

ADVANCEMENT IN MATCHING OF SPOT IMAGES BY INTEGRATION OF SENSOR GEOMETRY AND TREATMENT OF RADIOMETRIC DIFFERENCES

Emmanuel P. Baltsavias, Dirk Stallmann

Institute of Geodesy and Photogrammetry
Swiss Federal Institute of Technology
ETH-Hoenggerberg, CH-8093 Zurich, Switzerland
Tel.: +41-1-3773042, Fax: +41-1-3720438, e-mail: manos@p.igp.ethz.ch

Commission IV

ABSTRACT

This paper presents a matching algorithm for automatic DTM generation from SPOT images that provides dense, accurate and reliable results and attacks the problem of radiometric differences between the images. The proposed algorithm is based on a modified version of the Multiphoto Geometrically Constrained Matching (MPGC). It is the first algorithm that explicitly uses the SPOT geometry in matching, restricting thus the search space in one dimension, and simultaneously providing pixel and object coordinates. This leads to an increase in reliability, and to reduction and easier detection of blunders. The sensor modelling is based on Kratky's polynomial mapping functions to transform between the image spaces of stereopairs. With their help epipolar lines that are practically straight can be determined and the search is constrained along these lines. The polynomial functions can also provide approximate values, which are further refined by the use of an image pyramid.

Radiometric differences are strongly reduced by performing matching not in the grey level but in gradient magnitude images. Thus, practically only the information in stripes along the edges is used for matching. Edges that exist in only one image can be detected by subtracting quasi registered images in the upper levels of an image pyramid and by consistency checks. The points to be matched are selected by an interest operator in preprocessed gradient images. Gross errors can be detected by statistical analysis of criteria that are provided by the algorithm.

The results of an extensive test using a stereo SPOT model over Switzerland will be reported. Different cases of radiometric differences, and matching with different options and the qualitative comparison of the results based on forty thousand check points will be presented.

KEY WORDS: SPOT, automatic DEM generation, image matching, geometric constraints, radiometric differences, accuracy analysis

1. INTRODUCTION

The motivation behind this research was the aim to improve matching of SPOT images by integration of the sensor geometry and the treatment of radiometric differences. As far as the authors know, none of the published matching methods explicitly exploits the SPOT geometry to restrict the search space. Since SPOT does not have a perspective geometry in flight direction, accurate epipolar images can not be generated without the use of a DTM and thus the existing matching algorithms perform a 2-D search. The authors investigated Kratky's polynomial mapping functions (PMFs) (Kratky, 1989a) which transform from image to image, image to object and object to image space. For the computation of the coefficients of the PMFs the results of a rigorous bundle adjustment are used. The PMFs are much faster and almost equally accurate as the rigorous transformations. When the ray of an imaged point is projected by using the PMFs onto the other image of a stereo pair (epipolar line) it is with a very good

approximation a straight line. If a straight line is defined by the projection of a small ray segment which is centred at the correct point position in object space, then the deviation of the epipolar line from the straight line would be 0.25 pixel for a height error of more than 7 km. More details on the characteristics of the PMFs can be found in Baltsavias and Stallmann, 1992.

Attempts to circumvent the problem of radiometric differences have concentrated on using images acquired within a short time interval. However, this is difficult to achieve and does not solve the problem. The along-track stereo will strongly reduce these problems but can not eliminate them due to different perspective views, clouds and occlusions. Fusion and matching of multitemporal and multisensor data, as in change detection applications, will retain their importance even in the era of along-track stereo. Thus, the authors decided to attack this problem, which has been up to now treated only to a limited extent. The idea is to use gradient magnitude images, thus eliminating radiometric differences in areas of low

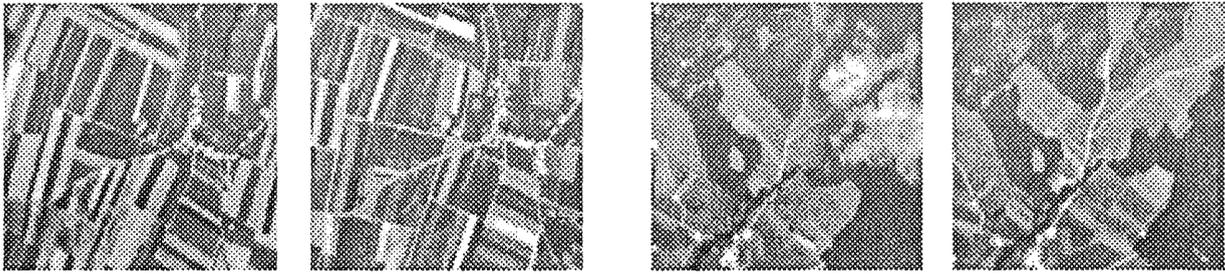


Figure 1 Radiometric differences due to agricultural activities (left pair) and due to clouds and shadows (right pair).

texture. Preliminary investigations have shown that the majority of the edges remain stable. However, different edges exist due to clouds, shadows, different perspective views, new edges within fields due to agricultural activities, human intervention, water level, snow coverage, changes in the tree canopies etc. (Figure 1). A method should be developed to try to detect the different edges through consistency checks.

