Matching of Ikonos Stereo and Multitemporal GEO Images for DSM Generation

Li Zhang, Maria Pateraki, Emmanuel Baltsavias
Institute of Geodesy and Photogrammetry, ETH-Hoenggerberg, CH-8093, Zurich, Switzerland
(zhangl, maria, manos)@geod.baug.ethz.ch

Abstract
Investigations on the use of stereo GEO and multitemporal GEO Ikonos images for DSM generation are reported. Both an own developed matching algorithm and the digital photogrammetric systems LH Systems DPW770 and VirtuoZo were used. In our own matching method, we start with quasi-epipolar image generation, and derivation of approximations using grid point matching based on relaxation. The final matching is a modified multi-image geometrically constrained method. Test results on both methods are presented. Firstly, DSMs were generated from multitemporal GEO images, which theoretically should be more difficult to match, by using our own matching procedure. Different matching versions, with and without exclusion of the different cloud areas, were tried and compared to a reference DTM with an accuracy of ca. 3 m. The results showed a very promising RMS of slightly over 3 m without any manual editing. The second example refers to a DSMs generation from a GEO stereo-pair by matching using LHS DPW770 and VirtuoZo and compared to results from aerial imagery. Due to large multitemporal differences between the two datasets, stable points were manually selected and measured. These points and other points in stable open areas that were automatically matched in the aerial imagery using VirtuoZo served as reference data. The results showed a high relative accuracy of ca. 1 m but the presence of a larger systematic error of 2-3 m.

KEY WORDS: high-resolution satellite imagery, IKONOS, DSM, matching, accuracy analysis

1. Introduction
Although IKONOS imagery has been commercially available since early 2000, the use of this imagery for DSM generation has been restricted due to various reasons. Firstly, up until begiining of 2002 stereo imagery was available only for governments and national mapping organisations. Secondly, Space Imaging follows a closed policy and does not release the sensor model but provides rational polynomial coefficients instead. These are provided for stereo images and special versions of Geo imagery (so called OrthoKit) but at significantly higher prices. To these reasons, one should add the general problems of Ikonos imagery like cost, availability, lack of user control on imaging parameters and conditions etc. Since summer 2001, some commercial systems (Erdas Imagine, LHS Socet Set, Z/I Imaging ImageStation, PCI Geomatics OrthoEngine) support to one or the other extent IKONOS imagery for import, stereo viewing and processing, orthoimage and DSM generation by using the rational polynomial coefficients (RPCs) provided by SI, estimating RPCs from ground control or using alternative (bundle adjustment, DLT) and partly proprietary mathematical models. While estimating RPCs from GCPs is both expensive (high number of GCPs required) and sub-optimal, use of bundle adjustment due to the undisclosed IKONOS sensor model is impossible. DLT, as shown in Fraser et al. (2002) could be used with GCPs, while proprietary sensor models (PCI) are of unknown quality, while raw IKONOS data which would be most appropriate for use with a strict sensor model are not available. As shown in Baltsavias et al. (2001) and Fraser et al. (2002), a very simple alternative geometric model, needing at least 3 GCPs, can achieve better positioning accuracy than the RPCs and at sub-meter level.
Investigations on DSM generation from high-resolution imagery include the following. Ridley et al. (1997) evaluated the potential of generating a national mapping database of maximum building heights for buildings at least 5 x 10 m in planimetry by using DSMs extracted by matching of 1m aerial imagery. They report that matching has a potential to provide the requested information if the DSM had a spacing of 1-3 m but with lower accuracy (1.5 – 3 m RMS) and completeness compared to manual measurements. Muller et al. (2001) use simulated 1m-resolution and IKONOS data for DSM generation and landuse determination to estimate effective aerodynamic roughness for air pollution modelling and determine position of trees close to buildings that may cause soil subsidence for insurance risk assessment. The most extensive results on DSM generation from Ikonos up to now have been reported by Toutin et al. (2001). Dial (2000) also reports the expected mapping accuracy of Ikonos. There are no published investigations on the other two operational high-resolution spaceborne systems (EROS-A1 and Quickbird-2). As far as the authors know, this is the first investigation deriving DSMs from multitemporal Ikonos images. This data are much more available than stereo data, they do not include epipolar resampling which possibly degrades geometric accuracy and are cheaper than stereo Ikonos. Naturally, multitemporal differences make matching more difficult, but as it is shown in Baltsavias and Stallmann (1992) it is possible with intelligent techniques to achieve good results even in case of large image differences.

