

Automatic DSM Generation from TLS 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

KEY WORDS: Three-Line Scanner, Star Imager (SI) System, Cross-correlation with Adaptive Window, Relational Matching, Multi-Photo Geometrically Constrained (MPGC) Matching, DSM, DTM

ABSTRACT:

The StarImager (SI) system is a new digital camera system developed by Starlabo, Tokyo. Due to its flexibility in imaging and GPS/INS recording and the possibility of mounting it on different aerial platforms it can be used in a great variety of applications. This paper presents a matching procedure for automatic DSM generation from the SI image data. It can provide dense, precise and reliable results. The method uses a coarse-to-fine hierarchical strategy with a combination of several image matching algorithms and associated automatic quality control procedures. The triangular irregular network (TIN) based DSMs are generated by a combination of the matching results from feature points, specific edge points and grid points. Finally, a modified multi-photo geometrically constrained (MPGC) matching algorithm is employed to achieve sub-pixel accuracy. The sensor model used for the geometric constraints is based on modified collinearity equations appended by different trajectory models.

The proposed approach in this paper has been applied to different areas with varying textures and terrain types. The accuracy test uses TLS images and is based on the comparison between manually measured DSMs from single frame aerial images and the automatically extracted DSMs from TLS, and on visual inspection of the results. Although the development of the matching system is not finalized we can report here already about some very encouraging results.

1. Introduction

With the advent of large format digital aerial cameras an increased need for reliable automated image analysis functions emerges. The three-line-scanner concept provides for triple overlap in strip direction for every image point and as such basically for fairly good reliability characteristics. In photogrammetry, the basic capabilities of image matching techniques have so far not been fully utilized yet. This paper suggests a new strategy for image matching that has been developed specifically for the new three-line-scanner sensor, but may be used with some modifications for other imaging sensors as well.

We refer here to the StarImager (SI) system, developed by Starlabo Corporation, jointly with the Institute of Industrial Science, University of Tokyo (Murai, Matsumoto, 2000; Murai, 2001; Chen et al., 2001). The first generation TLS

(Three-Line-Scanner) captures digital image triplets in along-strip mode. The imaging system contains three parallel one-dimensional CCD focal plane arrays, with 10200 pixels of $7\mu\text{m}$ each (see Figure 1). The system produces seamless high-resolution images (5 - 10 cm footprint on the ground) with three viewing directions (forward, nadir and backward). In order to get highly precise attitude data and

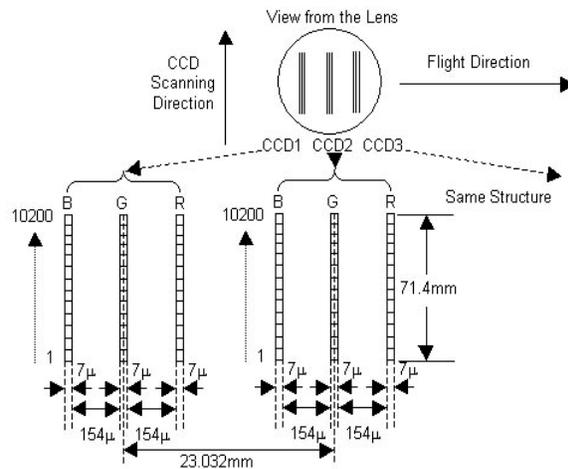


Figure 1: TLS sensor configuration

Table 1: TLS and SI-200 sensor parameters

	TLS	SI-200
Focal length	60.0 mm	65.0 mm
Number of pixels per array	10200	14404
Pixel size	$7\mu\text{m}$	$5\mu\text{m}$
Stereo CCD arrays	3	3
Multi-spectral CCD lines	3 RGB	3 RGB + 1 Infrared
Stereo view angle	21°	$21^\circ/30^\circ$ *
Field of view	61.5°	68.0°
Scan line frequency	500 HZ	500 HZ

* forward-nadir / nadir-backward stereo view angle

high quality raw-level image data from an aerial platform, a high quality stabilizer is used to stabilize the camera and outputs attitude data at 500 Hz. A Trimble MS750 serves as Rover GPS and collects L1/L2 kinematic data at 5 Hz and another Trimble MS750 serves as Base GPS on the ground.

Starlabo is currently developing a new generation camera system, called SI-200 (STARIMAGER-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 $5\mu\text{m}$ size. For the SI sensor and imaging parameters see Table 1.

Unlike with frame-based photography, the three-line geometry is characterized by nearly parallel projection in the flight direction and perspective projection perpendicular to the flight direction. The sensor model for the SI images is based on the collinearity equations and uses different forms of trajectory models. This sensor model is used for the improvement of the measured exterior orientation parameters for each scan line of SI images by a modified photogrammetric bundle adjustment, and for the derivation of the geometric constraints in our matching procedures. More details on our SI sensor model can be found in Gruen, Zhang, 2002a, 2003.

