SURFACE RECONSTRUCTION FROM DATA OF DIGITAL LINE CAMERAS BY MEANS OF OBJECT BASED IMAGE MATCHING

H. Diehl[1], C. Heipke[2]

[1] Department ZTA 21
Messerschmitt-Bölkow-Blohm GmbH
Post Box 80 11 09, D-8000 Munich 80, Germany
Tel: +49-89-607 28651; Fax: +49-89-607 25157

[2] Chair for Photogrammetry and Remote Sensing
Technical University Munich
Arcisstr. 21, D-8000 Munich 2, Germany
Tel: +49-89-2105 2671; Fax: +49-89-280 95 73; Telex: 522854 tumue d
E-mail: heipke@photo.verm.tu-muenchen.de

Commission III

ABSTRACT:

This paper deals with automatic surface reconstruction from data of digital line cameras in the framework of object based image matching. In this approach grid heights of a digital terrain model (DTM), the elements of exterior orientation and grey values of an orthoimage are simultaneously determined from the recorded image grey values in a least squares adjustment. In order to provide initial values for the unknowns a hierarchical image representation is used.

For the investigations simulated satellite images from a high mountain area (the Vernagtferner in the Austrian Alps) recorded with a 3-line camera are used. The elements of exterior orientation are assumed to be given for every set of three scan lines. The influence of white noise added to the grey values of the three image strips and of random errors in the exterior orientation is investigated.

The results show that this approach can tolerate a large amount of noise in the image (up to a standard deviation of more then 10 grey values), but is very sensitive to errors in the exterior orientation.

Key words: Image matching, 3-line camera, object surface reconstruction, algorithm

1. INTRODUCTION

In recent years the main topic in photogrammetric research was the investigation of digital or softcopy photogrammetry. There exist two different aspects: on the one hand, classical photogrammetric methods are supplemented or replaced by digital ones, and on the other hand new possibilities are discovered, which are especially suited for digital data.

This article deals with a task of digital photogrammetry: the extraction of height information from line imagery by means of object based image matching. There already exist a large number of investigations, publications and implemented procedures for image matching. Therefore, we first present a brief review of the different methods.

The methods of analytical photogrammetry have at first been transferred directly to digital photogrammetry, the computer had simply replaced the human operator. That is, the computer had to find conjugate points in stereo image pairs, but the subsequent calculations were still the same, regardless whether the image points had been found by man or machine.

For point matching well known methods from statistical mathematics have been developed like correlation of grey value distributions /Sharp et al. 1964; Helava 1976; Panton 1978; Hannah 1988/. Subpixel accuracy was

achieved with least square matching in combination with higher dimensional search spaces /Förstner 1982/. Also interest operators were developed to automatically find suitable points for image matching /Marr, Poggio 1979; Barnard, Thompson 1980; Förstner 1986/ leading to feature based image matching. It became apparent that automatic single point matching has a relatively high rate of blunders. Robust methods in the subsequent calculations help to reduce them, but a number of problems remained especially in weakly textured image regions.

Least squares image matching was extended in a number of ways: to single point matching in object space /Grün 1985/, multi point matching /Rosenholm 1986/ and to matching complete images using object space models /Ebner et al. 1987; Wrobel 1987; Helava 1988/. At the same time more than two images were processed simultaneously. A hierarchical approach using pyramid structures /Burt, Adelson 1983; Rosenfeld 1984/, for the image resolution as well as for the digital terrain model (DTM) grid was introduced to provide initial values for the unknowns.

Feature based matching algorithms derive features like edges or corners from each image separately /Hahn, Förstner 1988; Schenk et al. 1990/. These features are much denser in the images compared with the number of points, which are measured by a human operator. They are then matched in a further step.

Up to now investigations of object based image matching were exclusively conducted for frame images. Digitally acquired frame images taken with CCD cameras are being used for close range photogrammetry, but for airborne imaging no camera with an adequate number of pixels per frame exists. Hence conventional film must be digitized, a procedure which introduces additional errors in geometry /eg Boochs 1984/ and in radiometry /Diehl 1990/.

Line cameras produce digital data directly. They have proven to be successful in space applications (SPOT, MOMS 01) and will soon be used in other spaceborne missions (MOMS-02/D2, /Ackermann et al. 1989/ and MARS 94, /Albertz et al. 1992/) and airborne projects (DPA, /Müller 1991/). The present investigation deals with object based image matching using 3-line camera imagery.

