Automatic generation of digital surface models from IKONOS stereo imagery and related application

Abdalla Alobeid, Karsten Jacobsen

Institute of Photogrammetry and GeoInformation, Leibniz University Hannover Nienburger Str. 1, D-30167 Hannover, Germany [Alobeid, Jacobsen]@ipi.uni-hannover.de GORS 16th International Symposium, Damascus, Syria, 2008

Key words: very high resolution space sensors, image matching, DSM, accuracy

Abstract:

Digital surface models are an important tool for the generation of 3D-city models. With high resolution optical space sensors, like IKONOS, stereo pairs can be generated, which can compete to traditional aerial photos. By automatic matching of images having a ground resolution of 1m, it is not so easy to get sharp outlines of buildings. The traditional least squares matching leads to high accuracy of the object height, but the building shapes are blurred. So as alternative the automatic matching by dynamic programming has been investigated. It requires epipolar images for the reduction of the matching task from two-dimensional to just matching in the epipolar line. One of the used stereo pairs has been delivered by GeoEye as epipolar stereo pair, the other epipolar pair was made with an own program. By dynamic programming occlusion areas have been respected precisely, leading to sharp building outlines. The matching just in epipolar lines caused some problems of striping in the generated disparity image. This has been strongly reduced by median filtering of the disparity image just in the vector perpendicular to the epipolar line, so finally accurate and precise digital surface models could be generated in build up areas.

1. Introduction:

Digital surface Models (DMSs) are geometric descriptions of the height of the visible terrain surface, including elements like buildings and trees. DSMs became an important source for many applications such as change detection in urban area, scene analysis and urban planning.

The possibility of high resolution satellite sensors like IKONOS with 1m ground sampling distance (GSD) and Quickbird with 0.6m GSD to change their view direction against the orbit without loss of accuracy opens a new era for using such images instead of aerial images or laser scanner measurements for the generation of digital surface models (DSMs) of large areas in a short time with an acceptable accuracy to extract details on the surface.

These attributes make high resolution satellite imagery suitable for DSM generation, therefore, the need for a method of reliable image matching, prompted researchers to develop varying methods. The main problem of image matching in urban areas is caused by sudden changes of elevation and occlusions.

Several methods gave satisfactory results in urban areas from high resolution satellite imagery, (Poon et al., 2007; Krauss et al., 2005, 2008; Zhang, 2002, 2004; Zhang and Gruen, 2006; Jacobsen, 2006; Büyüksalih and Jacobsen, 2007). All the methods have some difficulties and limitations.

Here we compare two matching methods for generation DSMs in urban areas based on high resolution satellite imagery. The first method is the matching by least squares with region growing while the second method is based on the concept of dynamic programming.

2. Sensor orientation:

This step tries to find the relation between the ground coordinates and its corresponding position in the image. There are different orientation methods like geometric reconstruction, bias corrected rational polynomial coefficients (RPCs) and replacement models like 3D affine transformation.

Today most satellite images are distributed together with sensor oriented RPCs. These coefficients describe the relation between ground coordinates as normalized geographic coordinates and image coordinates (Grodecki, 2001, Jacobsen, 2005). Here we used the bias corrected sensor oriented RPCs with the Hannover program RAPORI as well as the geometric reconstruction with the Hannover program CORIKON.

3. Used test data:

The first test area covers a part of Maras city in Turkey. The city is located in rolling terrain, containing densely build up parts and some relative separate buildings with heights up to 45m. The IKONOS Geo stereo pair has been taken from the same orbit (table 1).