The SI system is generally designed for large-scale image acquisition; the image resolution is less than 10cm in most cases. To extract DSMs from this kind of imagery, one should specifically take into account occlusions, surface discontinuities such as very often caused by man-made objects, large areas with little or even no textures, repetitive patterns, etc.

SI imagery provides for some new characteristics and possibilities for image matching:

- The system delivers multiple images with multiple channels, it enables the multi-image matching approach, leading to a reduction of problems caused by occlusions, multiple solutions, and surface discontinuities and higher measurement accuracy through the intersection of more than two imaging rays.
- The system provides for precise interior and exterior orientation elements (especially after triangulation) that can be used to enforce geometric constraints, e.g. to restrict the search space along quasi-epipolar lines.
- Nearly parallel projection in flight direction. Therefore there are less occlusions on the nadir-view images.

The proposed method considers the new characteristics of the SI image data. Our matching goal is automatic DSM extraction from the SI raw level images. The DSMs are generated by a combination of the matching results from feature points, special edge points and grid points. A modified MPGC matching algorithm is employed to achieve sub-pixel accuracy for all the matched features. In the main parts of this paper we will report about our matching approach and software "SI-Matcher". We actually present a refinement of our matching approach shown in Gruen, Zhang, 2002b.

The method has been applied to different areas with varying textures and terrain types, including very complex city areas. The accuracy testing is based on the comparison of manually measured DSMs from single frame aerial images to automatically extracted DSMs from TLS imagery and on visual inspection of the results.

2. Matching Considerations

Automatic DTM/DSM generation through image matching has gained much attention in the past years. A wide variety of approaches have been developed, and automatic DSM 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 DTM generation are encountered with

- (a) Little or no texture
- (b) Distinct object discontinuities
- (c) Local object patch is no planar face in sufficient approximation
- (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 DTM

SI-Matcher aims to generate DSMs by considering specifically the problems (a)-(f). Figure 2 shows the strategy of our matching approach. We use the raw level SI images and the given or previously triangulated orientation elements. After production of image pyramids, the matcher uses three kinds of image features, i.e. general feature points, edge points and grid points. 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 approximation 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. Details of our approach will be described in chapter 3.

Among the usual matching techniques, area-based matching (ABM) and feature-based matching (FBM) are the two main ones applied in automatic DSM generation, but also relational matching is sometimes used. All basic matching techniques have 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. But even then the lack in image understanding capability will lead to problems, whose impact must be judged by the project specifications. Our matching approach is a hybrid method that combines ABM and relational matching. It uses a coarse-to-fine hierarchical strategy with a combination of several image matching algorithms and automatic quality control (see workflow in Figure 2).

We employ ABM (both in form of cross-correlation and least squares matching) to match feature points and grid points. Generally the performance and success rate of ABM mainly depends on the existence of sufficient image texture, the quality of the approximations and a set of matching parameters, such as the matching window size, the search distance and the acceptance threshold for the correlation coefficient. How to select a set of correct matching parameters is problematic, because the requirements for these parameter values are conflicting. These matching parameters are functions of many factors, including terrain type, image texture, image scale, disparity variations and image noise. SI-Matcher uses a set of adaptively determined matching parameters. This is done by analyzing the results of the higher-level image pyramid matching and using them at the current pyramid level (for details see chapter 3.2).

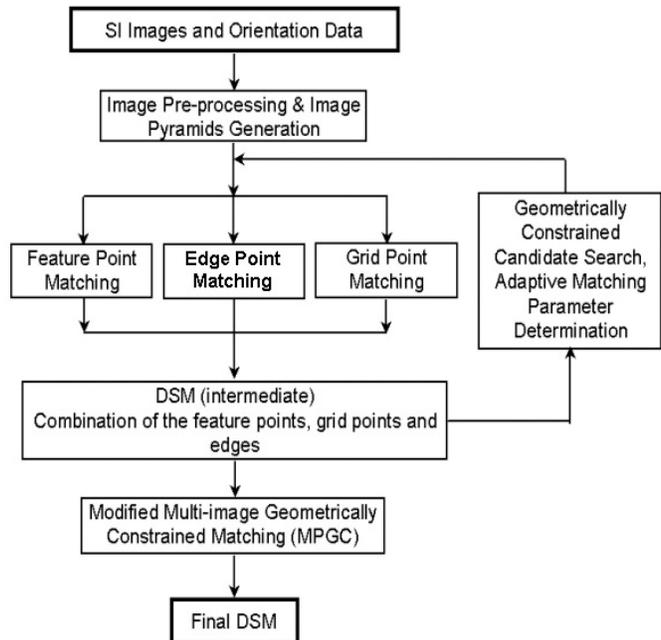


Figure 2: Workflow of the SI-Matcher

The performance of ABM is low if there is not enough image texture, in case of repetitive patterns and at surface discontinuities. Unfortunately these problems are very typical for large-scale images as provided by SI. In the first case, because of missing points, the ABM may lead to holes in the DSM. To overcome this problem, a global image matching technique, based on relaxation, is employed to match grid points, in order to bridge the poor texture areas. This essentially relational matching uses local smoothness constraints. In the last case, ABM generates smoothing effects on the surface discontinuities. Edge point matching is employed to match the edges located on the surface discontinuities and, these matched edges are used as break-lines to control the weights of the surface smoothness constraints in the global image matching. The smoothness constraints are prohibited to cross over edges.