2. TEST DATA

A stereo SPOT panchromatic level 1A model over W. Switzerland was acquired. The inclination of the sensor's optical axis was 23.4° R and 19.2° L respectively, leading to a B/H ratio of ca. 0.8. The acquisition dates were 20.7.1988 and 27.8.1988 with significant radiometric differences between the two images, particularly in agricultural areas. Figure 1 shows some typical image parts with large radiometric differences. The elevation range was 350 - 3000 m. The following preprocessing was applied to the original digital images:

- reduction of periodic and chess pattern noise
- Wallis filtering for contrast enhancement

136 control and check points were used with Kratky's rigorous SPOT model (Kratky, 1989b). 10 for the points were used as control points with a linear model of the attitude rates of change. The pixel coordinates were measured in one image manually and transferred to the second one by template matching. The RMS of the check points was 9 - 10 m in planimetry and 6 m in height.

3. MODIFIED MPGC

MPGC is described in detail in *Baltsavias, 1991*. It combines least squares matching (involving an affine geometric transformation and two radiometric corrections) and geometric constraints formulated either in image or object space. The constraints lead to a 1-D search space along a line, thus to an increase of success rate, accuracy and reliability, and permit a simultaneous determination of pixel and object coordinates. Any number of images (more than two) can be used simultaneously. The measurement points are selected along edges that are nearly perpendicular to the geometric constraints line. The approximations are derived by means of an image pyramid. The achieved

accuracy is in the subpixel range. The algorithm provides criteria for the detection of observation errors and blunders, and adaptation of the matching parameters to the image and scene content.

In the case of matching of SPOT images the geometric constraints were formulated as follows. First, given a measurement point in one of the images (template image) a height approximation is needed. If the existing approximations refer to the pixel coordinates, then the height is computed by using the pixel coordinates in the reference image, the x pixel coordinate in the second image and the image to image PMFs. This height Z is altered by a height error ΔZ . Using the heights $Z + \Delta Z$, $Z - \Delta Z$, the pixel coordinates in the template image are projected by the image to image PMFs in the second image where they define the geometric constraints (epipolar) line. The centre of the patch of the second image which is used for matching is forced to move along this line by means of a weighted observation equation of the form

$$v_c = (x + \Delta x) \cos\beta + (y + \Delta y) \sin\beta - p \quad (1)$$

where (x, y) the approximate pixel coordinates of the corresponding point in the second image and $(\Delta x, \Delta y)$ the unknown x-shift and y-shift.

Equation (1) is equivalent to the distance of a point $(x + \Delta x, y + \Delta y)$ (the patch centre of the second image) from a straight line. The epipolar line is expressed by the normal equation of a straight line, where p is the distance of the line from the origin and β is the angle between the perpendicular to the line and the x-axis.

If the patch of the second image does not lie on this line, then it jumps onto the line right in the first iteration. With our data, the epipolar lines are approximately horizontal, i.e. any error in the y-direction will be eliminated right in the first iteration. An example is shown in Figure 2. Since the epipolar lines are horizontal, the measurement points must be selected along edges that are nearly vertical in order to ensure determinability and high accuracy. Some advantages of the geometric constraints will now be presented. SPOT images include due their small scale a high degree of texture, i.e. edges. Measurement points lying along nearly straight edges can not be safely determined with other matching techniques, but with our approach they can as they lie at the intersection of two

nearly perpendicular lines. Figure 3 illustrates such an example. Another usual problematic case is that of multiple solutions. With geometric constraints side minima can only result if they fall along the epipolar line. Figure 4 shows an example with and without geometric constraints.

An extension of our approach, which has not been implemented yet, is the application of such constraints

for the 4 corners of the patch of the template image. This patch represents a part of the object surface which can be modelled by different surfaces (e.g. horizontal plane, inclined plane, 2nd degree surface). Depending on the type of the surface, the heights of the 4 corner points can be defined analytically. For example, if an inclined plane is selected, the heights at the 4 corner points are a function of the height at the patch centre and the two

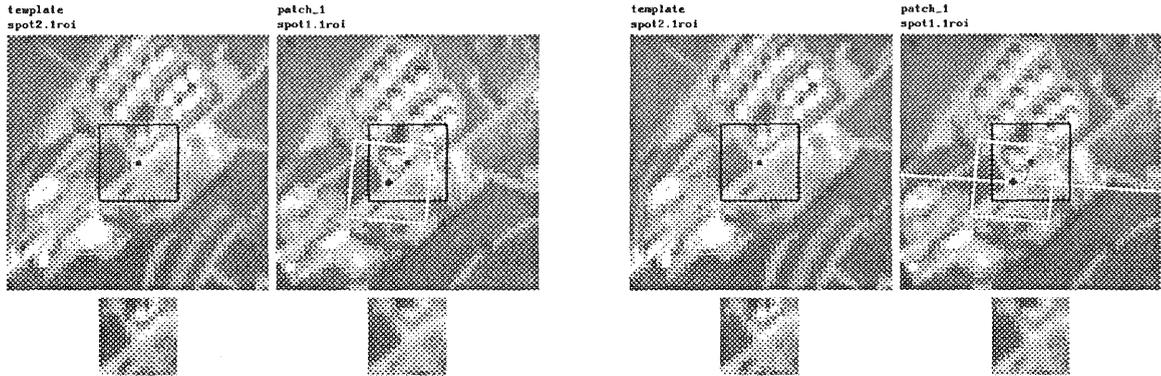


Figure 2 Matching without (left) and with (right) constraints. 10 iterations without and 5 iterations with constraints were needed. The “epipolar line” is the white line in the right image. The black frame is the initial position and the white frame with the black centre cross the final position.