2. Test Data
2.1 Melbourne Dataset
The first dataset comprised a stereopair of epipolar-resampled Geo PAN images of Melbourne. As indicated in Table 1, the sensor and sun elevation angles for the stereopair (imaged in winter) were less than optimal. The azimuths of sensor and sun of the left image differed considerably, leading to strong shadows in non-occluded areas. The base/height ratio of the stereopair was 1.2. GCPS measured with an accuracy of 1-2 dm in both image and object space existed and could be used for our Ikonos sensor model and aerial image orientation. More details on the dataset and the GCPs can be found in Fraser et al. (2002). An existing stereopair of 1:15,000 scale wide-angle colour aerial photography was used to measure reference data.

2.2 Nisyros Dataset
This dataset consisted of two Pan-Sharpened Geo images of a Greek island with little man-made objects and sufficient vegetation. It is an non-active volcano with elevation range 0-700 m. The parameters of the images are listed in Table 2. As it can be seen, they were similar except of the sensor azimuth. The different viewing angle let to illumination differences between the two images. The fact that shadows were mostly in occluded areas was an advantage. The major multitemporal differences were due to different clouds

and their shadows. 38 GCPs measured with GPS were available with an object accuracy of less than 1 dm. About half of them were not well-defined in the image. The reference DTM had a grid spacing of 2 m and was interpolated from 1:5,000 map contours and additionally measured breaklines. Its accuracy, estimated using the GCPs, was ca. 3.3 m RMS. More details on the data can be found in Vassilopoulou et al. (2002).

	Left Stereo	Right Stereo
Date, Time (local)	16/7/2000, 09:53	16/7/2000, 09:53
Sensor azimuth (deg)	136.7	71.9
Sensor elevation (deg)	61.4	60.7
Sun azimuth (deg)	38.2	38.3
Sun elevation (deg)	21.1	21.0

Table 1. Acquisition parameters of the Melbourne dataset.

	Image 1	Image 2
Date, Time (local)	8/4/2000, 10:35	28/3/2000, 10:34
Sensor azimuth (deg)	134.9	72.1
Sensor elevation (deg)	73.5	72.7
Sun azimuth (deg)	136.6	138.7
Sun elevation (deg)	53.4	49.3

Table 2. Acquisition parameters of the Nisyros dataset.

3. Own Matching Method and Test Results with Nisyros Dataset

3.1 Quasi-Epipolar Image Generation

Unlike frame-based imagery, where all pixels in the image are exposed simultaneously, each scan line of the Ikonos image is collected in a pushbroom fashion at different instant of time. Thus, epipolar lines with linear CCDs become curves. Using Orun and Natarajan' sensor model (Orun and Natarajan, 1994), which models the position of the sensor and its yaw angle variation as second-order polynomials of scan line (or time), and assuming the pitch and roll angles to be constant, the derived epipolar curve for a certain point shows a hyperbola-like shape. It can be approximated by a straight line for a small length but not for the entire image. An epipolar curve for the entire image can be approximated only by piecewise linear segments. In our approach, the epipolar curve for point (x_l, y_l) in the left Ikonos image is approximated by a quadratic polynomial and has the following form:

$$y_r = (A_1x_l+A_2y_l+A_3) + (A_4x_l+A_5y_l+A_6)x_r + (A_7x_l+A_8y_l+A_9)x_r^2 \qquad (1)$$

where, (x_r, y_r) are the pixel coordinates in the right Ikonos image and $A_{1\text{-}9}$ are unknown parameters. Using Eq. (1), the quasi-epipolar image pair can be generated by re-arranging of the original Ikonos image pair. However, the computation of the unknown parameters of the epipolar geometry needs a certain number of well distributed conjugate points. These conjugate points are extracted by the following feature-point based matching:

- The Moravec interest operator is used to select well defined feature points that are suitable for image matching. For this, the left Ikonos image is divided into small image windows of 21 × 21 pixels and then only one feature point, which has the highest interest value is extracted in each image window.
- Conjugate points are generated using the maximum of the normalized correlation coefficient. The positioning of the search areas is determined by using the already known control points in the neighborhood (image pyramids and a matching strategy based on region growing, which takes the already manually measured control points as seed points are used to get these approximate points). For reliability, the threshold of acceptable normalized correlation coefficients is 0.9.
- Least squares matching is finally used to refine the image coordinates of these points in order to achieve sub-pixel accuracy.

This procedure results in several hundreds of conjugate points. These points can be used to recover the epipolar geometry and interpolate the approximate values for the following grid point matching procedure.