The epipolar curves derived from the SI sensor model are used to restrict the search range to only one direction. The residual y-parallax grid (see 3.5 for details) is used to compensate some of the errors of the raw image data.

3. The Matching Approach

3.1 Input Data and Pre-processing

The input data includes the SI strip which normally consists of three images of the forward, nadir and backward view CCD arrays, and information about their interior and exterior orientation parameters (for details see Gruen, Zhang, 2002a, 2003). SI-Matcher employs the raw level SI images, because a high quality stabilizer is used in image acquisition. The Wallis filter is used to reduce the image noise and at the same time enhance the weak texture patterns. The image pyramid starts from the original resolution SI 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.

3.2 Feature Point Matching

Feature points are generally more or less randomly distributed over the image. We use the Foerstner interest operator to extract well-defined feature points in the nadir image. For that the nadir image is divided into small image patches and only one feature point will be extracted per image patch. In our implementation, the threshold for the parameter roundness is set to 0.50, the gray value variance of the image window is not allowed to drop below 8. These relatively low threshold values make the distribution of the collected feature points as uniform as possible. The density of the feature points can be controlled by the size of the image patches.

The determination of the correspondences to the other view images is carried out using the geometrically constrained cross-correlation method (see Figure 3). Given a point on the nadir image, an image ray that connects the instant perspective center and this image point can be determined. Given an elevation approximation Z_a and an elevation variation ΔZ , the coordinates of three object points $(X_u, Y_u, Z_a + \Delta Z)$, (X_a, Y_a, Z_a) and $(X_l, Y_l, Z_a - \Delta Z)$ can be computed by using the pixel coordinates and orientation elements. By projecting these three object points back to the forward and backward view images, search windows along the epipolar curves can be determined. These search windows are assumed to be rectangular for a small region in first approximation. Their width is ± 5 pixels. The correspondences are computed by cross-correlation with the adaptively determined image window size w_s , threshold of the correlation coefficient c_t , and within the search window with a distance related to terrain elevation variation ΔZ . The TIN based approximated DSM data that was derived from the higher-level image pyramid is used to estimate these parameters adaptively.

The SI-Matcher incorporates a method to select an optimal window size by evaluating the local variation of the image intensity and the disparity through a statistical model of the disparity distribution within the matching window, as proposed by Kanade, Okutomi, 1994. The variations of the disparities can be calculated from the elevation variation ΔZ through the image geometry and for the variations of the image intensities we choose the summation of the squared first derivatives of the intensity values. As a result, in flat areas with small intensity variations, the window size w_s increases, while in areas of large terrain elevation variations the window size w_s decreases. The threshold of the correlation coefficient c_t should also change according to the variation of the terrain elevation (curvature 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 ΔZ and the image geometry. Let $\Delta Z'$ be the computed elevation variation from the approximated DSM in a small neighborhood of the given point, we choose ΔZ is two or three times of the value $\Delta Z'$. Therefore, the size of the searching window decreases in flat areas and vice versa. By adaptively selecting the search window size, we can both reduce the processing time and increase the probability for correct matches. However, we cannot completely avoid the ambiguity problem due to reasons like repetitive image patterns. SI-Matcher takes n (≤ 5) matching candidates, whose correlation coefficients are above the threshold c_t , by sorting the candidates according to the size of the correlation coefficients.

As a result, for each feature point on the nadir-view image several matching candidates can be obtained. The correct match is determined by considering the following quality measures sequentially:

- a) Size of correlation coefficient. If the first correlation peak is more than two times that of the secondary peak, the first candidate should be the correct match.
- b) Using the same matching parameters, a feature point can be matched from the other images to the nadir view image. If the differences between this inverse matching and the normal one is less than 1.5 pixels, the candidate should represent the correct match.

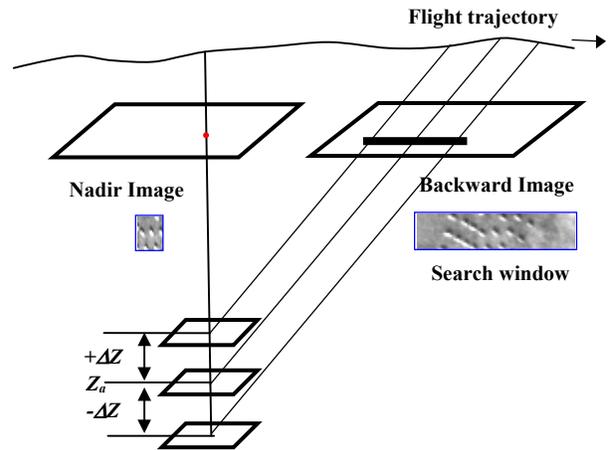


Figure 3: Determination of the search window for cross-correlation of given points

c) Under the condition that the feature point appears on all three images, the standard deviations of the object space coordinates of the photogrammetric forward intersection should not be larger than 2-3 times the standard deviations of the triangulation adjustment.