Image name	po_34003_pan_0000000.tiff	po_34003_pan_0010000.tiff	
Date of acquisition	2000-03-30 08:07	2000-03-30 08:07	
Pixel Size	1.00 m	1.00 m	
Scene size [km]	13.5 x 8.6 km	13.5 x 8.6 km	
Scene [pixel]	13680 x 8624	13680 x 8624	
Nominal Collection Elevation	78.30753°	82.94680°	
Nominal Collection Azimuth	230.3915°	267.6522°	
Sun Angle Elevation	50.80870°	50.77190°	

Table1: specification of IKONOS stereo pair, Maras city

The other test area is located in Istanbul, Turkey. It contains densely build up regions. The IKONOS stereo pair, taken from the same orbit, was achieved as epipolar-resampled Geo PAN Images (table 2).

Image name	Left scene	Right scene	
Date of acquisition	2005-05-16 09:05 GMT	2005-05-16 09:05 GMT	
Pixel Size	1.00 m	1.00 m	
Scene size [km]	13.7 x 11.5 km	13.7 x 11.5 km	
Scene [pixel]	13908 x 12212	13908 x 12212	
Nominal Collection Elevation	74.76850°	68.24700°	
Nominal Collection Azimuth	198.78960°	3.6940°	
Sun Angle Elevation	65.51112°	65.51112°	

Table2: specification of IKONOS stereo pair used in the DSM generation, Istanbul city

Within each stereo model a number of test areas with different characteristics have been selected.

4. Matching methods

The goal of automatic image matching is to determine 3D locations of the matched points in object space, building a digital surface model (DSM). Under optimal conditions the buildings should show their 3D-location as it is in the object space, but caused by occlusions and the size of used sub-matrixes the building form may be blurred and the buildings may look like hills. Many Matching algorithms have been published, but up to now no perfect automated solution has been found.

Here the following matching methods have been used:

- 1. Least squares with region growing
- 2. Dynamic programming

4.1 Least squares with region growing

The region growing strategy, published by (Otto and Chau.1989, Heipke et al., 1996), based on few seed points leads to the required start values for the Least squares matching.

We here make the assumption that the context of the templates are homologous.

Let $d'_i(x, y)$ be the density value of the template in left image and $d'_i(x, y)$ the density value of the submatrix in right image, the least squares method tries to find the geometric and radiometric transformation parameters of one of sub-matrix to the other.

$$d'_{i}(x, y) - v_{i}(x, y) = r_{o} + r_{1} \cdot d''_{i}(x', y')$$
(1)

$$x' = a_0 + a_1 x + a_2 y (2)$$

$$\mathbf{y}' = b_o + b_1 \mathbf{x} + b_2 \mathbf{y}$$

$$v_i(x, y) = r0 + r1.d''(a0 + a1.x' + a2.y', b0 + b1.x' + b2.y') - d'(x',y')$$
 (3)

$v_i(x, y)$: Residuals of gray level differences between template and search window.

The position and relation of the matching windows are changed until the gray level difference between the transformed window and the template reach a minimum. The image matching is based on Hannover program DPCOR, which has been developed by (Heipke and Jacobsen 2008). The size of the used sub-matrixes is specified in a control file as window size, and the determined conjugated points are located in the center of the window. The similarity of the transformed window and the template is described by the correlation coefficient. As tolerance limit for the correlation coefficient the value 0.8 was used in first test area and 0.6 in second test area.

The size of the sub-matrixes for matching in both test areas was 10 by 10 pixels. Every pixel in line and also column direction has been matched, which leads to high correlation of neighbored points. The size of the sub-matrixes has been found as optimal, while smaller sub-matrixes are influenced by noise, although they may give a more precise building shape. On the other hand larger sub-matrixes are smoothing the DSM. These parameters are based on intensive tests (Jacobsen, 2006, Alobeid, 2008).

The properties of the IKONOS image combination in the first test area, such as the small base length, same orbit and same sun elevation resulted in very high correlation coefficients and a good success rate of matching between 89% and 94% of the possible points. Moreover, between 86% up to 93% of the points, showed a correlation coefficient exceeded 0.95. The shown results are based on the software Surfer 8 from Golden Software; the DSMs have been interpolated by Delauny triangulation to a point raster with 1m spacing. The results have been improved by median filter with 3x3 pixels. Figure 3 shows results obtained in the first test area.