3.2 Derivation of Approximations by Using Grid Point Matching based on Relaxation

The matching procedure can be treated as a labeling problem and solved by the relaxation technique. That is, we regard the template image (in our case, the left quasi-epipolar image) as the model and another image as the scene, and then use the feature points in the model as a set of labels to label the feature points extracted from the scene. Relaxation is one of the efficient methods to solve the labeling problem. The work of Hancock and Kittler (1990) that uses Bayesian probability theory has provided a theoretical framework and rigorous basis for the relaxation method.

The important aspect of the relaxation matching algorithm that distinguishes it from the single point matching is its compatible coefficient function and its smoothness constraint satisfaction scheme (Baltsavias, 1991; Zhang et al., 1992). This is most important for areas with homogeneous or only little texture. In such areas, the single point matching is unable to match images because of the lack of information. With the smoothness constraint, such areas can be bridged over, assuming that the terrain surface varies smoothly over the area.

Firstly, the match points are selected and distributed in the form of a regular grid in the template image. Given a grid point in the template image, a search window in the search image (the right quasi-epipolar image, in our case) can be determined. The correct match of this point should lie in this search window. However, due to repetitive texture or poor texture information, there could be several candidate matches appearing in the search window. These candidate matches are located along the epipolar curve. They can be derived by traditional cross-correlation technique, and the candidate matches are selected if their correlation coefficient lies above a certain user-defined threshold (we choose this threshold value as 0.7). The approximate centers of the search windows can be derived from the matching results of the previous pyramid level.

Let I_i be one of the grid points on the template image and I_j ($j=1,\ldots,m$) its candidate matches in the search image. *P(i,j)* is the probability of match I_i? I_j. Moreover, let I_k be one of the points located in the neighborhood of point I_i and I_l ($l=1,\ldots,m$) be its corresponding candidate matches.

In order to link the matching results of the neighboring grid points to each other, we define the following compatible coefficient function *C(i,j;k,l)* which quantifies the compatibility between the match I_i? I_j and a neighboring match I_k? I_l:

$$C(i,j;k,l) = \frac{T}{\exp(\Delta p^2 / b)} \qquad where \quad \Delta p = (x_j - x_i) - (x_l - x_k) \qquad (2)$$

In equation (2), *?p* expresses the difference of the x-parallaxes in point I_i and its neighboring point I_k. The bigger the *?p*, the smaller the compatibility. This corresponds to a smoothness constraint on the image matching results, and it provides an ability to bridge over

the poor texture areas. *T* is a value quantified by the texture information and it is defined as inversely proportional to the minimum of four gray value variances (horizontal, vertical, and along the two main diagonals) at the image window around the point I_i. Normally, if one point is located in rich texture areas or at linear features, this value is small. This value can be treated as the weight of the smoothness constraint and it provides the possibility to control the continuity of the terrain surface. *ß* is a constant and its value was set to 400 empirically.

In the relaxation scheme, the so-called global consistency of matching can be achieved by an iterative scheme where the probabilities *P(i,j)* in iteration n+1are updated by the following rule:

$$P^{(n+1)}(i,j) = \frac{P^{(n)}(i,j)Q^{(n)}(i,j)}{\sum_{s=1}^{m} P^{(n)}(i,s)Q^{(n)}(i,s)} \qquad where \quad Q^{(n)}(i,j) = \prod_{I_k \in \Omega(I_i)} \sum_{l=1}^{m} P^{(n)}(k,l)C(i,j;k,l) \tag{3}$$

C(i,j;k,l) is the compatible coefficient function defined as above, $O(I_i)$ is the neighbourhood of point I_i (can be its 8 or 24 neighboring points), and n is the iteration number. The quantity $Q^{(n)}(i,j)$ expresses the support the match I_i? I_j receives at the n^{th} iteration step from the matches I_k? I_l in its neighbourhood $O(I_i)$.

The iteration scheme can be initialized by assigning the normalized correlation coefficient to $P^{(0)}(i,j)$ and, ideally the process will terminate when an unambiguous match result is reached, that is when each point I_i is matched with one candidate with probability 1, the probabilities for all other candidate matches for this point being zero. In practice, we terminate the process if any one of the following two conditions holds:

- For each grid point I_i, one of the match probabilities *P(i,j)* (*j=1,…,m*) exceeds 1-*e*, where *e* << 1 (for example, we set the value of *e* to 0.1).
- A pre-defined number of maximum iterations has been reached.

When the iterative procedure is terminated, the match which has the highest probability *P(i,j)* (*j=1,…,m*) is selected as the actual match. For speeding up the processing, reduction of the search range and higher reliability, an image pyramid is used. The image pyramid is generated from the quasi-epipolar images by a resolution reduction factor 3. The pyramid level number is a pre-defined value, this value can be a user input or determined according to the height range of the imaged area. The advantage of this method is that it can achieve reasonable matching results even in areas of little or no texture. Its disadvantage is that the matching results only have pixel-level accuracy. The relaxation matching results can be projected to the original Ikonos images and further refined by the following modified MPGC matching procedure.