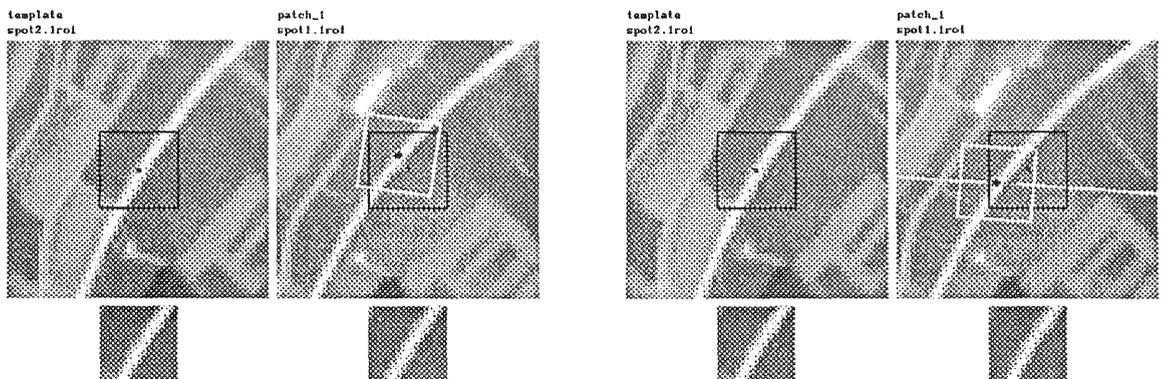


Figure 3 Matching along edges without (left) and with (right) constraints.

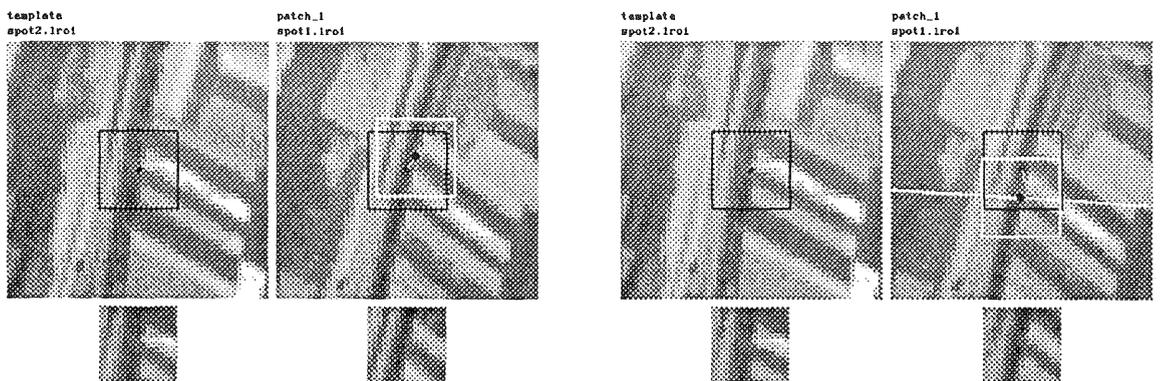


Figure 4 Multiple solution matching without (left) and with (right) constraints.

inclination angles of the plane. Thus, geometric constraints can be formulated by using the known pixel coordinates of all corner points, their heights (which are a function of the height at the patch centre and the plane parameters) and the PMFs (note that their coefficients are the same for all points). The unknowns to be solved for are the plane parameters and the height at the patch centre. This approach is an indirect object-based matching approach. It has the advantage that permits the determination of the local surface parameters (and not just the height at the patch centre), and that it constrains the 4 corners, i.e. restricts the affine geometric transformation. Using an inclined plane model, which is implicitly implied by the affine transformation of the current approach, the new approach also results in computational advantages as the number of unknowns reduces from 6 to 3. However, the problem to be solved is the a priori determination of a suitable surface model.

4. DATA PREPROCESSING AND SELECTION OF MEASUREMENT POINTS

First, the gradient magnitude images are computed. To reduce weak edges due to noise, which is very noticeable in SPOT images, all gradients with a magnitude less than a threshold T can be set equal to 0. The threshold is selected as a function of the mean and the standard deviation of the gradient magnitude image (in this case $T = \text{mean} - \text{standard deviation}$). The same function should be used for both images to ensure equal treatment. The threshold should not be too high otherwise (a) useful texture is deleted, and (b) the edges are broken and significant differences between the two images occur due to different edge strength. This approach eliminates noise but also low texture which is however not very likely to lead to accurate matching results. An example is shown in Figure 5.

As already mentioned, the measurement points are selected along edges nearly perpendicular to the epipolar lines. In order not to reduce the number of the selected points too much (and thus their density, which influences the DTM accuracy), points along edges with an angle of $\pm 45^\circ$ with the perpendicular to the epipolar line should also be selected. To avoid clustering of good points a

thin-out window for non-maxima suppression can be defined. To avoid selecting points lying at small and faint noisy edges the points are selected in the first level of the image pyramid. Our approach is to match the same number of points in all pyramid levels. Thus, a selected point must have the aforementioned properties in all pyramid levels. Generally, the approach to be followed is to detect good points in all levels of the image pyramid of the template image and keep the points that appear in all pyramid levels. However, these SPOT images had a lot of texture and this was expressed in all pyramid levels. By going up in the image pyramid, the relative number of selected points was actually increasing.

