COMPARISON OF DSM GENERATION METHODS OF URBAN AREAS FROM IKONOS IMAGES

Thomas Krauß, Manfred Lehner, Peter Reinartz,
German Aerospace Center (DLR), Remote Sensing Technology Institute
PO Box 1116, 82230 Wessling, Germany
thomas.krauss@dlr.de

KEYWORDS: VHR data, Stereo Data, DSM generation, dynamic programming

ABSTRACT:

In this paper two methods, an area based matching and a so called dynamic line warping, for generating digital surface models (DSMs) from IKONOS image pairs are compared and their limitations and usability for the rapid generation of good DSMs are discussed. The new dynamic line warping method is based on techniques used in speech recognition for warping corresponding epipolar lines of the stereo images onto each other in order to extract so parallaxes and object heights. One critical parameter for the derivation of DSMs is the convergence angle of the two stereo images. A small convergence angle gives similar images with smaller parallaxes and is better for area based matching techniques but is said to result in larger height errors and lacks of views from vertical features like walls. A large convergence angle – as used in the standard stereo products of most satellite image providers – on the other hand leads in urban areas to many features hidden in one or the other image and very different viewing conditions leading to more problems for intensity based matching. Therefore three IKONOS scenes with different convergence angles are used and the results are compared qualitatively.

1 INTRODUCTION

Three dimensional models of cities are important for many applications. Besides generating maps and deriving data for urban planning such data become essential in cases of catastrophes like earth quakes, simulation of tsunami events for urban areas at the coast or the simulation of floodings in cities near rivers. Since these events often take place in developing countries, the availability of such models in this area is very rare and the low update rates are a worldwide problem.

Existing methods for the generation of three dimensional models are based on high resolution aerial images or laser scanner measurements and additional external data like cadastral information. Since such data are not available in most of the described scenarios other data sources and methods have to be used. One possible approach for obtaining the required accuracy at a short term all over the world is the usage of stereo image pairs from very high resolution (VHR) satellites like IKONOS or QuickBird.

Based on such image pairs the generation of a high resolution digital surface model (DSM) and subsequently the three dimensional city model is possible for nearly every position of the world.

Subject of this paper is the first step of the generation of three dimensional city models: the derivation of a DSM from high resolution satellite stereo image pairs. Satellite companies offer stereo pairs with relative high convergence angles of about 30° to 60°. In urban areas such high convergence angles result in large occlusions and image parts looking very different in the two images.

Therefore two methods – a classical area based matching and a dynamic programming approach based on epipolar geometry – are applied to three different stereo pairs acquired with three different convergence angles.

The first method starts with a classical hierarchical intensity based matching not assuming epipolar geometry (Lehner and Gill, 1992). Such area based matching depends on comparing small windows in the images. Areas with steep slopes and large convergence angles are problematic since the two corresponding parts of the stereo images are very different e.g. in the neighbourhood of occlusions. Therefore such area based matching methods often fail in such scenes.

For this reason a good handling of geometric changes and occlusions is more important for urban scenes than for a softly undulating landscape.

The second method used in this investigation is based on epipolar images from which parallaxes and therefore the height information is extracted via a dynamic programming approach. The method is based on the observation how human recognition tries to fit stereo images onto each other. This process feels like distorting the two images first in a coarse level and then step by step in finer levels so that they fit piecewise together. In this process the brain automatically masks out inconsitent areas (occlusions) like walls seen only in one of the images.

An algorithm which maps corresponding lines of the stereo image one to the other was found in an approach called "dynamic time warping" based on dynamic programming which is well known in speech recognition (Ney, 1982; Sakoe and Chiba, 1978; Itakura, 1975). In speech recognition recorded samples of words rarely fit to actual spoken words. The algorithm matches an actual sound to learned words. Therefore it warps a time-sequence of characteristics onto an other and a measure for the needed transformations is returned as a sort of distance between the words.

Due to the fact that the requirements are very comparable – warp two sequences of values to each other – this method is also a good approach for warping epipolar image lines of one image to the corresponding lines in the stereo partner. Similar algorithms are already used in the meantime in computer vision (van Meerbergen et al., 2002; Hirschmüller, 2005).

These two methods are applied to all of the three test image pairs representing different exposure situations and the results are compared and discussed qualitively. For a quantitative investigation of the Munich test case see (Krauß et al., 2006). The feasibility of the methods for generating DSMs under certain circumstances and requirements for the stereo image pairs for a highly automatic three dimensional representation of urban areas are discussed.