3. 3 Modified MPGC

MPGC is described in detail in Baltsavias (1991). It combines least squares matching and geometric constraints formulated either in image or object space. The constraints lead to a 1-D search space along an epipolar line, thus to an increase of success rate, accuracy and reliability, and permit 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. The algorithm also provides criteria for the detection of observation errors and blunders, and adaptation of the matching parameters to the image and scene content.

Usually, the sensor model used for Ikonos is the rational function model, other models like polynomial mapping functions (Baltsavias and Stallmann, 1992) or DLT can be treated as a special case of the rational polynomial functions. For description of the RPCs see Grodecki (2000). Their form is given by:

$$x_n = F_x(X_n, Y_n, Z_n) = \frac{f_1(X_n, Y_n, Z_n)}{f_2(X_n, Y_n, Z_n)}; \qquad y_n = F_y(X_n, Y_n, Z_n) = \frac{f_3(X_n, Y_n, Z_n)}{f_4(X_n, Y_n, Z_n)} \tag{4}$$

where, (x_n, y_n) and (X_n, Y_n, Z_n) are normalised image and object coordinates respectively.

The geometric constraints were derived from Eq. (4) and have the following form:

$$\begin{aligned} v_x &= -\frac{\partial(x_n)}{\partial x}\Delta x + \frac{\partial F_x}{\partial X}\Delta X + \frac{\partial F_x}{\partial Y}\Delta Y + \frac{\partial F_x}{\partial Z}\Delta Z + (x_n - F_x(X_n, Y_n, Z_n))^0 \\ v_y &= -\frac{\partial(y_n)}{\partial y}\Delta y + \frac{\partial F_y}{\partial X}\Delta X + \frac{\partial F_y}{\partial Y}\Delta Y + \frac{\partial F_y}{\partial Z}\Delta Z + (y_n - F_y(X_n, Y_n, Z_n))^0 \end{aligned} \tag{5}$$

Where, (*x, y*) and (*X, Y, Z*) are image and object coordinates respectively. Equations (5) can be treated as weighted observation equations in MPGC, where (*?x, ?y*) are the unknown shifts in pixels, which are the common unknowns, appearing also in the affine parameters of the MPGC.

Weighted geometric constraint force the matching to search for a conjugate point only along the epipolar curve. If the initial match of the point in the search image does not lie on this epipolar curve, at the first iteration of MPGC matching, it jumps onto this curve. The grid point matching results based on relaxation provide quite good approximations for the MPGC procedure and increase the convergence rate. In our implementation of the modified MPGC, the adjustment starts only with the two shift parameters and after the first convergence all affine transformation parameters are used. Also, at the first two iterations, the weight value for the geometric constraints takes a large value in order to speed up the convergence and then it decreases to consider remaining errors in the orientations/epipolar geometry and allow matching slightly off the epipolar curve. The modified MPGC matching procedure has the potential of obtaining sub-pixel matching results and exploiting more than 2 Ikonos images simultaneously.

3.4 Test Results

For the Nisyros dataset, the DSM was generated for the entire island, including a part on the sea. The approach starts with a feature-point based matching to obtain conjugate points for epipolar geometry recovery and approximation values for relaxation matching. This stage provides 114 conjugate points that are refined by least squares matching without geometric constraints. The refined matching results are used for generation of quasi-epipolar images by using bilinear interpolation of the grey values. Then, the grid relaxation matching is performed on the quasi-epipolar image pair. The grid interval was 7×7 pixels and the patch size also 7×7 pixels. This step results in 852,615 grid points (493,916 left after cutting out the results of cloud covered areas). These points were then transformed to the original Ikonos images and 3D coordinates needed by MPGC were computed by spatial resection. The modified

MPGC procedure is performed by taking the relaxation matching results as approximations. For testing the potential accuracy and performance of our matching approach the following 5 different versions were executed:

Version 1: relaxation matching only, by cutting out the cloud covered areas manually.

Version 2: modified MPGC 2D matching, patch size is 21×21 pixels and no geometric constraints, by cutting out the cloud covered areas manually.

Version 3: modified MPGC matching, patch size is 21×21 pixels and with geometric constraints, by cutting out the cloud covered areas manually.

Version 4: modified MPGC matching, patch size is 11×11 pixels and with geometric constraints, by cutting out the cloud covered areas manually.

Version 5: modified MPGC matching, patch size is 21×21 pixels and with geometric constraints, without cutting out the cloud covered areas manually.

