. The correspondences of these grid points could be computed by using the method presented in chapter 3.2. Compared to the feature points, the choice of grid points is blind and thus many grid points are located in areas with weak or no texture. The search for the match of a given grid point has more possibilities to yield multiple candidates, or even no candidate.

To solve this problem, we use a global image matching method with relaxation technique. This method examines the candidates by computation how much support they receive from their local neighborhood and select the candidate that gains the highest support as the correct match. Here we use Prazdny's "coherence principle" model (Prazdny, 1985). We incorporate this idea in our global image matching and get the solution by relaxation technique.

Firstly, the points are selected in form of a regular grid in the reference image. Their matching candidates on the search images are computed. Together with all the matched feature points and edges they construct a TIN. It should be noted here that all the matched points can be categorized into three classes: Points having reliable matches, points having several candidates, and points without matching candidates. In the first case, they are treated as having only one matching candidate and, they serve as "anchors" for the global matching procedure. For the last case, they will be given several "false" candidates (with a very small correlation coefficient value) evenly distributed within the search window. The matched edges serve as break-lines in the TIN structure. They control the weights of the local smoothness constraints.

This method is performed on stereo pairs. The key point of this method, that distinguishes it from the single point matching, is its compatible coefficient function and its smoothness constraint satisfaction scheme. With the smoothness constraint, areas with homogeneous or only little texture can be bridged over, assuming that the terrain surface varies smoothly over the area. In the meantime, the surface discontinuities can be preserved because the smoothness constraints cannot cross the edges. For details of this procedure see (Gruen, Zhang, 2003).

3.5 Matching Through the Image Pyramids

A triangular irregular network (TIN) based DSM is constructed from the matched features on each level of the pyramid, which in turn is used in the subsequent pyramid level for the approximations and adaptive computation of the matching parameters. The matched edges are used as breaklines such that no triangle crosses these edges. The TIN maintains the original matching results without any interpolation. The surface discontinuities of the terrain can be well captured and preserved.

The initial DSM for the highest level of image pyramid can be extracted by standard cross-correlation based on a "region growing" matching strategy. This method uses the already measured control and tie points as seed points and matches the points under the assumption that points in a local neighborhood should have similar disparities (Otto, Chau, 1988). This method is justified because the disparity surface can be treated as continuous and smooth on the lowest resolution image pyramid level. In some difficult areas like very rough alpine terrain, some

manually measured points can also be introduced as seed points. It will give better approximations for the matching.

3.6 Refined Matching Based on the Modified MPGC

MPGC (Multi-Photo Geometrically Constrained Matching) was developed by Gruen, 1985, and is described in detail in Baltsavias, 1991. It combines least squares matching and geometric constraints formulated either in image or in object space and permits a simultaneous determination of pixel and object coordinates. Any number of images (more than two) can be used simultaneously. The achieved accuracy is in the sub-pixel range.

Our modified algorithm is an extension of the standard MPGC. It integrates the geometric constraints derived from the Linear Array sensor models. The geometric constraints force the matching to search for a conjugate point only along a small band around the epipolar curve and reduce the possibility of false matches.

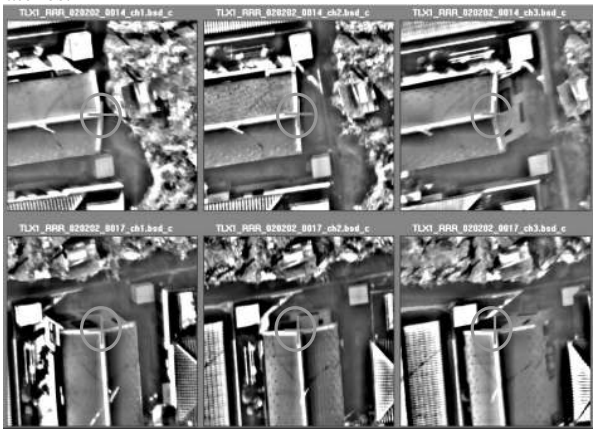

Figure 4: MPGC matching with multi-strip SI images.
Top: Images of a strip acquired from west to east
Bottom: Images of the cross-strip