To avoid selecting points at regions of radiometric differences, especially the ones with a large area extent (like clouds), the following approach can be used. Using the PMFs and an average height of the scene (derived either from a priori knowledge or from the average height of the control points used in the rigorous SPOT model), or a polynomial transformation derived from the pixel coordinates of the control points, the search image is registered with the template image. If the registration were perfect, a simple subtraction of the two images would give us the different edges. Since the registration is not perfect, an image pyramid is created so that at the highest level the misregistration error is within pixel range. Then through subtraction, the different edges are detected by binarising the difference image with an absolute threshold. This binary image can eventually be dilated in order to avoid selecting points whose patch would partially fall inside areas with radiometric differences. These disturbance areas are projected in all pyramid levels and convolved with the selected points in order to clean the selected points. An example is shown in Figure 6.

5. DERIVATION OF APPROXIMATIONS

In this test the approximations were either given manually or derived from a given DTM. The proposed general approach is the following. After the PMFs are computed an average height is used in order to determine the position of the selected points in the search image. To check the quality of these approximations the 136 points

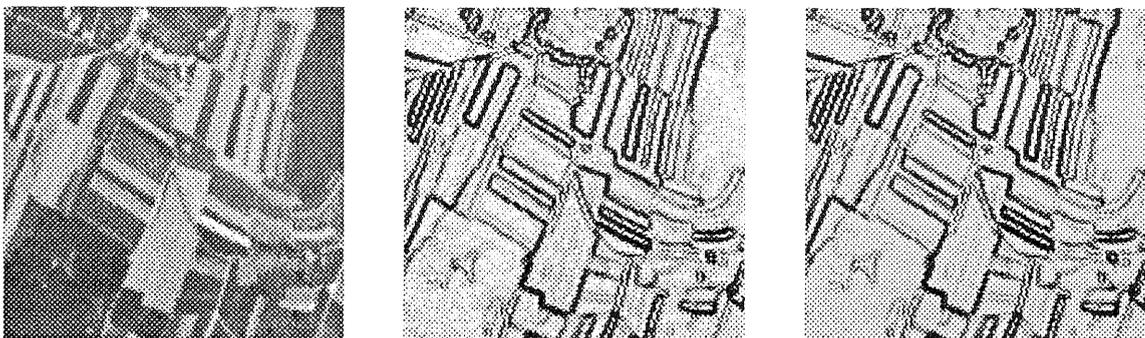


Figure 5 Grey level image (left), gradient magnitude image (middle), thresholded gradient magnitude image (right)

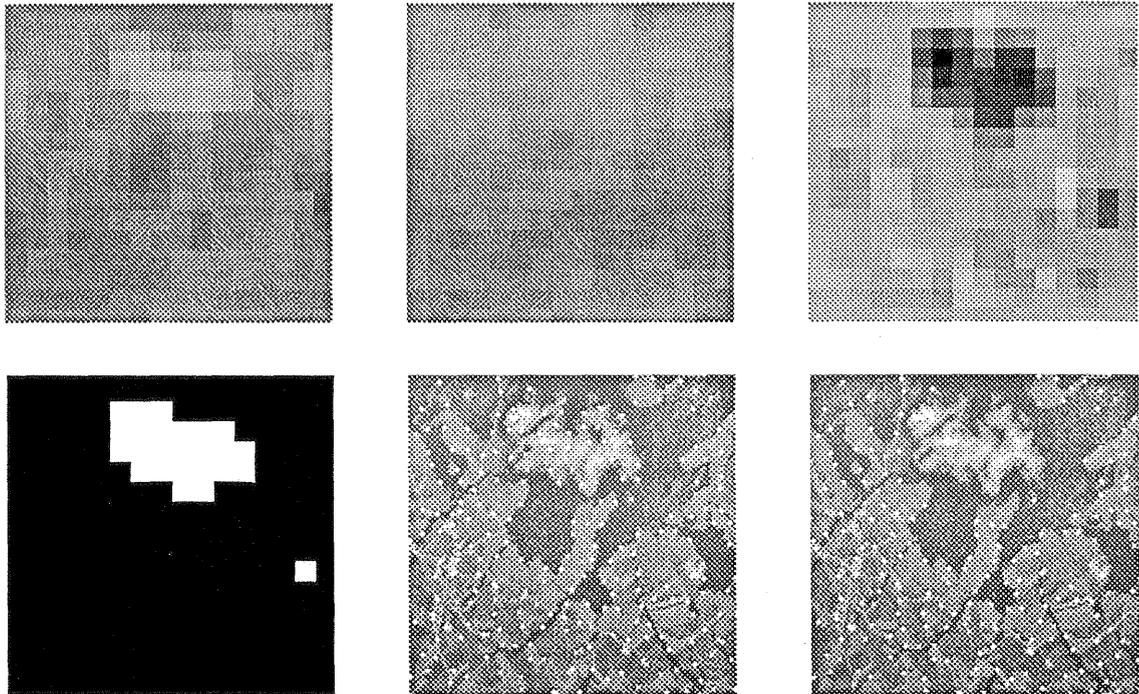


Figure 6 Top row: left (left) and right (middle) SPOT image at 4th pyramid level and normalized difference image (right). Bottom row: binarized difference image (left), image with selected points (middle), image with cleaned selected points (right)

were projected onto the search image by using an average height of 1000 m, and these pixel coordinates were compared to the known ones. The RMS differences were 32 pixels in x and 2 pixels in y, with the maximum error being 72 and 5 pixels respectively. Thus, a refinement of these approximations by an image pyramid approach is necessary. An alternative approach would be to actually transform and resample the search image by using the PMFs and the average height. In this case, the disadvantages are (i) the computational costs for the transformation and the resampling, and (ii) the degradation of the data. The advantages include: (i) matching can be performed using only shifts, thus resulting in computational gains which in case of many points exceed the losses, (ii) detection of radiometric differences can be applied as proposed above, and (iii) since the y-parallax of the co-registered images is very small, the images can be viewed stereoscopically (which anyway is required in digital photogrammetric workstations).