Versions 2-5 were applied to the raw images, but after noise reduction, contrast enhancement and radiometric equalisation. In all versions, the same match area and manually selected cloud areas were used. 3D coordinates were computed using the terrain-corrected affine transformation (Baltsavias et al., 2001) as sensor model. The affine parameters for each image were derived from just 4 GCPs. Table 3 shows the accuracy test results of different versions (matching – reference DTM). The results are according to the expectations with version 3 delivering the highest accuracy of ca. 3 m RMS, a remarkable result if one considers the multitemporal nature of the images. The positive bias is probably because matching measures on top of buildings, trees etc. In Fig. 1, an image area with contours derived from the matched results of Version 3 is shown.

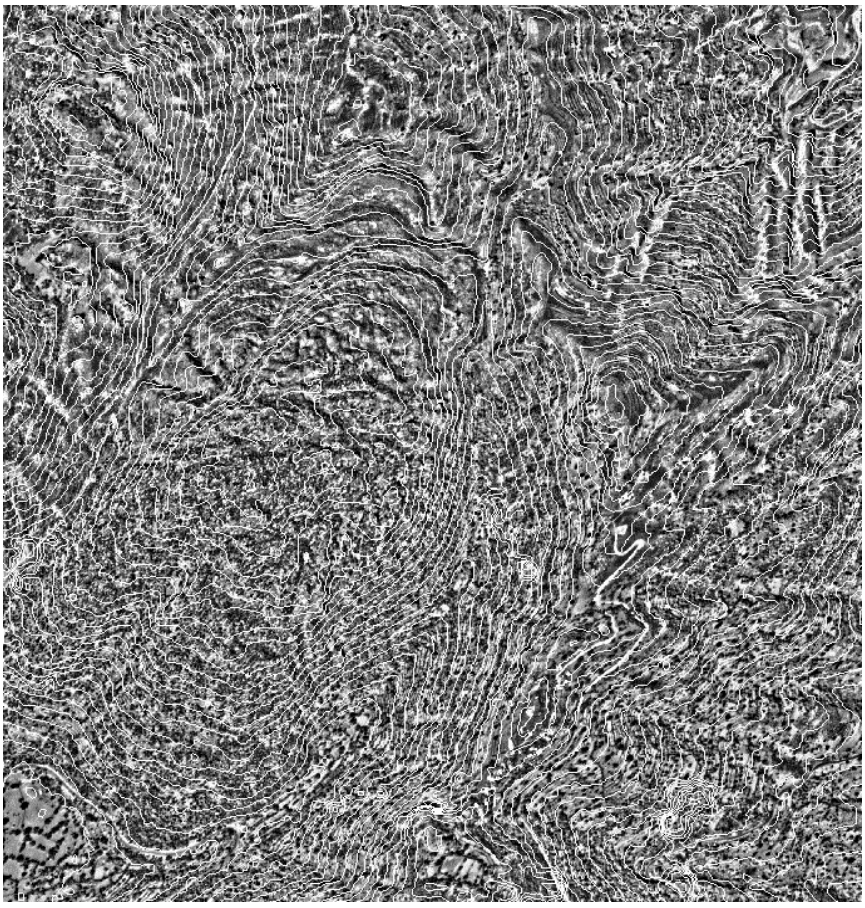

Fig. 1. An image region overlaid with contours derived from Version 3 matching results and 10 m interval.

Version	Number of Points	RMSE (m)	Standard Dev. (m)	Mean with Sign (m)	Max. Abs. (m)	% 0-1 m	% 1-2 m	% 2-3 m	% 3-4 m	% 4-5 m	% 5-8 m	% > 8 m
V. 1	492,109	3.963	3.733	1.332	172.365	27.03	23.71	17.81	11.67	7.30	9.09	3.38
V. 2	417,675	3.354	3.013	1.475	62.673	30.99	25.05	17.08	10.01	5.99	8.09	2.79
V. 3	417,060	3.186	2.883	1.356	97.385	31.66	25.68	17.58	10.26	5.89	6.90	2.04
V. 4	410,924	3.442	3.077	1.324	114.861	30.39	25.09	17.32	10.61	6.24	7.61	2.75
V. 5	451,091	5.406	5.227	1.380	332.726	30.68	25.18	17.51	10.43	6.14	7.50	2.57

Table 3. DSM accuracy quantitative evaluation. Matching minus reference data results from IKONOS, Nisyros. The percentages show frequency of absolute errors.