Figure 3: DSM generated by least squares matching in Maras city (window 540x465 points)

As result the least square matching usually provides dense and accurate disparity maps from stereo scenes with sub pixel accuracy and only few blunders. Unfortunately, the fixed template size is not able to track building outlines, which will be deformed into blobs, because part of template size may show the street level

or façade of building and the other part may show the roof top of the building, where LSM will take an average height between top of building and street level. In other words, only one disparity value, representing the object height is determined in the window by least square matching.

On the other hand, the region growing strategy may stop in areas that contain large height differences (large x-parallaxes) caused by buildings, where the region growing fails to find corresponding points over roofs top of buildings due to sudden change of height (sudden change of the x-parallax). In such a case, new seed points must be provided manually or by a method like based on the SIFT method, published by (Nunes et al, 2008)



Figure4: DSM generated by least squares matching in test area Maras city (Window: 400x600 pixels)

If the distance of neighbored buildings is not large enough, on the other hand this leads to missing ground points, causing a merging of closely neighbored building by interpolation of the height model. This happens often in the case of large sub-matrixes for matching (figure4).

Automatic image matching in build up areas is difficult because of the different view direction to the object. In the case of a smaller angle of convergence, the images of a stereo scene are more similar, improving the matching quality. The stereo images of the first test area have a height to base relation of 7.6 that means only 7.5° convergence angle, improving the matching, as it can be seen in figure 4.

In the next section, we tested an IKONOS stereo pair in the city area of Istanbul having a height to base relation of 1.5 (37° angle of convergence). The larger angle of convergence by simple theory improves the height accuracy, but it enlarges the disparity, causing some matching problems. The quite different view direction leads to different corresponding images, different view to shadows and vertical building parts.

In comparison in the case of analogue aerial cameras, normal angle with a height to base relation of 3.2 is preferred in cities, while in open areas wide angle cameras with a height to base relation of 1.6 are used. The normal angle cameras allow the stereoscopic view also to narrow streets. Since narrow streets between buildings disappear with large nadir angle.



Istanbul left window 460x470 pixels, right window with 450x450 pixels

The dark parts in the DSM shown in figure 5 are caused by the fact, that the different images show different facades of the same building.

As results, the least squares matching with region growing shows following dependency:

- It is strongly dependent upon the number and distribution of selected seed points, whereby the number of
 required seed points is smaller for more similar images and has to be enlarged with increasing stereo
 angles or time interval between imaging.
- Occlusions, causing matching problems, are larger for larger angle of convergence and larger nadir angles.
- The roof shape of buildings cannot be determined clearly with satisfying accuracy and reliability with the size of required sub-matrixes for matching.
- The fixed template size is also not able to track building outlines.
- DSMs generated by this method in city areas cannot be used for all purposes.

4.2 Matching by Dynamic programming

As mentioned before, some of these drawbacks prompted us to think about another method, which could overcome some of these problems like not sharp detection of building outlines. Further on, we liked to replace the manual measurement of seed points by a fully automated operation.

As solution the dynamic programming described in (Birchfield and Tomasi, 1998) has been chosen. The algorithm requires epipolar images, reducing the general 2D matching problem into a faster 1D search. No

windows are required for matching only individual pixels in corresponding lines are used and occlusions are determined and respected. This helps to find building outlines with satisfying accuracy.