Measures (a) and (b) can be performed on both pairs of the three SI images: On the forward- and nadir-view image (F+N) and on the nadir- and backward-view images (N+B). The point that passes test (a), (b) on both (F+N) and (N+B) and test (c) is the correct and unique match. The point that only passes test (a) and (b) on either (F+N) or (N+B) is also taken as successful point because it may appear only on two images. The point that cannot pass these tests may have multiple solutions, and at this point we cannot determine the correct match. This matching ambiguity will be solved by the following global image matching method through the use of local smoothness constraints (see chapter 3.4).

3.3 Edge Point Matching

In images of urban areas, edges play a major role, either as representations of surface discontinuities or as distinct gray level gradients on otherwise smooth surface patches. Therefore any successful matcher must pay special attention to this fact. In doing so we must consider the following problems that appear frequently in images of urban areas:

- a) Edges may break up into more than one segment due to noise, occlusions and the deficiencies of the feature extraction algorithms
- b) Conjugate edges on different images may have different shapes due to the projection distortions
- c) There may exist many similar features in a particular search area.

This is why we decided at this point not to match extracted edges as such but dominant points on the edges. These points are characterized by large local curvature values. Once the point is selected matching is done with ABM using variable size matching windows. The disparities of the other edge pixels are then interpolated by linear interpolation based on the reasonable assumption that the disparities along the edges will vary smoothly.

The Canny operator is used to locate the intensity discontinuities and the edges are traced through a local processing that analyses the characteristics of these pixels in a small neighborhood. This approach is carried out independently on three images. For each edge on the nadir-view image a set of dominant points are computed. For long straight lines we select some evenly distributed edge pixels as dominant points.

The dominant points along the edges on the nadir-view image are matched with the intersections between the candidate edges and the epipolar curve within the search windows on other view images (Figure 4). The search window can be determined by using the same method as that in chapter 3.2. Normally there are several matching candidates within the search window. To solve this ambiguity problem we could use the standard cross-correlation method, but this method may fail since it has very low correlation coefficients due to local discontinuities and/or different illumination conditions. SI-Matcher uses therefore the following adaptive cross-correlation procedure:

1) Computation of the dominant points on the edges
 2) Selection of a suitable window (size & shape)

2.1) Start with a small 9×9 window centered at the dominant point, computation of the correlation coefficients between the given point and its matching candidates. Exclude those candidates that have a very low correlation coefficient (threshold e.g. 0.4).

2.2) Expansion of the window by 8 pixels in one direction, e.g. to the right, and computation of the correlation coefficients for the expanded window. If the expansion decreases the correlation coefficients significantly for all the candidates ($\geq 30\%$), the direction is prohibited from further expansions. Repeat the same process for the other three directions.

2.3) Repetition of step 2.2 until the window size reaches a limit.

3) Determination of the correct match for the given dominant point by using the quality measures mentioned in chapter 3.2.

4) Iteration of steps 2 and 3 for all dominant points, taking those passing the quality measures as successfully matched points and recording the size and shape of their corresponding matching window (for the purpose of later refinements through the modified MPGC matching).

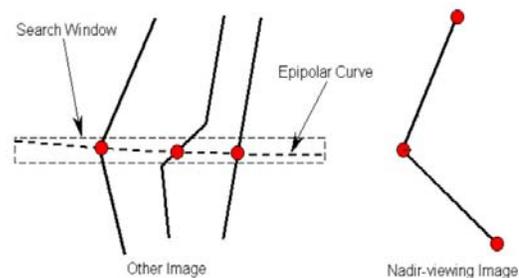


Figure 4: Dominant edge point matching. Matching candidates in the search window

5) Interpolation of the disparities for each edge pixel from the successfully matched dominant points. Figure 5 shows two examples of our edge matching, illustrating the problems (a), (b), (c) mentioned above.