2 IMAGE AND REFERENCE DATA

For experiments the following Ikonos data sets provided by Space Imaging were used:

- Southern France acquired 2004-01-30, 10:51 GMT with a ground resolution of 83 cm, with declination angles of

$+6.8°$ and $-1.1°$ (fig. 1) and a convergence angle of $6.29°$

- Munich acquired 2005-07-15, 10:28 GMT, level 1A images (only corrected for radiometry and inner orientation) with a ground resolution of 83 cm and declination angles of $+9.25°$ and $-4.45°$ (fig. 2) giving a convergence angle of $9.95°$
- Athens acquired 2004-07-24, 9:24 GMT as a standard stereo image pair (level 1B, epipolar) with a ground resolution of 88 cm and declination angles of $-19.99°$ and $+13.17°$ (fig. 3) and a convergence angle of $30.28°$

The southern France scene is a special non-standard Ikonos stereo acquisition with a small stereo angle of about only 6°. A part of the scene and a section from this part are shown in fig. 1.

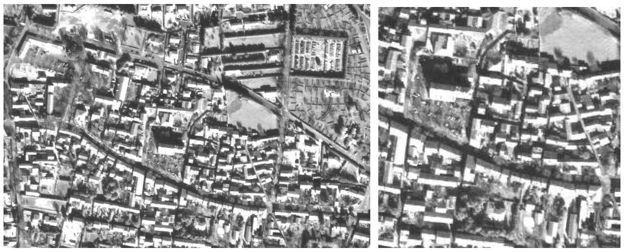

Figure 1. Ikonos image of test area “Southern France” (600 m×400 m) and section (250 m×250 m)

The Munich scene is also a non-standard Ikonos stereo acquisition with a small stereo angle of about only 10° as shown in fig. 2.

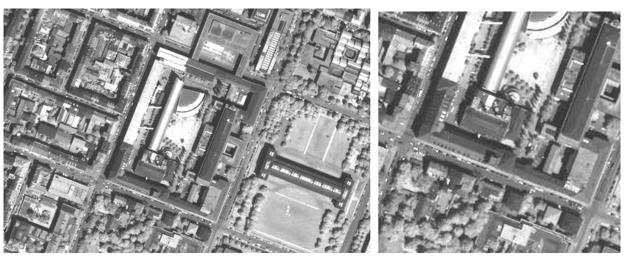

Figure 2. Ikonos image of test area “Munich TUM” (600 m×400 m) and section (250 m×250 m)

The Athens scene was acquired as a so called “Ikonos standard stereo pair” which means declination angles of about 10° to 20° each in one orbit and delivery as epipolar images (actual stereo angle was about 30°). A part of the scene and a section from this part are shown in fig. 3.

Figure 3. Ikonos image of test area “Athens” (600 m×400 m) and section (250 m×250 m)

2.1 RPC and Epipolar Geometry

Each high resolution Ikonos image is accompanied by two rational polynomial functions (RPC – rational polynomial camera model) which map object space coordinates longitude X, latitude Y and ellipsoid height Z into image coordinates $(x, y)^T$. One of these rational polynomial functions is given as a quotient of two third order polynomials with 20 coefficients each (Jacobsen et al., 2005; Grodecki et al., 2004).

Based on these polynomials an image location can be calculated for every object point and vice versa longitude and latitude for a given height by recursive fitting from the image coordinates. In fig. 4 the satellite positions for the two images of the Munich-scene are calculated in this manner for three selected image scanlines (orbit height reduced by a factor 20, longitude and latitude of satellite by a factor 3). The figure shows besides the orbit clearly the acquisition method of the second image in reverse mode. This means that the satellite first acquires the southernmost scanline traversing over the scene from north to south but scanning from south to north.

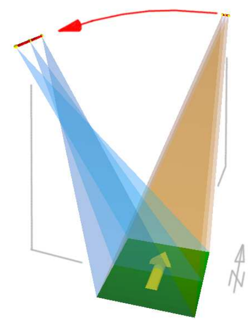

Figure 4. As an example the Ikonos acquisition geometry for the Munich scene was calculated from the provided RPCs (first image taken from north (brown), second in reverse mode from west (blue))

The tie point determination used in the first method does not require epipolar geometry. The second method in contrast is depending on epipolar geometry. For VHR satellite imagery quasi epipolar images can be derived using algorithms based on the provided RPCs for each of the images given in (Morgan, 2004). For standard stereo configurations Ikonos images can be ordered already in epipolar geometry from SpaceImaging (test case Athens).