The modified MPGC is used to refine the matched features in order to achieve sub-pixel accuracy. The DSM derived from the approaches (3.2)-(3.4) provides quite good approximations for the MPGC procedure and increases the convergence rate. The initial values of the shaping parameters in MPGC can be predetermined by using the image geometry and the derived DSM data. The corresponding rays of the four corners of the matching window in the reference image are intersected with the derived DSM and their object coordinates can be determined. Through imaging geometry, the corresponding image coordinates in the search images can also be determined. The initial values of the shaping parameters can easily be computed from these four corner points and their correspondences. By this way, features can be matched with multiple images or even with multiple image strips that have different flight directions and image scales. Figure 4 gives an example for the case of cross-strips.

For edge matching the parameters of a spline function of the 3D object edge are directly estimated together with the matching parameters of edges in the image spaces of multiple images. Some points, especially those grid points in non-texture or little texture areas, will fail in MPGC matching. These points are also kept but they are assigned low reliability indicator values.

4. EXPERIMENTAL RESULTS

In order to evaluate the performance of our approach for DSM generation it has been applied to different areas with varying textures, terrain types and image scales. In the following we will report about DSM accuracy test results with SI, IKONOS and SPOT5 HRS images.

4.1 SI Image Dataset, GSI area, Japan

In Japan's GSI (Geographical Survey Institute) test area, both SI images and aerial photos are available. The evaluation is based on the comparison between the manually measured DSM from aerial photos and the automatically extracted DSM from the SI images.

The GSI test area is roughly 650×2500 m^2. It is relatively flat with natural and artificial objects. There are a lot of small geomorphological features, but also small discontinuities like cars, isolated trees and large discontinuities and occlusions due to buildings. For the success of a matcher it is very important how it can handle local discontinuities (e.g. buildings or other man-made objects, vertical cliffs, etc.).

Figure 5 shows two image windows from the nadir-viewing SI images. We have 48 control points. They are signalized marks on the ground or on the top of buildings. They appear both in the SI and aerial photos.

Two stereo pairs of color aerial images of 1:8000 image scale, acquired with a film camera of 153 mm focal length, have been used for manual collection of reference data on an Analytical Plotter. The RMSE of the exterior orientation is reported as 5 cm in planimetry and 3 cm in height. Manually measured points, distributed at about 50 cm distance in both X and Y directions, are used as reference data.

Assuming a measurement accuracy of 0.1% of the flying height (best case situation for natural points) we can expect an accuracy of the reference data of 0.12 m as a best case scenario.

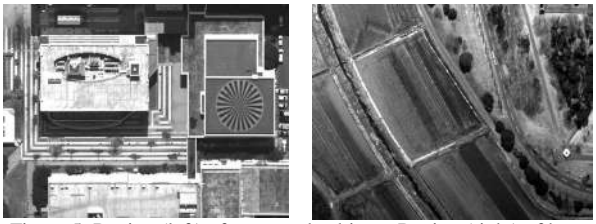

Figure 5: Region (left) of man-made objects. Region (right) of bare terrain area (including some sparse trees). The image patches are from the nadir SI-100 images with the ground resolution of 5.6 cm.

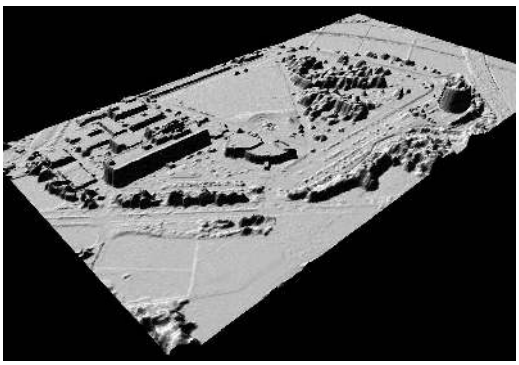

Figure 6: 3D visualization of the shaded DSM of the GSI test area.