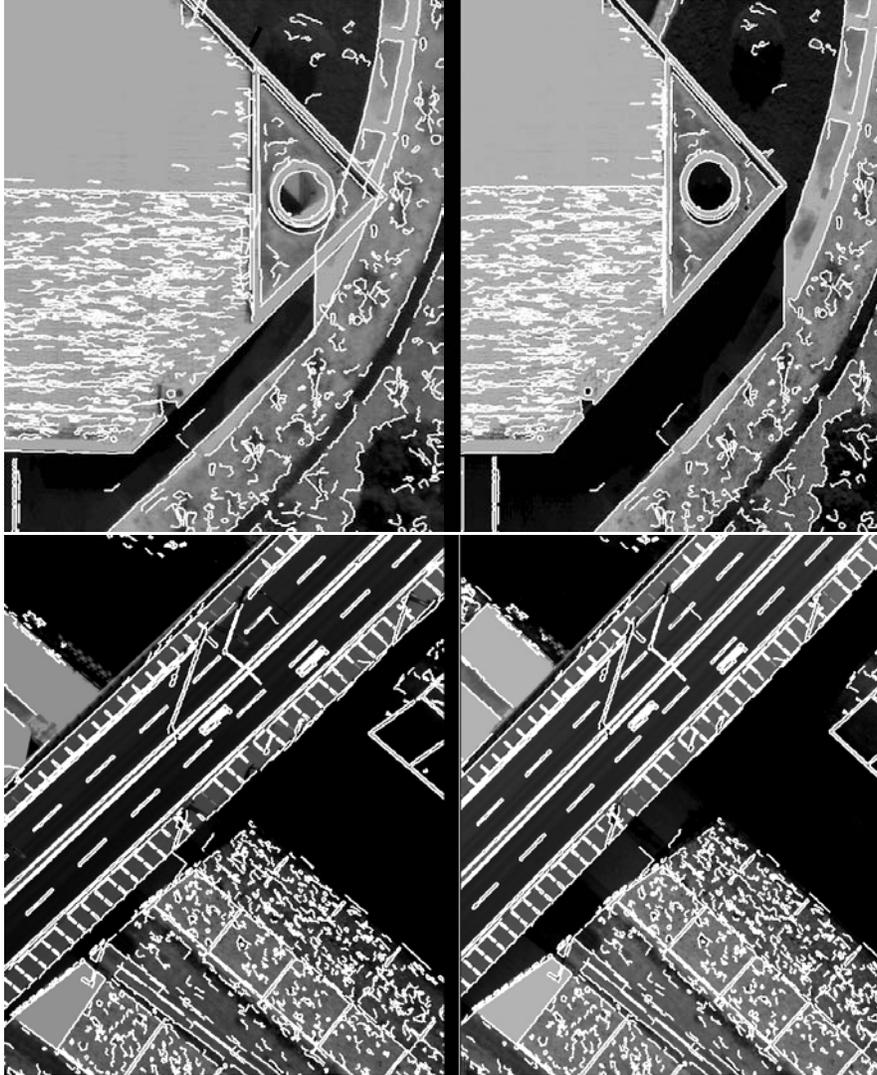


Figure 5: Two examples of the edge matching (forward- and nadir-view images)

3.4 Grid Point Matching

The purpose of grid point matching is primarily to create uniformly distributed points over the whole images even in the very little or non-texture areas. The correspondences of these grid points could be computed by using the method presented in 3.2. Compared to feature points, the choice of grid points is blind and thus many grid points are located in areas with weak or no texture. Therefore the probability for multiple candidates or no candidate at all increases for such points.

To solve this problem, we use a global image matching method based on relaxation technique. This method may be interpreted as relational matching. It 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 (Hancock, Kittle, 1990; Zhang, et al., 1992). Here we use Prazdny's "coherence principle" model (Prazdny, 1985), which states that the neighboring features should have similar disparities if they correspond to the same 3D object. These features should support each other, while features with dissimilar disparities should not interact with each other. We incorporate this principle in our global image matching and solve it by relaxation.

Firstly, the points are selected in form of a regular grid in the nadir-view image. Their matching candidates on other images are computed. Together with all the previously matched feature and edge points we construct a TIN, whereby the image edges are treated as discontinuities in Delauney triangulation. All the used points can be categorized into three classes: The points that having unique and correct matches, the points which did not pass the quality measures presented in 3.2 but have several candidates, and the points that have no matching candidates at all. For the first case, they are treated as having only one matching

candidate, for the last case, they will be given several “false” candidates (with a very small correlation coefficient value) that are evenly distributed within the search window.

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

In order to link the matching results of the neighboring points to each other, we define the following compatible coefficient function $C(i,j;k,l)$, which quantifies the compatibility between the match $I_i \leftarrow I_j$ and a neighboring match $I_k \leftarrow I_l$, according to Prazdny’s rules (Prazdny, 1985).

$$C(i,j;k,l) = \frac{T}{\beta d_{ik}} \exp\left(\frac{\Delta p^2}{\beta^2 d_{ik}^2}\right) \quad (1)$$

$$\text{where } \Delta p = (x_j - x_i) - (x_l - x_k)$$

In equation (1), Δp expresses the difference of the x-parallaxes in point I_i and its neighboring point I_k . d_{ik} is the distance of two points I_i and I_k . The bigger Δp or the larger d_{ik} , becomes the smaller the compatibility will be. This corresponds to a smoothness constraint on the image matching results. T is a weighting factor to control the continuity of the terrain surface. If the point I_i is located on an edge or if the connecting line between I_i and I_k crosses the edge, T will be set to a very small value, otherwise T equals 1 (Figure 7). β is a scaling value and is set to a constant experimentally.