3 DESCRIPTION OF THE DSM GENERATION METHODS

3.1 Hierarchical intensitiy based matching

Hierarchical intensity based matching consists of two major steps: First the matching process uses a resolution pyramid (Lehner and Gill, 1992; Kornus et al., 2000) to cope even with large stereo image distortions stemming from carrier movement and terrain. Large local parallaxes can be handled without knowledge of exterior orientation which is often not available with sufficient accuracy. The selection of pattern windows is based on the Foerstner interest operator (Foerstner and Guelch, 1987) which is applied to one of the stereo partners. For selection of search areas in the other stereo partner(s) local affine transformations are estimated based on already available tie points in the neighborhood from a coarser level of the image pyramid.

Tie points with an accuracy of one pixel are located via the maximum of the normalized correlation coefficients computed by sliding the pattern area all over the search area. These approximate coordinates of tie points are refined to sub-pixel accuracy by local least squares matching (LSM). The number of tie points found and their final sub-pixel accuracy achieved depend mainly on image similarity and decrease with increasing stereo angles or time gaps between imaging.

The second step uses a region growing first published by (Otto and Chau, 1989) based on the implementation of the Technical University of Munich (TUM) (Heipke et al., 1996). It combines LSM with a strategy for local propagation of initial conditions of LSM.

Various methods for blunder reduction are used for both steps of the matching:

- Given threshold for correlation coefficient
- 2-directional matching and threshold on resulting shifts of the coordinates
- Threshold on residuals in image space from forward intersection based on the rigorous modeling of the imaging process or on rational polynomial functions (RPC).

In areas of low contrast the propagation of affine transformation parameters for LSM in region growing leads to high rates of blunders. In order to avoid intrusion into homogeneous image areas (e.g. roof planes without structure) the extracted image chips are subject to low thresholds on variance and roundness of the Foerstner interest operator. This and the many occlusions found in densely built-up areas imaged with a large stereo angle create lots of insurmountable barriers for region growing. Thus for this type of stereo imagery the massive number of seed points provided by the matching in step one turns out to be essential for the success of the region growing.

Test area	**Stereo angle (degrees)**	**b/h ratio**	**Percentage of tie points**	
			Base: interest op. points	**Base: total pixels**
France	6.29°	0.11	98 %	89 %
Munich	9.95°	0.17	86 %	84 %
Athens	30.28°	0.54	46 %	41 %

Table 1. Percentages of tie points in relation to interest operator points (original image resolution of image pyramid) and total number of pixels in image (region growing)

The numbers of tie points found and their sub-pixel accuracy is highly dependent on the stereo angle. A large stereo angle and thus a large base to height ratio b/h leads to poorer numbers of tie points and to lower accuracy in LSM via increasing dissimilarity of correctly extracted image chips. Thus, a large b/h cannot be recommended for stereo imaging of city areas. This contradicts the normal imaging practice for Ikonos and QuickBird stereo acquisitions. Table 1 gives the percentages of tie points from the two matching steps for the three test areas.

In a first approximation the standard deviation of stereo height measurement is proportional to the quotient of the standard deviation of the image coordinates from matching and the base to height ratio. Thus, a smaller b/h can be compensated by improved matching accuracy. This effect can be seen in table 2 for our test areas where the smaller b/h seems to be even more than compensated by better matching conditions. Thus, in case of small stereo angles more tie points with overall better accuracy can be determined in densely built-up areas resulting in a better DSM.

Area	c	σ **[pixel]**
France	0.95	0.094
Munich	0.90	0.137
Athens	0.78	0.175

Table 2. Mean correlation coefficient c from LSM and standard deviation σ of residuals in RPC forward intersection (after blunder reduction)

In order to compute a DSM from the tie point cloud from matching the object space coordinates of the tie points are computed by forward intersection based on the rigorous modeling of the imaging process or on RPC. If ground control points are available they are used in bundle adjustment or correction of RPC. The irregular 3D point cloud is transformed into a regular DSM grid by using the triangulation and interpolation as described in (Hoja et al., 2005).

3.2 Dynamic line warping

For the second DSM generation method the images are assumed to be available as a stereo pair with parallaxes in image line direction (epipolar images, horizontal epipolar lines). Creating a dense digital surface model (DSM) from such image pairs is based on warping corresponding image lines of the two images onto each other and calculating local heights from the relative displacements of this warping (parallaxes).