Table 1: DSM accuracy evaluation.
Manually measurement minus automatically extracted DSM

Area	RMSE (m)	Mean (m)	% 0.0-1.0 m	% 1.0-2.0 m	% 2.0-3.0 m	% 3.0-4.0 m	% > 4.0 m
(1)	0.28	0.04	98.65	0.91	0.13	0.09	0.03
(2)	1.01	-0.16	74.70	12.15	7.32	4.84	2.36

The processed SI data include three panchromatic images with a footprint of ca. 5.6 cm. As a result of triangulation, 2.8 cm and 5.0 cm absolute accuracy in planimetry and height (as computed from checkpoints) were obtained.

For analysis of the matching accuracy, we divided the reference data into two classes, i.e. (1) the bare terrain area with some sparse trees and small artificial objects (including the large parking areas); (2) the area with man-made objects and trees. Our analysis has been performed for these two classes separately.

A very dense raster DSM with 15 cm interval was interpolated from the automatically matched point cloud and edges. The points with low reliability values were given a small weight in the interpolation procedure. Figure 6 shows the 3D visualization of the extracted DSM.

Comparing the reference points with the raster DSM in two different areas (1) and (2) leads to the results shown in Table 1. The bare terrain area shows much better results, but it still suffers from problems like multi-temporal differences between

aerial photos and SI images, and some sparse trees. In the man-made object area, the accuracy becomes worse. Here 2.4 % of all points points have more than 4 meters difference. These points are almost all positioned on the border of buildings and trees. This may be caused by the fact that manually measured points and automatically extracted points are referring to different object points due to small errors in orientation procedure and 3D modeling problems.

4.2 IKONOS Image Dataset, Thun, Switzerland

The test covers the area around the town of Thun, Switzerland. It is about 17×20 km^2. The terrain elevations range from 600 m to 2200 m, with some very steep and high cliffs.

One IKONOS stereo pair was acquired on 2003-12-11 10:29 GMT over an area of approximately 11×20 km^2. Another IKONOS triplet was acquired on 2003-12-25, 10:39 GMT over area of approximately 11×20 km^2. The two areas have nearly 50 % overlap. All images are IKONOS Geo products. The sensor and sun elevation angles (ca. 19 degrees) were less than optimal. The low elevation angle of the sun causes strong shadows, especially in the south part of the images and in general low contrast images. A 2 m resolution reference DSM generated from airborne LIDAR in the year 2000 was obtained from the Swiss Federal Office of Topography, Bern. The reference DSM only covers the south part of the whole area.

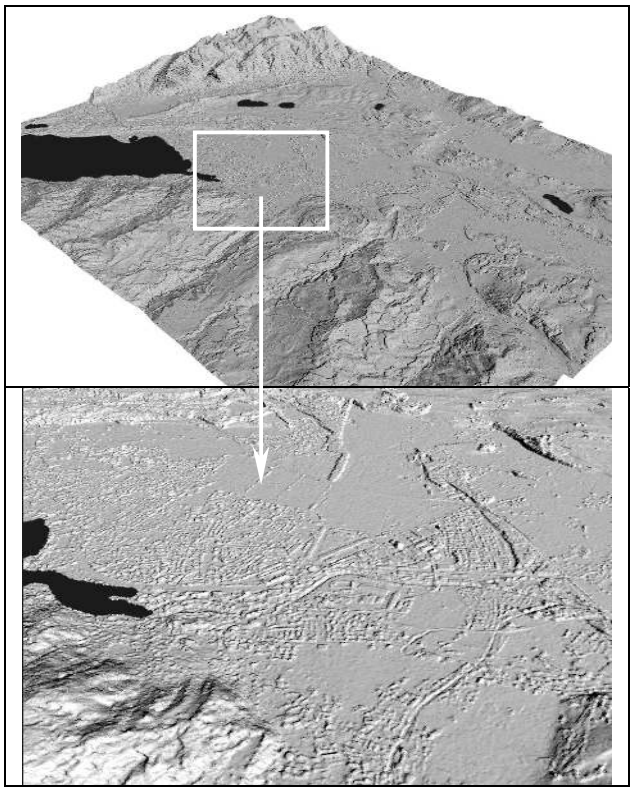

Figure 7: Extracted DSM
Top: The whole area; Bottom: Town of Thun

Firstly, the original 12 bit images were preprocessed and resulted in enhanced images for further image matching. Then the images were orientated with the help of about 40 GCPs measured by GPS. The orientation accuracy is about 0.41 m in planimetry and 0.68 m in height. Finally, a 5 m raster DSM was generated from the matched mass points and the edges. Some areas like lakes and rivers are defined as "dead areas" manually.