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

$$P^{(n+1)}(i,j) = \frac{P^{(n)}(i,j)Q^{(n)}(i,j)}{\sum_{s=1}^m P^{(n)}(i,s)Q^{(n)}(i,s)} \quad (2)$$

$$\text{where } Q^{(n)}(i,j) = \prod_{I_k \in \Omega(I_i)} \sum_{l=1}^m P^{(n)}(k,l)C(i,j;k,l)$$

$C(i,j;k,l)$ is the compatible coefficient function defined as above, $\Omega(I_i)$ is the neighbourhood of point I_i , and n is the iteration number. The quantity $Q^{(n)}(i,j)$ expresses the support which the match $I_i \leftarrow I_j$ receives at the n^{th} iteration step from the matches $I_k \leftarrow I_l$ in its neighbourhood $\Omega(I_i)$.

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 $P(i,j)$ ($j=1, \dots, m$) exceeds $1-\varepsilon$, where $\varepsilon \ll 1$ (for example, $\varepsilon = 0.1$).
- The predefined number of iterations has been reached.

When the iterative procedure is terminated, the match which gains the highest probability $P(i,j)$ ($j=1, \dots, m$) is selected as the actual match.

This method is performed by using stereo pairs, in a combination of the forward- and nadir-view and the nadir- and backward-view images.

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. At the same time the surface discontinuities can be preserved because the smoothness constraints are deactivated across edges (see Figure 6).

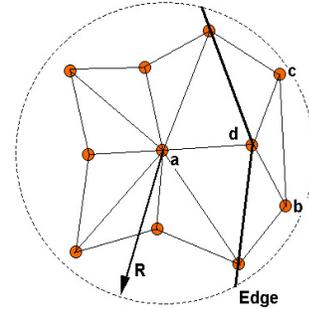


Figure 6: A neighborhood of the point **a** with radius **R**

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 Delaunay triangulation is used to construct the TIN in order to connect all the matched feature points, grid points and the dominant points on the edges. The matched edges are used as break-lines such that there is no triangle that crosses the edges. The TIN maintains the original matching results without any interpolation and the surface discontinuities of the terrain can be preserved.

The initial DSM of the highest level of the image pyramid can be extracted by standard cross-correlation based on a “region growing” matching strategy. This method uses the already measured control points as

seed points and matches the other 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.

Residual y-parallaxes caused by the imprecise SI sensor model and orientation elements degrade the success rate of the matching significantly (Zhang, Miller, 1994). To compensate for the residual y-parallax, we use residual y-parallax grids, which are derived from reliable feature point matching results on the higher pyramid level. The residual y-parallax for each point on the current pyramid level is interpolated from those residual y-parallax grids.

3.6 Refined Matching based on the Modified MPGC

MPGC (Multi-Photo Geometrically Constrained Matching) was developed by Gruen, 1985; Gruen, Baltsavias, 1988, and is investigated in detail in Baltsavias, 1991. It combines least squares matching and geometric constraints formulated either in image or in object space. The collinearity constraints lead to a 1D-search space along epipolar lines, thus to an increase of success rate, precision 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 the adaptation of the matching parameters to the image and scene content.

Our modified algorithm is an extension of the standard MPGC. It integrates the geometric constraints derived from the SI sensor model (for details see Gruen, Zhang, 2002b). The geometric constraints force the matching to search for a conjugate point only along a small band around the epipolar curve. If the initial match of the point in the search images does not lie on this epipolar curve, it jumps onto this curve at the first iteration of MPGC. Figure 7 shows an example.

The modified MPGC is used to refine the matched feature and edge points in order to achieve sub-pixel accuracy. The grid points are not treated accordingly. The DSM derived from the approaches (3.2)-(3.4) provides quite good approximations for the MPGC procedure and thus increases the convergence rate.

The initial values of the shaping parameters (here we use the affine transformation) in MPGC can be predetermined by using the SI sensor model and the derived DSM data. The corresponding rays of the four corners of the matching window in the nadir-view image are intersected with the derived DSM and the corresponding object coordinates are determined. Through the imaging geometry, the corresponding image coordinates in other images can also be determined. The initial values of the shaping parameters can thus be computed from these four corner points and their correspondences. The matching window parameters, including the size and shape, can be used to generate a window mask in order to exclude some pixels in the matching window. This is important when performing modified MPGC matching in cases where points are located on edges or if occlusions are involved.

In our implementation, at the first two iterations, the weight for geometric constraints takes a large value in order to speed up the convergence and then it decreases to consider residual errors in the orientations.

This method is employed to refine the matching results of the feature points, the grid points and the dominant points on the edges. 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 reliable indicator values.

4. Experimental Results

In order to evaluate the performance of our approach for DSM generation, several TLS data sets with different terrain type and different image textures are used. Here we report about the results of Japan's GSI (Geographical Survey Institute) test area. For some other, earlier results see Gruen, Zhang, 2002b. Both SI image and aerial photos are available in the test area. The evaluation is based on the comparison between the manually measured DSM from aerial photos and the automatically extracted DSM from the SI images.