Because of the assumption of epipolar images there exist only horizontal parallaxes and no vertical shifts between the two images. Each line of an image is represented as a profile of gray values as shown in figure 5 top and bottom respectively.

Such gray value sequences have to be mapped onto each other by stretching and compressing parts to achieve an optimal local fit. For this the “dynamic time warp” algorithm used in speech recognition (Ney, 1982) was implemented to warp image lines one to each other (Krauß et al., 2005). This method calculates spectral characteristics for short overlapping parts of the audio samples, calculates distances between each of these parts and uses dynamic programming to receive a so called “minimal total distance” for the given sample with the compared sample of the dictionary. Only this minimum total distance is used further for determining the most probable sequence of words.

For explanation of the method let’s take as an example two arbitrary sequences of values I and I' as: $I^T = (1 \quad 0 \quad 2 \quad 1 \quad 0)$ and $I'^T = (0 \quad 1 \quad 0 \quad 2 \quad 1)$. For the implementation of the algorithm first a “distance” has to be defined. In the case of a direct comparison of the sequences of gray values of the two epipolar images this can be in a first approach the absolute value of the gray value distance $M_{i,j} = |I_i - I'_j|$.

In the next step – the dynamic programming approach for simplifying an exhaustive recursion – the rows and columns are cumulated to a matrix D filling the first line and column according to

$$D_{1,j} = \sum_{k=1}^{j} M_{1,k}, \qquad D_{i,1} = \sum_{k=1}^{i} M_{k,1}$$

and the rest of the matrix $D_{i,j}$ $(i, j > 1)$ according to

$$D_{i,j} = M_{i,j} + \min\left\{D_{i-1,j}, D_{i,j-1}, D_{i-1,j-1}\right\}$$

yielding D in the example from above:

$$D = \begin{pmatrix} \mathbf{1} & \mathbf{1} & 2 & 3 & 3 \\ 1 & 2 & \mathbf{1} & 3 & 4 \\ 3 & 2 & 3 & \mathbf{1} & 2 \\ 4 & 2 & 3 & 2 & \mathbf{1} \\ 4 & 3 & 2 & 4 & \mathbf{2} \end{pmatrix}$$

In this Matrix D the overall distance is defined as the rightmost bottom element – in the example 2. This value is a measure for all needed shifting, stretching and squeezing operations for one sequence to fit onto the other. In speech recognition it's sufficient finding the dictionary sample with the smallest minimal total distance to the given voice input.

But besides this minimal total distance a so called "minimal path" can be defined. This path connects the endpoints of the compared sequences in a manner describing what parts of one sequence have to be shifted, stretched or squeezed to fit onto the other.

This minimal path is required by our approach to extract the wanted parallaxes and thus the heights. Starting from the rightmost bottom element going back always using the smallest possible next neighbour to the top, left or top-left gives the minimal path (marked above in bold).

Calculating the matrix, picking a small area of the result and showing the extracted correlations between the input gray value profiles yields fig. 5. As can be seen well areas with different widths in the profiles are correctly mapped to each other.

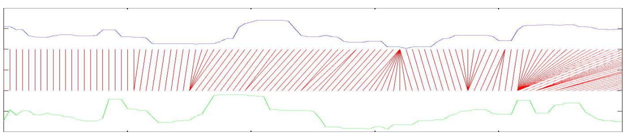

Figure 5. Found correlations between the two input gray value profiles (top and bottom of graphic)

A perfectly diagonal line as minimal path represents no relative displacements between the two profiles. All deviations from this line indicate the searched parallaxes. If this minimal path is represented by pairs of coordinates $\mu_k = (i,j)_k$ it's possible to fit a parabola through $D_{i-1,j+1}$, $D_{i,j}$ and $D_{i+1,j-1}$ to calculate the parallaxes as the minimum of this parabola in subpixel accuracy.

Using small window areas of diameter w around every line position for calculating the distance $M_{i,j}$ gives better results but introduces on the other hand smoothing effects as resulting from area based methods. The distance $M_{i,j}$ of line y with images $I_{x,y}$ and $I'_{x,y}$ and window size w is then calculated as

$$M_{i,j} = \sum_{\lambda=-\lfloor w/2 \rfloor}^{w-\lfloor w/2 \rfloor -1} \sum_{\mu=-\lfloor w/2 \rfloor}^{w-\lfloor w/2 \rfloor -1} \left| I_{i+\lambda,y+\mu} - I'_{j+\lambda,y+\mu} \right|$$

This means that over a window of size $w \times w$ all absolute values of gray value distances between the corresponding pixels in the two images are summed up ($\lfloor x \rfloor$ is the largest integer smaller than x).