Figure 7 shows the 3D visualization of the generated DSM (the whole test area and part of the town area). The results show that even small geomorphological features are extracted and surface discontinuities are well preserved.

Table 2: DSM accuracy numbers (west part of test area).
O-Open areas; C-City areas; V-Tree areas; A-Alpine areas.

Area	No. of Compared Points	Mean (m)	RMSE (m)	< 2.0 m	2.0-5.0 m	> 5.0 m
O+C+V+A	29,210,494	1.21	4.80	60.7%	16.8%	21.3%
O+C+A	17,610,588	1.11	2.91	77.0%	13.9%	10.1%
O+A	14,891,390	1.24	2.77	79.8%	12.2%	8.0%
O	11,795,795	1.00	1.28	90.3%	8.5%	1.2%

Table 3: DSM accuracy numbers (east part of test area).
O-Open areas; C-City areas; V-Tree areas; A-Alpine areas.

Area	No. of Compared Points	Mean (m)	RMSE (m)	< 2.0 m	2.0-5.0 m	> 5.0 m
O+C+V	20,336,024	-0.45	4.78	57.7%	21.3%	20.9%
O+C	13,496,226	0.33	3.38	68.7%	20.8%	10.3%
O	3,969,734	0.97	1.54	83.0%	15.0%	2.0%

Table 2 and 3 give the DSM accuracy test results. We compute the differences between the interpolated heights from our DSM minus the heights of the reference DSM. The accuracy of the generated DSM is at 1.3 – 4.8 pixel level. It depends on the terrain type. Higher accuracy can be achieved in open areas. In urban and tree areas the accuracy becomes worse. The analysis shows that points with more than 6 m differences are almost all distributed in the tree and urban areas. In open areas (with some sparse trees and small cluster of houses), more than 70 percent of the points have differences of less than 1meter. The results show significant biases. This is caused by different point definitions in laserscan and photogrammetric surface points. Also, the different acquisition times of the IKONOS images and the laserscans may play a role . For more details see (Eisenbeiss, et al., 2004).

Table 4: Reference DEMs. Height accuracy in meter

DEM Name	Location	DEM Spacing	Source	DEM Size	Height Accuracy
DLR-DEM-01	Prien	5 × 5	Laser Scanner	5km × 5km	0.5
DLR-DEM-02	Gars	5 × 5	Laser Scanner	5km × 5km	0.5
DLR-DEM-03	Peterskirchen	5 × 5	Laser Scanner	5km × 5km	0.5
DLR-DEM-04	Taching	5 × 5	Laser Scanner	5km × 5km	0.5
DLR-DEM-05-1	Inzell-North	25 × 25	Laser Scanner	10km × 1.3km	0.5
DLR-DEM-05-2	Inzell-Sourth	25 × 25	Contour lines	10km × 7.7km	5.0
DLR-DEM-06	Vilsbiburg	50 × 50		50km × 30km	2.0

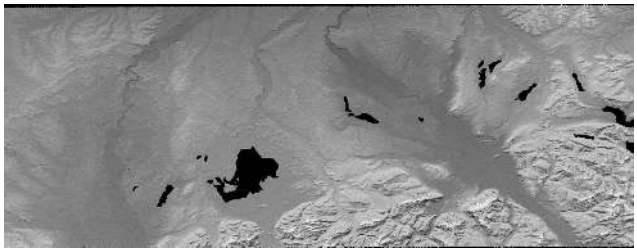

Figure 8a: Color shaded extracted DSM (the whole test area)