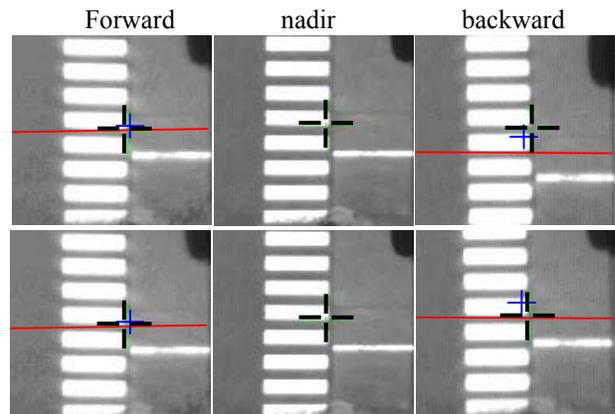


Figure 7: MPGC matching without (above) and with (below) geometric constraints. The line is the corresponding epipolar curve. The small crosses are the initial positions and the big crosses are the final positions

The GSI test area is roughly $650 \times 2500 \text{ m}^2$. The nadir image of Figure 8 gives an overview of the test area. It is relatively flat with natural and man-made objects. There are a lot of small geomorphological features, small discontinuities like cars, isolated trees, large discontinuities and occlusions due to buildings (some of them are over 20 meters high with some complex small infrastructures on top). Also, there are occlusions due to car and people motions. Figure 9 shows two image windows from the nadir-view image. All 48 control points are signalized marks on the ground or on the top of buildings, they both appear in the TLS images and aerial photos. The control points were measured using GPS and conventional total stations. The obtained accuracy was reported as 2 cm for the horizontal and 3 cm for the vertical components.



Figure 8: TLS Image from the GSI strip
(nadir-view image with a ground resolution of 5.6 cm)

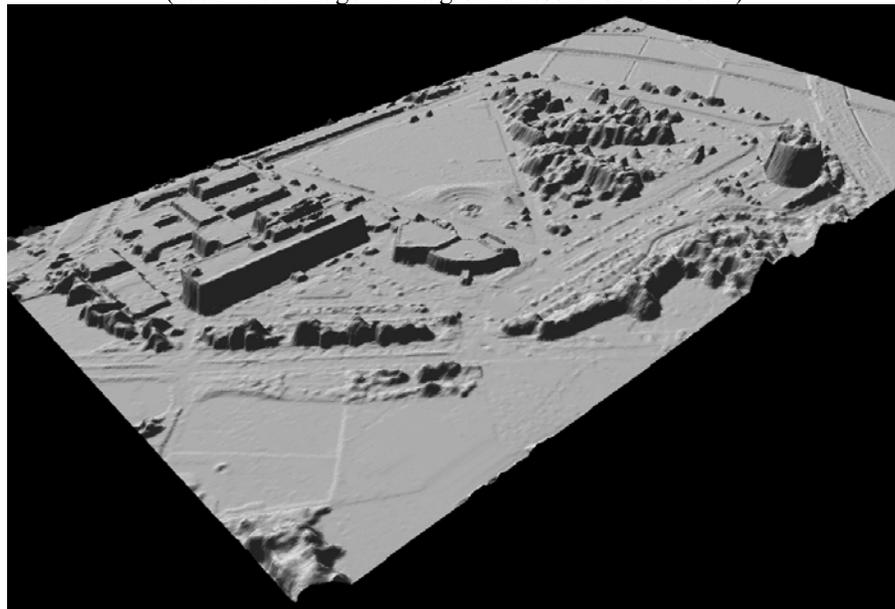


Figure 9: 3D visualization of the shaded DSM of the GSI test area

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 the Analytical Plotter Wild AVIOLYT AC-3. The RMS of the exterior orientation is reported as 5.1 cm in planimetry and 3.3 cm in height. For reference data a 50 cm grid of points has been measured manually.

The processed TLS data includes three panchromatic images with a footprint of 5.6 cm. For the determination of the position and the attitude values for each image scan-line, a photogrammetric triangulation procedure which models the flight trajectory with piecewise polynomials is performed with the help of the signalized control points. The tie points are measured using semi-automatic least squares matching. As a result of triangulation, 2.8 cm and 5.0 cm absolute accuracy in planimetry and height are obtained.

For the analysis of the matching accuracy we divide the reference data into four classes: (1) Bare terrain, (2) points on top of buildings, (3) buildings, sparse trees, cars and (4) whole area with mixed objects.

For processing we follow the matching approach described in chapter 3. From feature, edge and grid points a very dense raster DSM with 15 cm ground interval was interpolated. The points that failed in the modified MPGC were given a small weight value in the interpolation procedure. Figure 9 shows the 3D visualization of the extracted DSM.

By visual inspection of the DSM one can see that the small geomorphological features are extracted, the little-texture areas are bridged smoothly, and the surface discontinuities are well preserved.