Afterwards a vertical median filter is applied to reduce line streaking and blunders resulting from the line by line correlation. For acceleration of processing time a maximum correlation window size can be chosen. Outside of this distance from the main diagonal the matrix M simply gets filled with a maximum distance value instead of the calculated distance.

For the final calculation of the DSM the above resulting disparity image is back-transformed from epipolar image space to original image space. Forward intersection using RPCs and subsequent interpolation produce a regular DSM.

4 APPLICATION OF THE METHODS TO THE TEST DATA

4.1 Test data set Athens

The first example is the test data set "Athens" as shown in fig. 3. Applying first the classical hierarchical algorithm gives as a result the DSM shown in fig. 6.

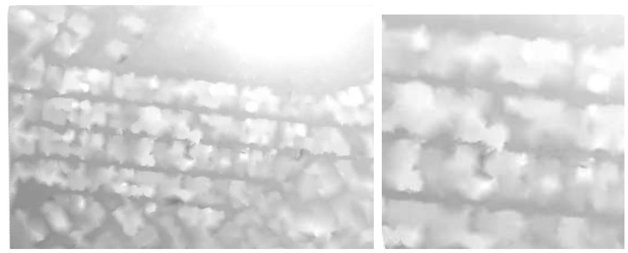

Figure 6. DSM of test area "Athens" (Fig. 3) generated by hierarchical intensity based matching (600 m×400 m) and section (250 m×250 m)

The DSMs shown here and in the following images contain always the same clippings from the Ikonos stereo pairs as shown in section 2. As can be seen in the original Ikonos images (fig. 3) the blocks of houses mostly contain narrow courtyards in the center. These courtyards are about 8 to 10 m wide whereas the houses are about 15 m and the streets are about 12 m wide.

Fig. 7 shows the DSM of the same section calculated with the line warping method using a window size $w = 3$. For the selection of the window size see (Krauß et al., 2005).

Figure 7. DSM of test area "Athens" (Fig. 3) generated by dynamic line warping

The DSM calculated with the classical hierarchical algorithm shows no large blunders but a much smoother DSM whereas the DSM calculated with the dynamic programming approach shows sharp edges but also blunders and streaking in epipolar direction because of the missing linkage between the calculations for each epipolar line. However the line warping method shows much more details especially at sharp ridges than the classical area based method but generates also more blunders.

As can be seen in the small sections the area based method smoothes out many yards which are much better preserved by the line warping method.

Cutting profiles through these DSMs as shown in fig. 8 illustrates the differences between the two methods (all units in the profiles are given in meters; heights from sealevel and horizontal distance along profile).

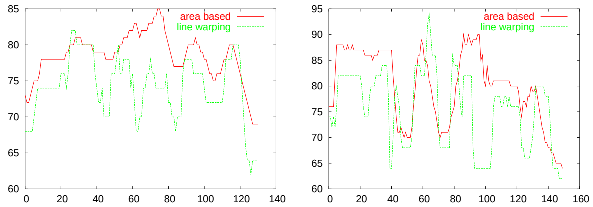

Figure 8. Profile in epipolar direction (left) and across (right) from the calculated DSMs (red: area based matching, dotted green: line warping)

As can be seen the area based method smoothes out most of the smaller courtyards because of the window size of 13 × 13 pixels used in region growing. The streets in this area are, however, wide enough to be found by the area based matching. The line warping method on the other hand reveals sharp edges reproducing buildings, courtyards and streets much more clearly whereas the line streaking effects leading to blurry edges can be seen in epipolar direction (left-right).

4.2 Test data set Munich

Applying the area based matching to the Munich scene shown in fig. 2 yields the DSM shown in fig. 9.