2. THEORETICAL BACKGROUND

2.1 3-line cameras

The concept of 3-line cameras for digital photogrammetry has been investigated for more than a decade /Hofmann et al. 1982; Ebner, Müller 1986; Heipke et al. 1990; Müller 1991/. For 3-line camera imagery there exist discrete projection centres for each image, but one image consists only of three lines. For the evaluation of the three image strips, the projection centre must be considered as a 6-dimensional function of time. An image strip is not expressed in terms of spatial parameters (x,y), but in time/space (t,y). In the following, the row (x) of an image strip will correspond to the time t, and the column (y) corresponds to the pixel number within the sensor line. With expensive equipment, the flight path can be recorded synchronously to the read out clock of the sensors. It can also be sampled in larger time steps and interpolated in between. In the latter case, additional interpolation errors arise.

2.2 Object based image matching

The extraction of DTM from stereoscopic images can always be seen as searching for an optimal geometric transformation $T_{1,2}$ between two images $T_{1,2}$: image 1 → image 2, or $T_{i,j}$ between several images. In single point image matching this transformation is evaluated for discrete points. Then, a grid DTM is generated.

In multi point matching, the transformation is calculated directly for (a part of) the grid. It is established in one of the images and transformed to the other (see fig. 2.1).

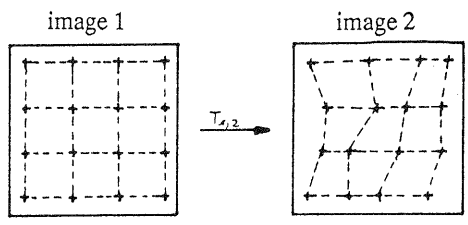

Fig. 2.1 Multi point transformation $T_{1,2}$ expressed by a regular grid in image 1 and the transformed one in image 2

If the matching process is formulated in object space (fig. 2.2), the grid is established there and transformations $T_{i,o}$ are set up between an image i and the object space o. Transformations between two images i,j can be derived from

$$T_{ij} = T_{i,o} \bullet T_{o,j} = T_{i,o} \bullet T_{j,o}^{-1} \qquad (1)$$

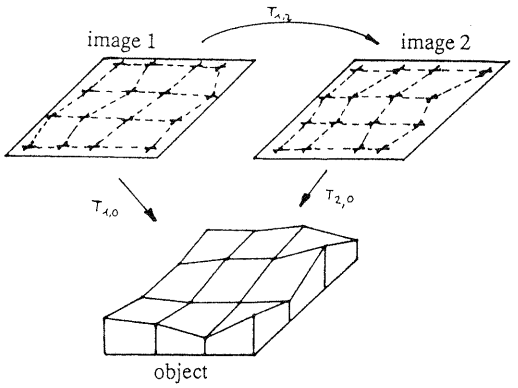

Fig. 2.2 Object based image matching transformations. The grid is established in object space, the inverse transformations $T_{1,o}^{-1}$ and $T_{2,o}^{-1}$ are the projections into the image planes.

There is a number of advantages in object based image matching:

- symmetry between all images (no master and slave images exist),
- epipolar constraints are given implicitly, even for more than two images,
- more than two images and different spectral bands can be introduced,
- point determination, DTM generation, and orthoimage computation are included in the approach.

2.3 3-line cameras and object based image matching

The transformations $T_{i,o}$ are inverse to the well known collinearity equations:

For line imagery, these equations have another meaning. The image coordinate x_b can't be calculated, it is constant and equivalent to the distance between the sensor and the principal point in image space. The elements of exterior orientation have to be expressed as time dependent functions. Instead of calculating the image coordinate x_b at a given moment t, we must calculate the moment t from equation (2a) using x_b = constant. In general this calculation must be carried out iteratively. Afterwards y_b is computed from equation (2b) with the value of t calculated from (2a).

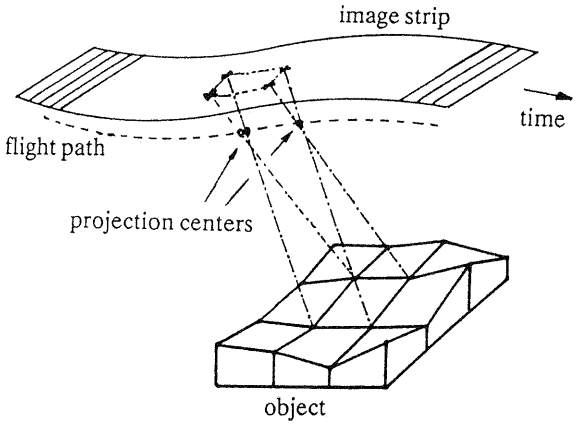

Fig. 2.3 Projection of a DTM mesh to an image strip.