6. ACCURACY TESTS

The accuracy of the matching algorithm was tested by using the 25 m DTM of Switzerland which is generated by the Bundesamt für Landestopographie. The DTMs of the 1:25,000 map sheets 1224 and 1225 were acquired. Each DTM has 701 x 481 nodes in E-W and N-S direction respectively. The DTMs are produced by bicubic interpolation in x and y direction, whereby the known heights are supplied by digitised contours, lake contours and spot heights. The accuracy of these DTMs

was checked by bicubic interpolation of the heights of ca. 1000 spot heights and comparison to their known values which have an accuracy of 1 - 2 m. The 1224 DTM was derived from ca. 107,000 height values and has an accuracy (RMS) of 1.9 m. The height range is 900 m but the terrain is generally smoothly changing. The 1225 DTM was derived from 252,000 height values, has an accuracy of 4.1 m and a height range of 1500 m. Although it is not the most extreme case that can be encountered in Switzerland, the terrain is in most parts steep. Forests cover ca. 20% of map sheet 1224 and 35 - 40% of map sheet 1225. In the latter there are also lakes covering ca. 4% of the area. Some clouds were present. The radiometric differences were larger in map sheet 1224 which included agricultural areas.

The aim of this test was to check the accuracy potential of the algorithm. Thus, good approximations derived from the given DTMs were used. The measurement points and their approximations were derived as follows. First, an orthophoto for each DTM using one of the SPOT gradient magnitude images was generated. The points were selected in the first pyramid level of the orthophoto by using a thin-out window of 3 pixels (27183 and 26064 points in the map sheets 1224 and 1225 respectively) and were projected into the original orthophoto images. After exclusion of the points at the areas of radiometric differences through subtraction with the orthophotos from the second SPOT image, 20,180 and 22,592 points remained. The X, Y coordinates of these points were readily available (since the images were orthophotos) and the height was bilinearly interpolated from the given

DTMs. Using the object to image PMFs, pixel coordinates were derived and matching could be performed.

For matching the following 5 different versions were run:

Version 1: patch size 17 x 17, no geometric constraints, conformal transformation

Version 2: patch size 17 x 17, constraints, conformal transformation

Version 3: patch size 17 x 17, constraints, shifts only

Version 4: patch size 17 x 17, constraints, conformal transformation, grey level image

Version 5: patch size 9 x 9, constraints, shifts only

All versions used gradient magnitude images with the exemption of version 4 that used grey level images. The choice of these versions was based on preliminary investigations that were performed with some of the worst out of the 136 points. The aim was to compare constraints vs. no constraints, grey level vs. gradient magnitude images, conformal vs. shift transformation, and shifts with different patch sizes. The case of affine transformation was excluded a priori because in many cases it is not stable since the selected points lie at edges and thus two scales and one shear are often not determinable.

Some points were unsuccessfully matched either because they were transformed outside the search window or because they needed more than 20 iterations. Table 1 shows these results.

Table 1 Matching versions

Version	1224		1225	
	Successfully matched points	Iterations per point	Successfully matched points	Iterations per point
1	94.4%	5.6	97.1%	5.2
2	97.7%	4.1	98.5%	4.2
3	99.2%	3.4	99.2%	3.8
4	94.4%	4.9	97.2%	4.3
5	98.4%	3.6	98.9%	3.6

These results were analysed for automatic detection of blunders. The criteria that have been used for quality analysis are: standard deviation of unit weight from the least square matching, correlation coefficient between the template and the patch, number of iterations, x-shift (i.e. change from the approximate values), standard deviation of x-shift, y-shift, standard deviation of y-shift, and the size of the 4 shaping parameters (two scales, two shears). With the conformal transformation only two shaping parameters were used (one scale, one shear). After matching, the median (M) and the standard deviation of the mean absolute difference from the median (s(MAD)) were computed for each criterion. The median and the s(MAD) were used instead of the average and the standard deviation because they are robust against blunders. The threshold for the rejection of one criterion

was defined as $M + N \cdot s(\text{MAD})$. N was selected to be 3 for all criteria with the exemption of the number of iterations, the two shifts and the two scales which should be left to vary more ($N = 4$). A point was rejected (i) when one of its criterion did not fulfil the aforementioned threshold (relative threshold derived from the image statistics), or (ii) one of its criteria did not fulfil a very loosely set threshold, e.g. for the correlation coefficient 0.2 (absolute threshold, valid for all images). The same N and absolute thresholds were used for all versions. Table 2 gives information on the amount of rejected points.