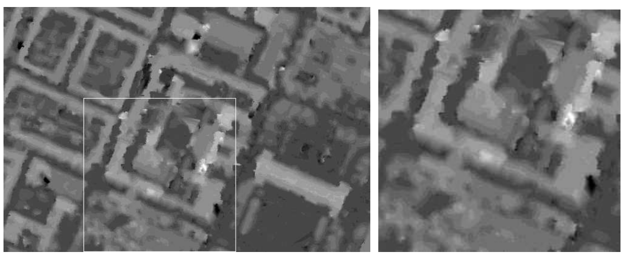

Figure 9. DSM generated by applying the hierarchical intensity based matching to the Munich scene (600 m×400 m) and section (250 m×250 m)

As can be seen the results are obviously better than in the Athens scene. Small courtyards are better resolved than in the previous test case. Fig. 9 shows some blurring at the edges of the buildings and smoothed out regions on the main building of the university.

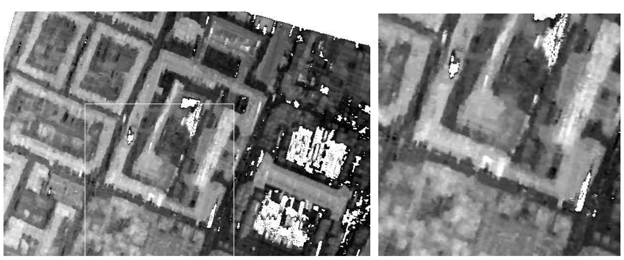

Figure 10. DSM calculated with line warping (600 m×400 m) and section (250 m×250 m)

After transforming the Munich scenes locally to epipolar geometry using some tie points the dynamic programming algorithm is applied as shown in fig. 10 (window size 3).

The DSM generated by line warping shows much less smoothing than the one generated with the area based matching (fig. 9) but on the other hand more blunders appear as can be seen for example in the grass area near the Old Pinakothek and streaking effects in epipolar direction (top to bottom).

The edges of the buildings are more straight in epipolar direction using line warping than with the area based method but across epipolar direction more jagged due to the line streaking effects. So the tower of the university in the middle of the southern buildings gets nearly wiped out due to the line streaking whereas on the other hand the rim of the western building and the ridge of the main building is badly reproduced by the area based method due to missing tie points on the street.

4.3 Test data set southern France

Fig. 11 shows the DSM of the test case “Southern France” calculated with the classical hierarchical area based method and fig. 12 the one calculated with dynamic line warping. This test area represents a case with an extreme narrow convergence angle of only $6.29°$ and only areas with few and low buildings.

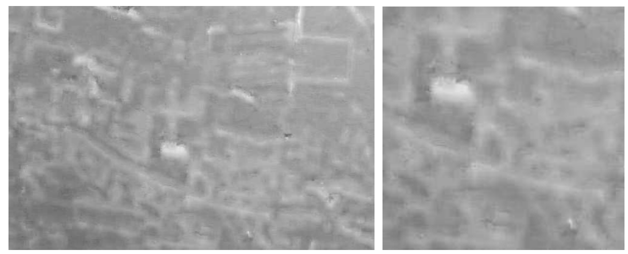

Figure 11. DSM of test area southern France generated by the hierarchical area based method (600 m×400 m) and section (250 m×250 m)

In this case the area based matching method shows all of its advantages in comparison to the line warping algorithm. The results of the latter algorithm are significantly poorer due to many grassy areas which are up to now not handled correctly by this algorithm.

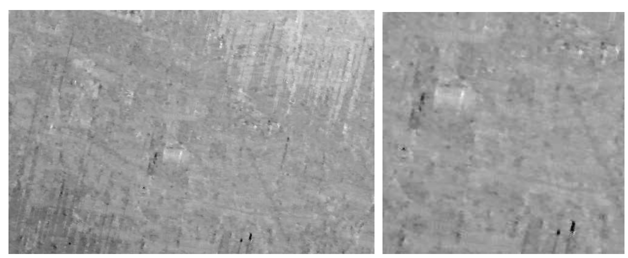

Figure 12. DSMs of test area southern France generated by line warping algorithm (600 m×400 m) and section (250 m×250 m)

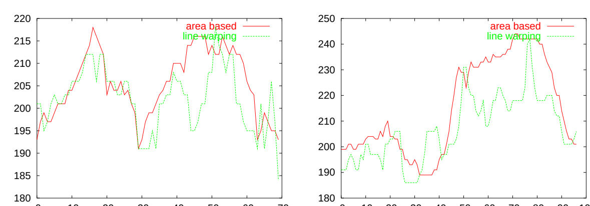

Figure 13. Profile in epipolar direction (left) and across (right) from the calculated DSMs (red: area based matching, dotted green: line warping)

Cross sections in epipolar direction (top-bottom) and across are shown in fig. 13. Investigation of these shows analogous results as in the previous test cases. The area based matching gives here a significantly better DSM due to the small convergence angle but narrrow streets and yards are also wiped out.