3. THE MATCHING ALGORITHM

In this chapter the developed object based image matching algorithm for line imagery is described. It based on work by Heipke /1990/ and Müller /1991/. First we introduce a model for the object space, the flight path and the imaging process. Then we describe the least squares adjustment to compute values for the independent unknowns of our approach. At last we show how the necessary initial values for the unknowns can be calculated using a pyramid structure.

$$x_b = -c\,\frac{r_{11}(t)\,(X - X_0(t)) + r_{21}(t)\,(Y - Y_0(t)) + r_{31}(t)\,(Z - Z_0(t))}{r_{13}(t)\,(X - X_0(t)) + r_{23}(t)\,(Y - Y_0(t)) + r_{33}(t)\,(Z - Z_0(t))} \qquad (2a)$$

$$y_b = -c\,\frac{r_{12}(t)\,(X - X_0(t)) + r_{22}(t)\,(Y - Y_0(t)) + r_{32}(t)\,(Z - Z_0(t))}{r_{13}(t)\,(X - X_0(t)) + r_{23}(t)\,(Y - Y_0(t)) + r_{33}(t)\,(Z - Z_0(t))} \qquad (2b)$$

3.1 The models

To model the object space we make the following simplifying assumptions: the considered object surface is rigid, not moving and opaque. It is illuminated by external light sources with time constant light.

In the geometric model we assume that the object surface is a function over the XY-plane, that is, for each arbitrary point (X,Y) in the coordinate plane there exists exactly one point (X,Y,Z) on the object surface. Thus, we can formulate the height Z as Z(X,Y). To describe this function with a finite number of height values we establish a grid in the XY-plane with nodes (X_k,Y_l) and grid heights $Z_{kl} = Z(X_k,Y_l)$. A height Z(X,Y) at an arbitrary point is interpolated from the neighbouring grid heights eg by bilinear interpolation. The mesh size is constant and is chosen based on the roughness of the terrain.

In the radiometric model we assign an object grey value G(X,Y) to every point on the object surface (X,Y,Z(X,Y)). G(X,Y) is assumed to be constant in time and independent on the direction of observation. To describe this grey value function with a finite number of grey values G_{ij}, we establish another grid in the XY-plane with nodes (X_i,Y_j) and grid values $G_{ij} = G(X_i,Y_j)$. G can be considered as an orthoimage of the terrain and the grid nodes (X_i,Y_j) as the centres of so called object surface elements. The distance between neighbouring nodes is constant, too. It should be chosen according to the pixel size of the original images. Usually the grey value grid has a much smaller grid size than the height grid (X_k,Y_l).

The flight path is a one-dimensional function, a function in time, with 6 values. It is modelled by an equidistant series of moments t_m with exterior orientation elements $p_{m,n}$, $n = 1..6$, for position and attitude. Arbitrary moments t are linearly interpolated. The duration between two moments t_m, t_{m+1} should be chosen according to the roughness of the flight path. The moments t_m are called the orientation moments or orientation images.

For the imaging process, we use the pixel-by-pixel model of computing an orthoimage /Mayr, Heipke 1988/. For each object surface element the surface height $Z_{ij} = Z(X_i,Y_j)$ is interpolated from the neighbouring grid heights. The resulting point (X_i,Y_j,Z_{ij}) is projected into the image strip using the collinearity equations yielding the position (t,y_b). For this position a bilinear interpolation of the surrounding pixels is used to obtain the grey value for the object surface element (X_i,Y_j).

3.2 The least squares adjustment

The observations in our approach are the grey values $(g_{ij})_s$ of the image strips, s = 1,2,3. The independent unknowns are the height values $Z_{k,l}$ in the DTM grid points and the elements of exterior orientation $p_{m,n}$ in the orientation moments t_m. The object grey values G_{ij} are treated as unknowns too, but they are not independent from the $Z_{k,l}$ and the $p_{m,n}$. For a given image strip, a set of height values and orientation parameters, they are determined unambiguously.

For every object surface element (X_i,Y_j) the $(g_{ij})_s$ can be computed for each image strip s from the Z_{kl} and the $p_{m,n}$. Thus, for 3-line camera imagery three functions $F_s:(Z_{kl}, p_{m,n}) \rightarrow (g_{ij})_s$, s = 1,2,3 can be formulated. Each F_s can be separated into the three parts. H_1 is identical for all strips, H_2 and H_3 are different. The flow of computation can be represented as follows:

Object surface element (X_i,Y_j) and DTM heights Z_{kl}

↓ H_1 = Bilinear height interpolation in object space

surface point height $Z(X_i,Y_j)$

↓ H_2 = Projection to image strip using the modified collinearity equations (2)

projected position (t,y_b)

↓ H_3 = Bilinear grey value interpolation in image strip s

grey value $(g_{ij})_s$

The optimization criterium is to minimize the sum of squares of differences between the unknown object grey value G_{ij} and the observations $(g_{ij})_s$ over all pixels:

$$\sum_{i,j} (G_{ij} - (g_{ij})_s)^2 \rightarrow \min \qquad s = 1,2,3 \qquad (3)$$

(3) can be rewritten as a system of observation equations and represents the functional dependence between the independent unknowns $Z_{k,l}$ and $p_{m,n}$ and the optimizing criterion. So far there exists no fundamental difference to the frame image approach, but there is one difficulty: H_2 can't be expressed explicitly (see chapter 2.1). It is only given implicitly and must be calculated iteratively. Control information is added to (3) and the unknowns are solved for using the standard least squares formulae.