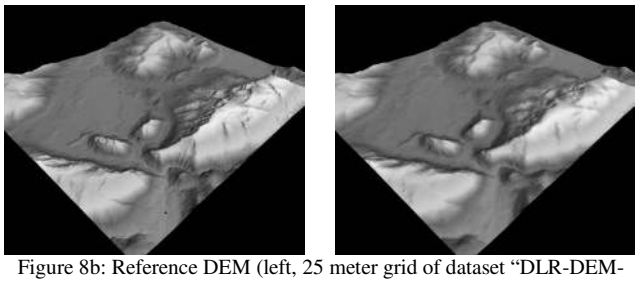

Figure 8b: Reference DEM (left, 25 meter grid of dataset "DLR-DEM-05-2") and the generated DSM (right, 25 meter grid)

4.3 SPOT5 HRS Image Dataset, Bavaria, Germany

Here we report the work carried out within the ISPRS–CNES Initiative on DEM generation from SPOT5-HRS stereo images. For details see (Poli et al., 2004)

The test area (No. 9) covers a part of South Bavaria and a part of Austria (approximately 120×60 km^2). Table 4 gives information on the reference DEMs. A stereo pair from SPOT5-HRS was acquired on 1 October 2002 in the morning. The

images have the resolution of 5 m in along-track and 10 m in across-track directions.
The test area includes a mountainous area (rolling and strongly inclined alpine area) and hilly areas (rough/smooth and weakly inclined areas). Our image matching software not only generates a large number of mass points, but also produces line features. The TIN based DSM was generated from the matched mass points and the edges (as break-lines).
Figure 8 shows the 3D visualization of the generated DSM. The results show that the shapes of our generated DSMs are similar to the references, but slightly smoothed. This can be expected because of the 5m resolution of the satellite images.
Tables 5 and 6 show the DSM accuracy test results. The orientation accuracy is about 6.3 m in planimetry and 2.6 m in height. We compute the differences between the heights of the reference DEM and the interpolated heights from our DSM. Table 6 shows the DSM accuracy test result by masking out the tree areas manually.
From Tables 5 and 6 it can be seen that:

- The accuracy of the generated DSM is more or less at the 1pixel level or even better. Only the datasets 5 give values at about 2 pixels, but these higher values are mainly caused by trees.
- All datasets still contain some blunders, which our procedures failed to detect.
- The results show systematic errors. In datasets 5-1 and 5-2 the biases are about 1 pixel. Except in case of dataset 6 all biases are significantly negative. This indicates that our generated DSMs are higher than the reference DEMs, an effect which could be expected..

Table 5: DSM accuracy, units are meter

DEM Name	Maximum Difference	Minimum Difference	Average	RMSE
DLR-DEM-01	25.1	-32.9	-2.6	5.7
DLR-DEM-02	29.1	-37.1	-1.2	5.0
DLR-DEM-03	20.7	-17.2	-0.5	3.2
DLR-DEM-04	13.6	-23.1	-2.5	4.7
DLR-DEM-05-1	19.2	-33.5	-5.8	8.3
DLR-DEM-05-2	136.8	-89.3	-4.3	9.5
DLR-DEM-06	26.8	-27.1	1.5	4.0

Table 6. DSM accuracy, units are meter
(excluding the tree covered areas)

DEM Name	Maximum Difference	Minimum Difference	Average	RMSE
DLR-DEM-01	15.4	-23.7	-1.7	4.6
DLR-DEM-02	29.1	-31.7	0.2	3.6
DLR-DEM-03	20.7	-13.6	0.1	2.9
DLR-DEM-04	10.5	-18.4	-1.2	3.2
DLR-DEM-05-1	19.1	-13.3	-1.7	4.9
DLR-DEM-05-2	49.8	-66.8	-1.3	6.7
DLR-DEM-06	26.8	-25.9	2.1	4.4