4. DSM Generation and Test Results using Commercial Digital Photogrammetric Systems

A DSM was generated with matching using the LHS DPW770 and VirtuoZo digital photogrammetric systems. The IKONOS images used were pre-processed with the third approach mentioned in Baltsavias et al. (2001), but without Wallis filter and for VirtuoZo with reduction to 8-bit (VirtuoZo can not process 11-bit images for DSM generation). Wallis filter removes the low frequencies, and in the image pyramid generation process, which is used by both systems, this leads to homogeneous images with little texture. The

proper procedure is to generate the pyramid levels first and then apply Wallis. This could be done with LHS but not with VirtuoZo, which does not provide access to the pyramid levels, but to allow a more objective comparison of the two systems it was not used with LHS. LHS can read the RPCs of IKONOS images, and we assume that after matching these RPCs are used to compute a spatial resection, although no information is provided on how this is done. With VirtuoZo, the pixel coordinates were used to compute object coordinates using the RPCs and spatial resection. For both systems, the geographic coordinates (RPCs use geographic coordinates) were transformed to UTM, and then the RPC bias, computed from 28 GCPs, was subtracted and a regular 2m grid was interpolated. With LHS a so-called Adaptive Automatic Terrain Extraction (AATE) method was used, associated with a modified steep_1 matching strategy which is suitable for rough terrain. Nonadaptive strategies perform an epipolar resampling of the images first. Although the IKONOS images were already epipolarly resampled, the system could not recognise this, and thus was trying to transform the images, but unsuccessfully, and thus matching with nonadaptive strategies was failing (it is unknown whether AATE performed on-the-fly epipolar resampling or not, and if yes how). With AATE, 6 pyramid levels in 8 computational passes and a patch size of 9x9 pixels at the lowest level were used to generate a regular grid with 2m spacing. LHS uses a simple 2x2 averaging to generate the pyramid levels. This fast, but qualitatively poor, method leads to a smoothing of details in higher pyramid levels. For the Melbourne city scenes, this led to images that had very repetitive signal due to the repetitive scene content with buildings and city blocks. Another filter was used for pyramid generation (7x7 approximation of the ideal sinc filter, see Baltsavias, 1991) but this, although finally used in the matching, led only to minor improvement. With VirtuoZo, the number of pyramid levels is fixed (4, including the original image, whereby pyramid levels are generated with 3x3 averaging and a reduction factor of 3). A regular image grid with 7 pixels spacing was matched, using a patch size of 11x11 pixels.

The matching results were evaluated visually by superimposing contours on the stereopair (that was extremely slow and cumbersome with LHS) and DSM-shading (see Figs. 2 and 3), but no manual editing was performed. Visually, the results of VirtuoZo looked initially better (not necessarily when comparing Fig. 2 and 3, because of the different visualisation and the lower match point density of the VirtuoZo DSM), and the contours were following the objects (e.g. buildings) in most cases nicely. It should be noted that the Melbourne scene was a very difficult case for matching. Although the terrain is relatively flat, high buildings created parallaxes of over 200 pixels. Their occlusions and shadows created additional problems. The scene objects were highly discontinuous, leading to smoothed surface results, when using area-based matching as is the case with both systems. The repetitive image content in higher pyramid levels could easily lead to multiple solutions and wrong matching. Although the images had practically no y-parallax, with LHS a 1D matching in x-direction could not be enforced, while with VirtuoZo this was possible, but only because we had access to a separate, commercially unavailable, matching module.

For a quantitative analysis, we used the aerial imagery, which was covering only a part of the IKONOS stereopair. A DSM was generated using matching with VirtuoZo. The colour images were scanned at a Vexcel UltraScan 5000 with a pixel size of 25 microns (ca. 0.40 m pixel footprint) and were first transformed to B/W using a weighted combination of the RGB channels, whereby the weights were a linear function of the noise of each channel, determined by the method described in Baltsavias et al. (2001). Then, they were pre-processed using the third approach in Baltsavias et al. (2001), including Wallis filtering (the pyramid levels had sufficient signal in this case). The images were oriented using 8 GCPs from the Melbourne test-field in UTM, although their distribution