Table 2 Points rejected by automatic blunder detection

Version	1224		1225	
	Percentage over successful points	Remaining good points	Percentage over successful points	Remaining good points
1	16.1%	15987	15.7%	18504
2	11.4%	17485	12.7%	19417
3	9.2%	18173	9.0%	20391
4	18.1%	15606	17.2%	18183
5	10.8%	17714	8.7%	20394

As it can be seen from Table 1 and Table 2, the amount of successfully matched points decreases and the percentage of detected blunders increases when (i) no geometric constraints are used (version 1), and (ii) grey level images are used (version 4). From the remaining versions, the ones using shifts result in more successful points because they are more stable (robust) than the one using the conformal transformation. The conformal transformation includes a scale which is not always well-determinable. Constrained matching needs less iterations per point than unconstrained version, especially when only shifts are used. The differences between the two shift versions are minimal although their patch size differs considerably. The above results are valid and similar for both map sheets in spite of the different terrain form and land usage.

For the accuracy analysis two comparisons were made:

- The matched points are bilinearly interpolated in the reference DTM grid and the differences between the interpolated heights and the heights as estimated by matching are computed (Table 3).
- And a new DTM was derived from the matched points and compared to the reference DTM (Table 5).

Table 3 Differences of estimated heights (cleaned data) to heights bilinearly interpolated in the reference DTM

Version	1224		1225	
	absolute max.	RMSE	absolute max.	RMSE
1	31.7	7.2	42.9	8.9
2	33.8	8.4	44.8	9.4
3	38.7	9.5	47.6	11.2
4	40.9	9.6	48.0	10.0
5	41.7	10.2	52.7	10.7

Table 4 Differences of estimated heights (raw data) to heights bilinearly interpolated in the reference DTM

Version	1224		1225	
	absolute max.	RMSE	absolute max.	RMSE
1	203.3	11.5	158.8	12.6
2	175.1	12.6	157.5	13.3
3	251.2	14.0	236.8	15.9
4	198.8	19.5	224.1	17.1
5	248.7	19.0	298.3	18.3

Table 5 Differences between new and reference DTM

Version	1224		1225	
	absolute max.	RMS	absolute max.	RMS
1	94.3	8.9	271.5	19.9
2	93.6	9.5	235.6	18.6
3	94.7	10.4	191.6	18.5
4	107.8	10.9	233.0	19.3
5	98.6	10.4	52.7	17.0

Table 3 represents the accuracy of our matching approach. The accuracy is in the subpixel level! The figures of Table 5 are worse due to interpolation errors (330,000 points were interpolated from 16,000 - 20,000 points). Still the results for map sheet 1224 are close or less than 10 m. The results for map sheet 1225 are worse due to the mountainous terrain, many forests and the lake. With denser measurement points they should be close to the results of map sheet 1224 as Table 3 also indicates.

Version 1 (without constraints) is surprisingly good. The reason is that the approximations were very good. Additionally the points were chosen along nearly vertical edges. Thus, the precision in x-direction is good and errors in y (gliding along the edge) influence minimally the estimated heights due to the horizontal base. Additionally, the results of version 1 are based on fewer points due to many detected blunders (Table 2). This

reduced density, however, influence the accuracy as it can be seen for map sheet 1225 (Table 5). The advantages of the use of the constraints will become more apparent in a realistic case when the approximate values are poorer.

Version 4 is worse than the similar version with gradient magnitude images (version 2). The difference is not so big again due to good approximations and many reduced points for version 4. The shift versions (3 and 5) perform quite well. Version 5 gives the best results of Table 5 for map sheet 1225 due to the small patch size which models better the irregular terrain surface, and the large number of correct points which reduces the interpolation errors.

The improvement of the results due to blunder detection is remarkable. Table 4 shows the same results as Table 3 but for the raw data (including blunders). The results are as the average 37% worse than those of Table 3.

For visualisation the absolute differences d between the two DTMs which are higher than the threshold value t are combined with the orthophoto and marked as white areas (Figure 7 and Figure 8). The new DTM was derived from the points of version 2 and the threshold is defined by:

$$t = \bar{d} + \text{RMS}(d) \quad (2)$$

with \bar{d} mean of absolute differences.

Differences higher than the threshold can be found especially in three types of areas (Figure 7 and Figure 8 at a, b and c):

- (a) At the mountain-ridges and cliffs. At these regions there are surface discontinuities and forests. Additionally interpolation errors occur because the density of the selected points was low at these regions and thus the terrain surface could not be modelled correctly (see Figure 9 with the triangles used for DTM interpolation).
- (b) At forest areas, because the matched points are on the tree tops and the reference DTM refers to the earth surface.
- (c) On the lake surface. The selected points lied on either sides of the lake, and at certain places much higher than the lake surface. Thus, the large triangle that were used for the DTM interpolation (Figure 9) were lying much higher than the lake surface.