For comparison the reference raster DSM points were interpolated to the TLS point positions. The results, split up into four classes, are shown in Table 2. With a RMS of 12 cm the accuracy for bare terrain area is very high, given the fact that the accuracy of the manually measured reference data is of the same size. The accuracy of top-of-building points and tree and car points is degraded, but this has to be expected. Especially in the latter cases the systematic errors caused by wrong object definition (DTM in reference data, DSM in TLS data) let the results degrade. Also, as evidenced by the results of Table 2 the distribution of errors depends on the type of terrain and objects. For bare ground we get very short tails, why in the other cases the numbers of large errors (blunders) increase substantially. The large blunders (> 4 m) can all be found in the vicinity of buildings and trees. Here the problem is more that of 2.5D surface modeling rather than matching. Also, comparing heights instead of surface normals causes errors especially in cases where small planimetric errors play lead to big height deviations.

Area	Point Number	RMS (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)	188856	0.12	0.02	99.80	0.05	0.00	0.00	0.00
(2)	50879	0.30	-0.03	96.81	2.47	0.56	0.07	0.01
(3)	89456	0.56	0.06	95.59	3.09	0.66	0.19	0.05
(4)	394159	0.44	0.01	97.49	1.43	0.38	0.32	0.28

5. Conclusions

In cooperation with Starlabo Corporation, Tokyo we are developing a suite of new methods and software for the processing of StarImager (SI) data. In this paper we have reported about our current matching approaches for fully automated DSM generation. The methods and results are of preliminary nature and will be refined further. We have developed a matching strategy combining feature point and variable window edge point matching with relational matching based on grid point matching with neighborhood smoothness constraints. In the last step 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 SI sensor model, which has been published in Gruen, Zhang, 2002a, 2003. This strategy allows us to bridge areas with little or no texture and at the same time maintain the important contribution of object/image edges.

As evidenced by a visual inspection of the test results we can reproduce also small geomorphological features. For bare ground areas we obtain an RMS accuracy of 12 cm. With RMS values of 30 cm and 56 cm the accuracy is reduced for top-of-building points and tree/car areas. This is to be expected, not so much because of matching errors, but because of the used approach of 2.5D surface modeling and the associated comparison of heights instead of surface normals. This introduces methodological errors into the test procedure.

The results from this accuracy test indicate that the presented concept has some potential for future refinements, with respect to the matching algorithm but even more with respect to using a technically sound method of comparison of results from different data sets.

After all, and especially in city modeling, the produced point cloud will rarely be the final result. Any method of matching has also to be evaluated with respect to its capability of generating data which can be turned into structured 3D object models without too much extra effort.

Acknowledgment

The authors would like to thank Starlabo Corporation, Tokyo for their project support and provision of the test SI image data sets.

References

- Baltsavias, E. P., 1991, Multiphoto Geometrically Constrained Matching. Dissertation, IGP, ETH Zürich, Mitteilungen No. 49, 221 pages.
- Chen, T., Shibasaki, R., Morita, K., 2001. High Precision Georeference for Airborne Three-Line Scanner (TLS) Imagery. 3rd International Image Sensing Seminar on New Developments in Digital Photogrammetry, Sept. 24-27, Gifu, Japan, pp. 71-82
- 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., Bühner, Th., 2000: DTMs Derived Automatically From DIPS - Where Do We Stand? Geoinformatics, Vol.3, No.5, July/August, pp. 36-39
- Gruen, A., Baltsavias, E. P., 1988, Geometrically Constrained Multiphoto Matching. Photogrammetric Engineering and Remote Sensing, Vol. 54, No. 5, pp. 633-641
- Gruen, A., Zhang, L., 2002a. 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., 2002b. Automatic DTM Generation from Three-Line-Scanner (TLS) images. IAPRS, Vol. 34, Part 2A, Graz, Austria, pp. 131-137
- Gruen, A., Zhang L., 2003. Sensor Modeling for Aerial Triangulation with Three-Line-Scanner (TLS) Imagery. Journal of Photogrammetrie, Fernerkundung, Geoinformation (PFG), 2/2003, pp. 85-98
- Hancock, E. R., Kittler, J., 1990, Discrete relaxation. Pattern Recognition, Vol. 23 pp. 711-733
- 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., 2001. Development of Helicopter-borne Three Line Scanner with High Performance of Stabilizer and IMU. 3rd International Image Sensing Seminar on New Development in Digital Photogrammetry, Sept. 24-27, Gifu, Japan, pp. 1-3
- 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.
- Prazdny, K., 1985. Detection of binocular disparities. Biological Cybernetics, Vol. 52, pp. 93-99
- Zhang, Z., Zhang, J., Wu, X., Zhang, H., 1992, Global Image Matching with Relaxation Method. Proceedings of the International Colloquium on Photogrammetry, Remote Sensing and Geographic Information Systems, 11-14 May, Wuhan, China, pp. 175-188
- Zhang, B., Miller, M., 1997. Adaptive Automatic Terrain Extraction. SPIE Vol. 3072, pp. 27-36