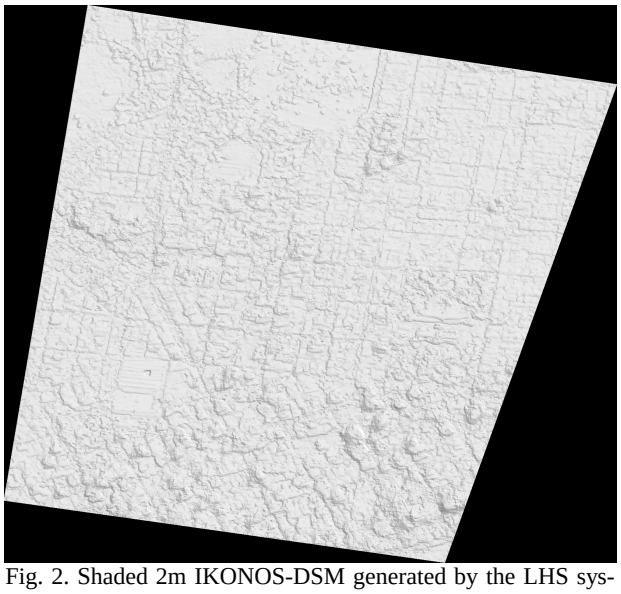

Fig. 2. Shaded 2m IKONOS-DSM generated by the LHS system with AATE. The oval University of Melbourne campus is at the top left part.

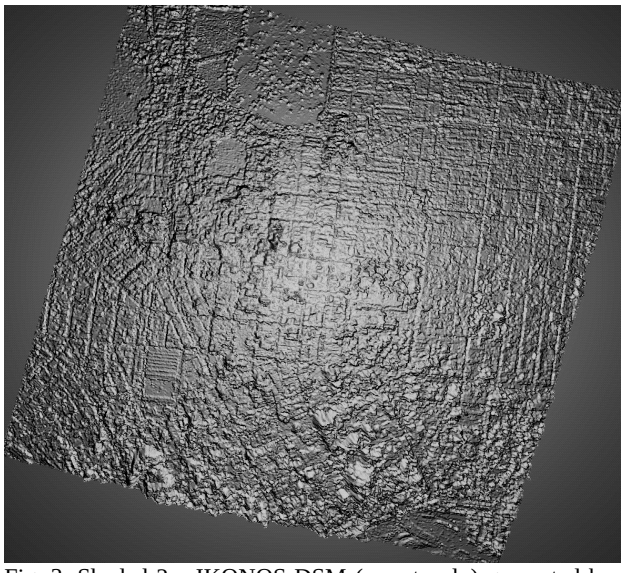

Fig. 3. Shaded 2m IKONOS-DSM (a part only) generated by VirtuoZo. The oval University of Melbourne campus is at the top left.

was not optimal, and the RMS of the orientation residuals in X, Y and Z were 0.31, 0.70 and 0.35 m respectively. GCPs were measured in one image manually and transferred to the second one semi-automatically using matching. A regular image grid with 7 pixels spacing was matched, using a patch size of 11x11 pixels, and then a 2m regular grid was interpolated. DSM matching with the aerial images had the same problems as Ikonos, with the exception of reduced repetitive patterns, due to the higher resolution of the aerial imagery. Thus, the use of this DSM as a reference dataset would not be optimal. Another problem in comparing the DSMs from aerial and IKONOS imagery were the multitemporal differences. The aerial imagery was acquired on 11/1/1993. Thus, there were many differences in buildings, roads and vegetation, while the occlusions and shadows were also different. This problem could be avoided by manually measuring in the aerial images only points that were visible in both image datasets and had remained stable (72 points). In addition, points in open areas matched with VirtuoZo were also used leading to a total number of 8627 reference points.

With LHS, a Figure Of Merit (FOM) is provided to indicate suspicious points. 22% of the matched points had poor FOM (7% had high elevation slope, 7% low (less than 0.45) correlation coefficient, ca. 3% each had large signal power difference between the two images and excessive correlation shift, and ca. 2% were elevation spikes). A dataset without poor-FOM points was also used (marked

"cleaned" in Table 4). As with previous DSM tests with LHS, the results without poor-FOM points are better but large blunders still remain in the presumably good points. A quantitative analysis of the DSMs gave the results listed in Table 4, whereby the reference data were interpolated in the raw matching data of both systems.
The results of Table 4 are to a certain extent deteriorated by some remaining errors in the reference DSM or differences between Ikonos and aerial images (e.g. due to some isolated trees existing only in Ikonos imagery). Due to the latter, the bias, as also most of the blunders, is positive since the IKONOS DSMs was higher. The statistics for VirtuoZo are worse, contrary to the visual checks, but this may be due to aerial DSM errors and the partly different DSM coverage for the two systems.

Method	Number of Points	Standard Dev. (m)	Mean with Sign (m)	Max. Abs. (m)	% 0-1 m	% 1-2 m	% 2-3 m	% 3-4 m	% > 4 m
VirtuoZo	8627	1.189	2.453	7.019	15.0	17.3	28.7	33.0	5.9
LHS	8627	1.013	2.096	7.811	13.7	19.0	36.1	25.7	5.5
LHS (cleaned)	8627	0.941	1.993	7.811	13.8	18.8	36.8	25.0	5.6

Table 4. DSM accuracy quantitative evaluation. Matching results minus reference data for IKONOS, Melbourne. The percentages show frequency of absolute errors.