5 DISCUSSION

Comparing the DSMs generated by the two methods for the three test scenes shows pros and cons for both methods. So the classical area based matching relies heavily on finding many good tie points in the first step since the region growing can't proceed to large areas in urban scenes due to occlusion and other effects. Missing tie points result in a smoothing out and no straight and sharp edges.

Comparing the DSMs of the scenes shows clearly an effect of the different convergence angles. With narrow convergence angles like in the France test case the area based method shows much better results than the line warping method. Under wide convergence angles as used in the Athens case the results of the line warping method are significantly better than those of the area based method. In this case the DSM quality of the area based method drops significantly due to more missing important tie points with increasing convergence angle and therefore a smoothing out of much more small details like streets or courtyards.

The area based matching method is already in use since many years and gives very good results in smoother terrain like the southern France test case or with stereo data of coarser resolution. But in densely built up urban areas with steep walls and stereo pairs with larger convergence angles, missing tie points lead to more and more significant errors in the DSM as can be seen in the Athens test case.

The line warping method suffers in the actual algorithm still from streaking effects and large blunders in grass areas. On the other hand no such significant degeneration of the produced DSMs with increasing convergence angles can be found.

Further work will be needed to reduce the streaking between lines through interconnection of the epipolar lines in the DSM generation. Those interconnection is also justified by the fact that epipolar projection is to a certain extent an approximation. For example in the full Munich scene (13000×15000 pixels) the epipolar projection reports for the 813 input tie points a zero mean of line parallax with a standard deviation of 0.44 pixel. Also analysis and elimination of large blunders in grass areas possibly due to BRDF effects have to be implemented.

REFERENCES

Foerstner, W. and Guelch, E., 1987. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In: ISPRS Intercommission Workshop, Interlaken.

Grodecki, J., Dial, G. and Lutes, J., 2004. Mathematical model for 3d feature extraction from multiple satellite images described by rpcs. In: ASPRS Annual Conference Proceedings, Denver, Colorado.

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

Hirschmüller, H., 2005. Accurate and efficient stereo processing by semi-global matching and mutual information. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Hoja, D., Reinartz, P. and Lehner, M., 2005. DSM generation from high resolution satellite imagery using additional information contained in existing DSM. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36 (1/W3). ISPRS Workshop, Hannover.

Itakura, F., 1975. Minimum prediction residual principle applied to speech recognition. IEEE Trans. Acoust. Speech signal process. ASSP 23, pp. 67–72.

Jacobsen, K., Büyüksalih, G. and Topan, H., 2005. Geometric models for the orientation of high resolution optical satellite sensors. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36 (1/W3). ISPRS Workshop, Hannover.

Kornus, W., Lehner, M. and Schroeder, M., 2000. Geometric inflight calibration by block adjustment using MOMS-2P 3-line-imagery of three intersecting stereo-strips. SFPT (Société Française de Photogrammétrie et Télédétection) 159, pp. 42–54.

Krauß, T., Reinartz, P., Lehner, M. and Stilla, U., 2006. Generation and evaluation of DSM from ikonos stereo pairs. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. PCV06, Sept. 2006, Bonn.

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: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36 (1/W3).

Lehner, M. and Gill, R., 1992. Semi-automatic derivation of digital elevation models from stereoscopic 3-line scanner data. ISPRS 29 (B4), pp. 68–75.

Morgan, M., 2004. Epipolar Resampling of Linear Array Scanner Scenes. PhD thesis, University of Calgary, Canada. UCGE Reports Nr. 20193.

Ney, H., 1982. Dynamic programming as a technique for pattern recognition. In: Proc. 6th Int. Conf. Pattern Recognition, pp. 1119–1125.

Otto, G. and Chau, T., 1989. Region growing algorithm for matching of terrain images. Image and vision computing (7) 2, pp. 83–94.

Sakoe, H. and Chiba, S., 1978. Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. Acoust. Speech signal process. ASSP 26, pp. 43–49.

van Meerbergen, G., Vergauwen, M., Pollefeys, M. and van Gool, L., 2002. A hierarchical symmetric stereo algorithm using dynamic programming. International Journal of Computer Vision 47 (1/2/3), pp. 275–285.