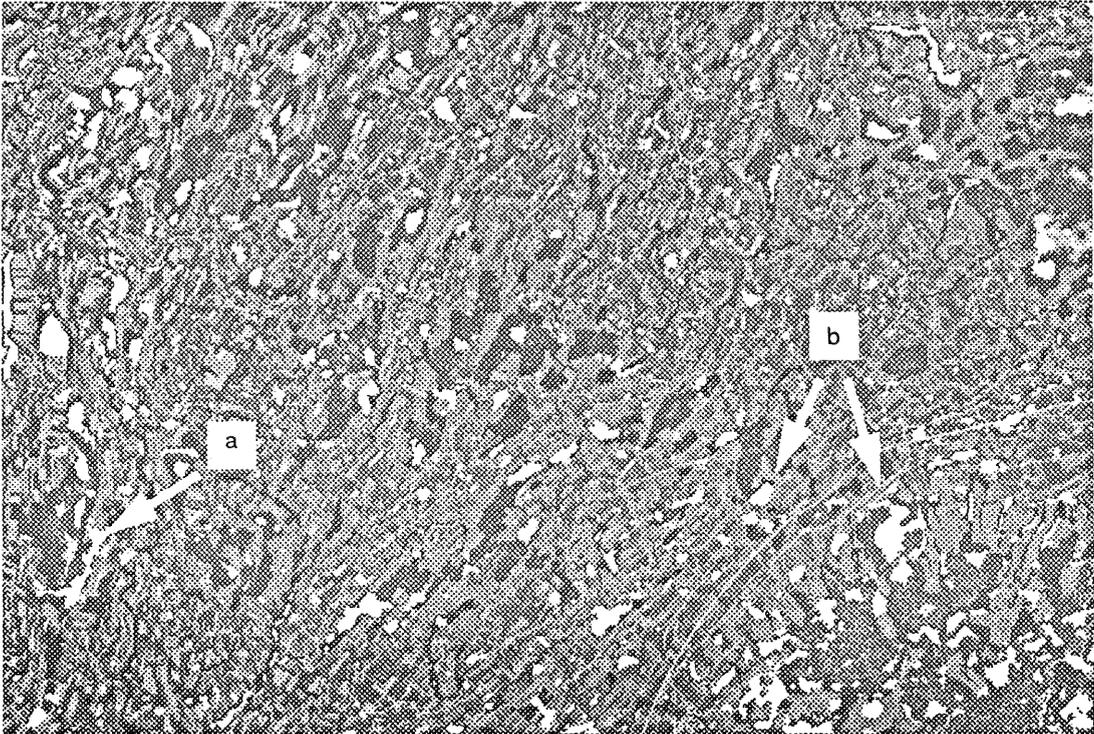


Figure 7 Orthophoto of map sheet 1224 with overlay of DTM differences > 17 m (white areas)

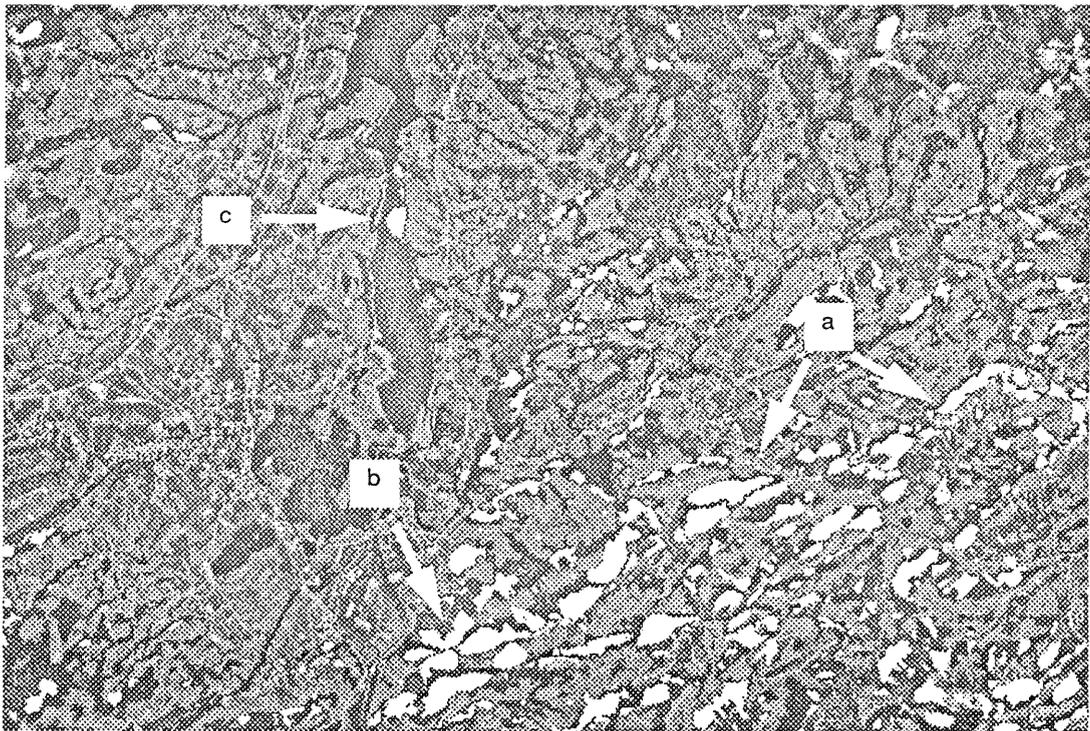


Figure 8 Orthophoto of map sheet 1225 with overlay of DTM differences > 30 m (white areas)