5. Conclusions

This investigation has shown that IKONOS Geo stereo imagery has high geometric integrity and the potential to yield accurate DSM results, even from the cheapest Geo version, without RPCs and even with multi-temporal images. Investigations on other linear-CCD space-borne sensors have shown that under ideal circumstances raw matched DSM data can be ca. 1 pixel accurate or even less for good B/H ratio. In these tests, the conditions were far from ideal and thus a lower accuracy was achieved, which is however still remarkable. An additional problem for Nisyros was the insufficient quality of GCPs and reference DTM that could not allow achievement of accuracies at 1-m level. GCPs and reference data in Melbourne would allow that, and the fact that in these results the standard deviation was close to 1-m may show the full potential of DSM generation from Ikonos.

Acknowledgements

The provision of imagery and ground control information by the Department of Geomatics, University of Melbourne and the cooperation with C. Fraser, H. Hanley and T. Yamakawa are gratefully acknowledged. The IKONOS image of Nisyros was made available by Prof. E. Lagios, University of Athens, Greece within the EU project Geowarn (www.geowarn.org).

References

Baltsavias, E.P., 1991. Multiphoto geometrically constrained matching. Ph.D. dissertation, Institute of Geodesy and Photogrammetry, ETH Zurich, Report No. 49.

Baltsavias, E.P, Stallmann, D., 1992. Metric information extraction from SPOT images and the role of polynomial mapping functions. International Archives of Photogrammetry and Remote Sensing, Vol. 29, Part B4, pp. 358-364.

Baltsavias, E., Pateraki, M., Zhang, L., 2001. Radiometric and geometric evaluation of Ikonos GEO images and their use for 3D building modelling. Proc. Joint ISPRS Workshop "High Resolution Mapping from Space 2001", Hannover, 19-21 September, 21p. (on CD ROM). Available also at http://www.ipi.uni-hannover.de/html/publikationen/2001/isprs_cd.pdf (accessed 28 June 2002).

Dial, G., 2000. Ikonos satellite mapping accuracy. Proc. ASPRS Annual Conference, Washington D.C., 22-26 May 2000, 8 p. (on CD-ROM).

Fraser, C., Baltsavias, E., Gruen, A., 2002. Processing of *Ikonos* imagery for sub-metre 3D positioning and building extraction. ISPRS Journal of Photogrammetry and Remote Sensing 56(3): 177-194.

Grodecki, J., 2001. Ikonos stereo feature extraction - RPC approach. Proc. ASPRS Annual Conference, St. Louis, 23-27 April 2001, 7 p. (on CD-ROM).

Hancock, E. R., Kittler, J., 1990, Discrete relaxation. Pattern Recognition, Vol. 23, pp. 711-733.

Muller, J.-P., Kim, J.R., Tong, L., 2001. Automated Mapping of Surface Roughness and Landuse from Simulated and Spaceborne 1m data. In: Automated Extraction of Man-Made Objects from Aerial and Space Images (III), E. Baltsavias, A. Gruen, L. Van Gool (Eds.), Balkema Publishers, Lisse, pp. 359-369.

Orun, A. B., K. Natarajan, 1994. A Modified Bundle Adjustment Software for SPOT Imagery and Photography: Tradeoff, Photogrammetric Engineering and Remote Sensing, 60(12): 1431-1437.

Ridley, H.M., Atkinson, P.M., Aplin, P., Muller, J.P., Dowman, I., 1997. Evaluating the potential of the forthcoming commercial US high-resolution satellite sensor imagery at the Ordnance Survey. Photogrammetric Engineering and Remote Sensing, 63(8): 997-1005.

Toutin, Th., Chénier, R., Carbonneau, Y., 2001. 3D geometric modelling of Ikonos GEO images. Proc. Joint ISPRS Workshop "High Resolution Mapping from Space 2001", Hannover, 19-21 September, 9 p. (on CD ROM).

Vassilopoulou, S., Hurni, L., Dietrich, V., Baltsavias, E., Pateraki, M., Lagios, E., Parcharidis, I., 2002. Orthophoto Generation Using Ikonos Imagery and High Resolution DEM: a case study on monitoring of volcanic hazard on Nisyros island (Greece). ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 57 (submitted for publication).

Zhang, Z., Zhang, J., Wu, X., Zhang, H., 1992, Global Image Matching with Relaxation Method. Proc. International Colloquium on Photogrammetry, Remote Sensing and Geographic Information Systems, 11-14 May, Wuhan, China, pp. 175-188.