5. CONCLUSIONS

In this paper we have reported about our current matching approaches for fully automated DSM generation from linear array images with different resolutions. We have developed a matching strategy combining feature point matching, grid point matching with neighborhood smoothness constraints, and robust edge matching. The strategy allows us to bridge over areas with little or no texture and at the same time maintain the important contribution of object/image edges. The modified MPGC is used to refine the matching results in order to achieve sub-pixel accuracy. The geometrical constraints are derived from the specific sensor models for the linear array imagery, which can be the rigorous sensor model for aerial and satellite images or the RF (Rational Function) model for satellite images.
As evidenced by a visual inspection of the results we can reproduce even small geomorphological features. The results from the quantitative accuracy test indicate that the presented concept leads to good results. If the bias introduced by trees and buildings is taken out, we can expect a height accuracy of one pixel or even better from satellite imagery (e.g. IKONOS and SPOT) as "best case" scenario. In case of very high resolution aerial images (footprint 8 cm and better) it is obvious that the "one pixel rule" cannot be maintained any more. Alone surface roughness and modeling errors will lead to large deviations, such that an accuracy of three to five pixels should be considered a good result. This is at the same level as laser scanning results. Of course, the photogrammetric data can also be produced with the same or even better point density. On the other hand, with these accuracies we are still operating at a coarser level than with manual measurements from analogue aerial images, but we do that with the advantage of great gain in processing speed.
A major problem left is the control and automated detection of small blunders, which still infest the results, despite the simultaneous matching of more than two images. This constitutes a relevant topic for further research.

ACKNOWLEDGEMENTS. We appreciate the support of the Swiss Federal Office of Topography, Bern, which provided the laserscan data. We also thank Henri Eisenbeiss, who helped in setting up the Thun area as a testfield for highresolution satellite image processing.

REFERENCES

Baltsavias, E. P., 1991, Multiphoto Geometrically Constrained Matching. *Dissertation, IGP,* ETH Zürich, Mitteilungen No. 49, 221 pages.

Eisenbeiss, H., Baltsavias, E., Pateraki, M., Zhang, L., 2004. Potential of IKONOS and QuickBird Imagery for Point Positioning, Orthoimage and DSM Generation. *IAPRS,* Vol. 35, Istanbul, Turkey (to be published)

Gruen, A., 1985, Adaptive Least Squares Correlation: A powerful Image Matching Technique. *South Africa Journal of Photogrammetry, Remote Sensing and Cartography*, 14 (3), pp. 175-187.

Gruen, A., Bär, S., Bührer, Th., 2000: DTMs Derived Automatically From DIPS - Where Do We Stand? *Geoinformatics,* Vol.3, No.5, July/August, pp. 36-39.

Gruen, A., Zhang, L., 2002. Sensor Modelling for Aerial Mobile Mapping with Three-Line-Scanner (TLS) Imagery. *ISPRS Commission II Symposium on Integrated System for Spatial Data Production*, Xi'an, P. R. China, August 20 – 23.

Gruen, A., Zhang L., 2003. Automatic DTM Generation from TLS data. *Optical 3-D Measurement Techniques VI*, Vol. I, Zurich, pp. 93-105.

Kanade, T., Okutomi, M., 1994. A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment. *IEEE Transactions on PAMI,* Vol. 16, No. 9, pp. 920-932.

Murai, S., Matsumoto, Y., 2000. The Development of Airborne Three Line Scanner with High Accuracy INS and GPS for Analysing Car Velocity Distribution. *IAPRS,* Vol. 33, Part B2, Amsterdam, pp. 416-421

Otto, G. P., Chau, T. K. W., 1988. A "Region-Growing" Algorithm for Matching of Terrain Images. *Proc. 4th Alvey Vision Club,* University of Manchester, UK, 31 Aug. – 2 Sept.

Poli, D., Zhang, L., Gruen, A., 2004. SPOT-5/HRS Stereo Image Orientation and Automatic DSM Generation. *IAPRS,* Vol 35, B 1, Istanbul, Turkey (to be published)

Saint-Marc, P., Chen, J-S., Madioni, G., 1991. Adaptive Smoothing: A General Tool for Early Vision. *IEEE Transactions on PAMI*, Vol. 13, No. 6

Prazdny, K., 1985. Detection of binocular disparities. *Biological Cybernetics*, Vol. 52, pp. 93-99.