In the epipolar image pair, the difference in the position of corresponding points is located only in the x'coordinate direction (direction of imaging base), the differences of the x'-coordinates are named as disparity (x'-x"), identical to the x-parallax. Our goal of using this algorithm is to compute a disparity map with acceptable accuracy from two epipolar lines of IKONOS epipolar pair. For the optimal determination of the corresponding points in epipolar lines, the cost λ (M) has to be minimized:

$$\lambda(\mathbf{M}) = \operatorname{Nocc.} \mathcal{K}\operatorname{occ} - \operatorname{Nm.} \mathcal{K}\mathbf{r} + \sum_{i=1}^{Nm} d(Xi, Yi)$$
(4)

Kocc	Occlusion constant	Nocc	number of occlusions (not the number of occluded pixels)
Kr	matching reward constant	Nm	number of matches
d(Xi,Yi)	dissimilarity between two corresponding epipolar lines		

The dissimilarity between pixels is computed as the minimum of this quantity and its symmetric counterpart:

(5)

$$d(Xi,Yi) = \min\left\{d(Xi,Yi,IL,IR), d(Yi,Xi,IR,IL)\right\}$$

Where: IL, IR: is the intensity in left and right corresponding epipolar lines

The computation of dissimilarity has been done as follows:

First we compute the average of the grey values of pixels in the right epipolar line and its neighboring pixel to

the left:
$$I_R = \frac{1}{2} \{ I_R(Y_i) + I_R(Y_i - 1) \}$$
 and $I_R = \frac{1}{2} \{ I_R(Y_i) + I_R(Y_i + 1) \}$ (6)

Then we let Imin = min {
$$\overline{I_R}$$
, $\overline{I_R}$, $I_R(Yi)$ } and Imax = max { $\overline{I_R}$, $\overline{I_R}$, $I_R(Yi)$ } (7)

With these quantities, we compute $\overline{d}(Xi, Yi, IL, IR), \overline{d}(Yi, Xi, IR, IL)$ as follows:

$$d(Xi, Yi, IL, IR) = \max\{0, IL(Xi) - I\max, I\min - IL(Xi)\}$$
$$\overline{d}(Yi, Xi, IR, IL) = \max\{0, IR(Yi) - I\max, I\min - IR(Yi)\}$$
(8)

Thanks to the structure of the matching cost, the optimal matching relation for each epipolar line pair can be determined by using dynamic programming technique. The optimal match sequence, leads to the disparity at each point of the epipolar line pairs.

In contrast to the least squares matching, the dynamic programming has been shown efficient in providing a dense coverage of corresponding points, particularly at building outlines. Moreover, the computation time was less. Unfortunately, due to the process of matching limited to just the corresponding epipolar lines, a streaking appears in epipolar direction, causing distortion of building borders as can be seen in figure 6,

The results have been improved by median filter limited to pixels in the y-direction. With 7x1 neighbored disparities the noise reduction has been found as optimal. Based on this vertical median filter, the shape of buildings became clearer. Figure 6 shows the DSM generated by applying the dynamic programming algorithm to both test areas. The optimization of this method is still in progress, so only preliminary results are presented.



Before y-direction median filter

Figure 6

After y-direction median filter

5. Discussion and conclusion

Optimal digital surface models are required for a variety of applications such as:
Automated updating of building data base from DSM.
Environmental change analysis of urban areas.
Automatic change detection based on DSM change.
Urban development and visualization.

The automatic generation of digital surface models from IKONOS stereo imagery with high precision is still a difficult task. Especially in densely build up urban areas most methods fail to provide high quality DSM due to occlusion and surface discontinuity. Here we have seen the results of two important factors in generating digital surface models. The first factor is the optimal extraction of building or building block outlines with satisfied accuracy and reliability for practical application. The DSM generated least squares matching is not satisfying for the detection of small buildings and has problems with the precise determination of building outlines, while the DSM calculated with dynamic programming shows sharp edges of buildings but it suffers from streaking in epipolar direction.

An important factor for image matching in build up areas is the height to base relation; the matching of images with a small angle of convergence like in the first test area (h/b=7.6) shows clearer building outlines than the matching with a large base as is the case in the Istanbul data set (h/b=1.6),.

The larger occlusions in the case of small height to base relation caused a loss of some of matched points on facades of buildings, which are not visible in both images, so both methods failed to find corresponding points on building facades.