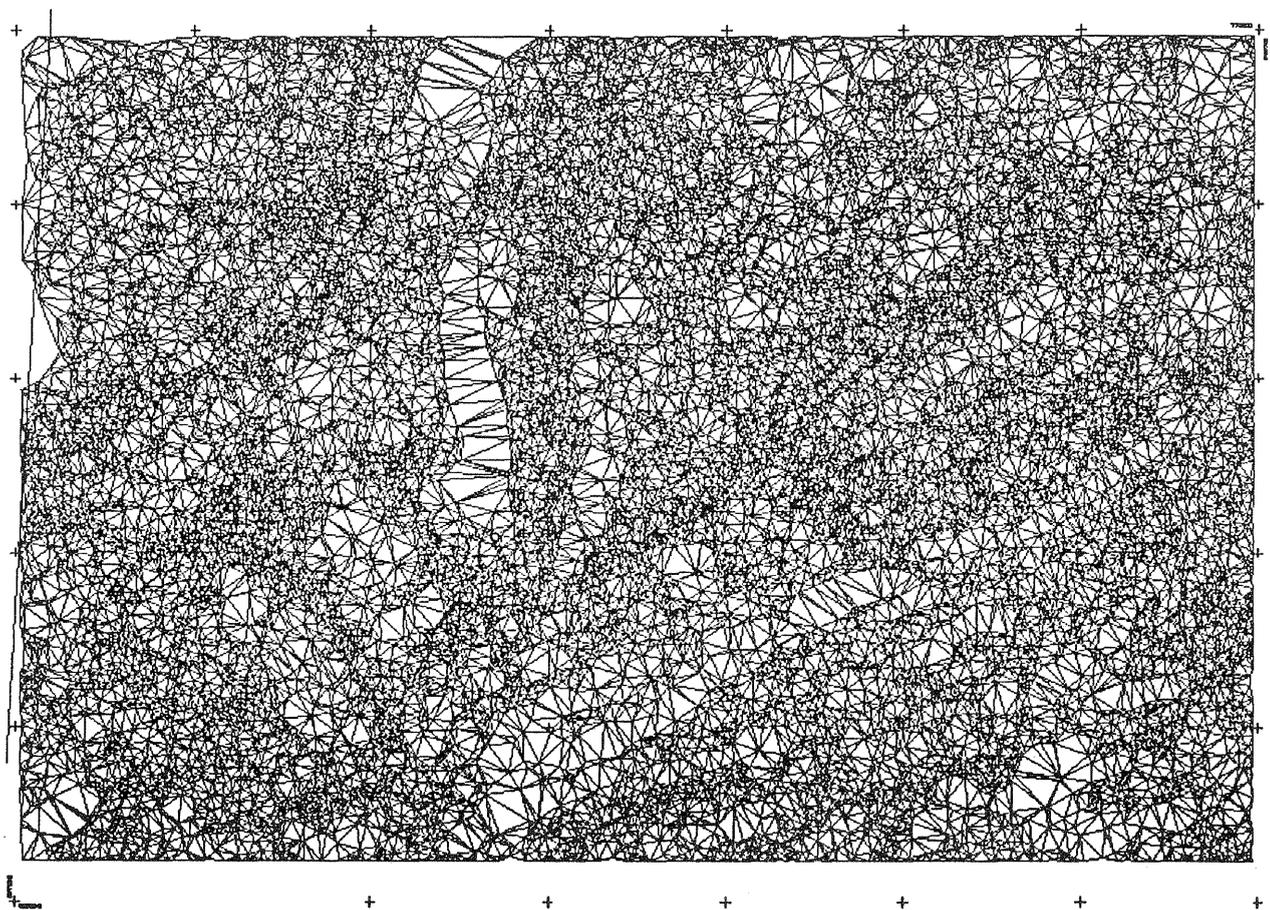


Figure 9 Triangulation mesh for DTM interpolation of map sheet 1225

7. CONCLUSIONS

A matching algorithm for SPOT images was presented that uses the SPOT geometry to impose constraints that reduce the search space from 2-D to 1-D. The algorithm severely reduces the problems caused by radiometric differences can simultaneously use any number of images (more than two) and determines in on step pixel and object coordinates. The use of gradient images magnitude instead of grey level images improves the results. A conformal transformation is suggested. However, the use of only shifts also leads to good results, slightly inferior to those of the conformal transformation.

Problematic cases like multiple solutions and occlusions are reduced and the computation times decreases due to the 1-D search. An intelligent blunder detection scheme is proposed that uses criteria derived solely from the statistics of the results.

The accuracy of the matching is excellent (RMS less than 10 m for 36,000 check points). The accuracy of the interpolated DTM depends clearly on the density of measurement points and for sufficient density it can also be in the 10 m range.

8. ACKNOWLEDGEMENTS

The authors express their gratitude to the Bundesamt für Landestopographie, Bern, Switzerland for providing the SPOT images and DTM data and appreciate their cooperation in the accuracy analysis of the results.

9. REFERENCES

- Baltsavias, E. P., 1992. Multiphoto Geometrically Constrained Matching. Ph. D. Dissertation, Mitteilungen Nr. 49, Institute of Geodesy and Photogrammetry, ETH Zurich, 221 p.
- Baltsavias, E. P., Stallmann, D., 1992. Metric information extraction from SPOT images and the role of polynomial mapping functions. In: Proc. of 17th ISPRS Congress, Commission IV, 2 - 14 August, Washington D. C., USA.
- Kratky, V., 1989a. On-line aspects of stereophotogrammetric processing of SPOT images. PERS, Vol. 55, No. 3, pp. 311 - 316.
- Kratky, V., 1989b. Rigorous photogrammetric processing of SPOT images at CCM Canada. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 44, pp. 53 - 71.