In spite of promising results, the DSMs generated by both methods, does not satisfy our ambitions so far, there is still scope for an improvement in these algorithms for gaining precise DSMs.

Further research will focus on DSM generation by semi global matching algorithm, which is an extension of dynamic programming with some variations that not only two epipolar line pairs can be used for matching, but the energy function can be used. Moreover mutual information can be used instead of the intensity.

6. Reference:

Birchfield, S. and Tomasi, C. 1996. Depth discontinuities by pixel-to pixel stereo. Technical Report STAN-CS-TR-96-1573, Stanford University.

Birchfield, S. and Tomasi, C. 1998a. Depth discontinuities by pixel to-pixel stereo. In Proceedings of the 6th International Conferenceon Computer Vision, pp. 1073–1080

Birchfield, S. and Tomasi, C. 1998b. A pixel dissimilarity measure that is insensitive to image sampling. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Büyüksalih, G.; Jacobsen, K., 2007: Digital Surface Models in Build up Areas based on very high resolution Space Images: ASPRS annual conference. Tampa, on CD, 8 pages also http://www.ipi.uni-hannover.de (last access September 2008)

Grodecki, J. (2001): IKONOS Stereo Feature Extraction- RPC Approach, ASPRS annual conference St. Louis 2001.

Heipke, C., Kornus, W. and Pfannenstein, A., 1996. The evaluation of MEOSS airborne 3line scanner imagery processing chain and results. Photogrammetric Engineering and Remote Sensing 62 (3), pp. 293–299.

Jacobsen, K., 2006: Digital surface models of city areas by very high resolution space imagery, EARSeL Workshop on Urban Remote Sensing. Berlin, also http://www.ipi.uni-hannover.de (last access September 2008)

Jacobsen, K., 2007: Orientation of high resolution optical space images: ASPRS annual conference. Tampa 2007, on CD, 8 pages also http://www.ipi.uni-hannover.de (last access September 2008)

Krauß, T., Lehner, M., Reinartz, P., 2008. Generation of coarse 3D models of urban areas from high resolution stereo satellite images, ISPRS2008 Beijing, IntArchPhRS. part B2, Beijing, 2008, pp. 1091-1098

Krauß, T., Reinartz, P., Lehner, M., Schroeder, M. and Stilla, U., 2005: DEM generation from very high resolution stereo satellite data in urban areas using dynamic programming. In: IntArchPhRS., Vol. 36 (1/W3). ISPRS Workshop, Hannover. http://www.ipi.uni-hannover.de (last access September 2008)

Nunes, J.L., Jacobsen, K., Feitosa, R., Mota, G., Heipke, C., Silveira, M., Amorim, S., 2008: Estimation of Height Change in Poor Urban Settlements of Rio de Janeiro from Multi-time, Multiscale, and Multi-platform Remotely Sensed Data IntArchPhRS., Vol. XXXVII, Part B1 (WG I/5) pp 835-840

Otto, G.P; Chau, T.K.W. 1989, Region-growing algorithm for matching of terrain images. Image and Vision, V. 7, N.2, pp. 83-94.

Poon, J., Fraser, C., Zhang, C., 2007. Digital surface models from high resolution satellite imagery. Photogrammetric Engineering and Remote Sensing, volume73, number11, November

Zhang, L., Gruen, A., 2004. Automatic DSM generation from linear array imagery data. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 35(B3): 128-133.

Zhang, L., Gruen, A., 2006, Multi-Image Matching for DSM Generation from IKONOS Imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 60(3), 195-211.

Zhang, L., Pateraki, M., Baltsavias, E., 2002: Matching of IKONOS Stereo and Multitemporal GEO Images for DSM Generation, Proc. Map Asia 2002, Asian Conference on GIS, GPS, Aerial Photography and Remote Sensing, 7 - 9 August, Bangkok, Thailand.