3.3 Initial values using a hierarchical approach

Non linear least squares adjustment needs initial values for the unknowns to start with. In image matching hierarchical procedures /Burt, Adelson 1983/ from coarse to fine have been used to provide them /eg Li 1989/. An image pyramid as well as a DTM pyramid are generated from level 0 (the original image and the DTM). For the next higher level 2*2 elements are combined into one element, thus the resolution is coarser by a factor of 2 and the amount of data by a factor of 4. In this way several levels are computed.

The adjustment process starts with some coarse initial height values (ideally a constant value) in the highest level of the image and DTM pyramid. In every level the iteration proceeds until an end condition is reached. Then the DTM values of the next lower level are calculated, eg by linear interpolation, and the computation proceeds on that level, until level 0 is processed.

4. CONDUCTED EXPERIMENTS

Real data from high resolution 3-line cameras with known flight path are not yet available. Therefore, the algorithm was tested with simulated data. The main goal of the simulations was to investigate the influence of white noise in the grey values and of random errors in the exterior orientation onto the matching results.

4.1 Input data

For the simulations we produced so called semi-synthetic image strips. We used a real orthoimage from the "Vernagtferner", a glacier in Austrian Alps /Rentsch 1992/ together with the corresponding DTM and generated three image strips by means of inverse orthoprojection. We used orientation parameters corresponding to a straight flight path with constant velocity. The simulation parameters were chosen as follows:

Ground elevation:	2620 m - 2950 m
DTM mesh size:	50 m * 50 m
Object surface element size:	3.125 m * 3.125 m
(eg 16 * 16 object surface elements per DTM mesh)	
Flight altitude:	300 km
Speed:	7700 m/s
Sensor pixel size:	10.4 μm
Sensor read out frequency:	2464 Hz
Radiometric resolution:	8 bit
Convergency angles:	2 x 20 grad
Calibrated focal length:	1.0 m

The used DTM can be seen in figure 4.1. In figure 4.2 the semi-synthetic image strip for the nadir looking sensor (1024 * 1024 pixels) is shown. The central part consisting of 256 * 256 pixels was used in the simulations. The image texture in this area is rather good.

Radiometric noise with different standard deviations was added to the grey values of all image strips. $\sigma_r = 3.0$ grey values corresponds to a well calibrated CCD sensor, $\sigma_r = 12.0$ to the noise, which must be expected in digitized films /Diehl 1990/. In the case of $\sigma_r = 0.0$ no noise was introduced. However, the generation of the image strips inherently induces quantization noise and interpolation errors.

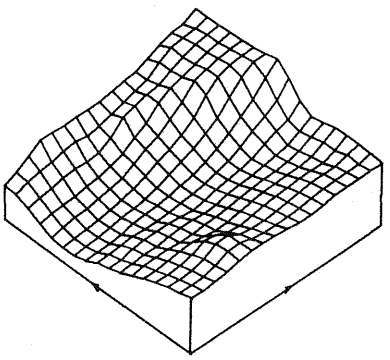

Fig 4.1 The Vernagtferner DTM used for the simulations

Fig. 4.2 The semi-synthetic image strip of the nadir channel

pyramid level	number of pixels	number of DTM meshes	size of object surface elements [m^2]
8	4 * 4	1	800 * 800
7	4 * 4	1	400 * 400
6	4 * 4	1	200 * 200
5	8 * 8	1	100 * 100
4	16 * 16	1	50 * 50
3	32 * 32	2 * 2	25 * 25
2	64 * 64	4 * 4	12.5 * 12.5
1	128 * 128	8 * 8	6.25 * 6.25
0	256 * 256	16 * 16	3.125 * 3.125

Table 4.1: Pyramid structure

Random errors with different standard deviations in position (σ_p = 0.4 m, 1.0 m, 2.0 m) and attitude (σ_a = 0.08 mgrad, 0.2 mgrad, 0.4 mgrad) were added to the elements of exterior orientation for every 10th line. These errors can be regarded as originating from interpolation errors in the flight path.

4.2 Pyramid structure

For each image strip and for the DTM a pyramid structure was generated. Level 0 contained 256 * 256 pixels in 16 * 16 DTM meshes, the next level 128 * 128 pixels in 8 * 8 DTM meshes and so on. Level 4 contained 16 * 16 pixels in one DTM mesh. In the levels 5 and 6 only the number of pixels was reduced. For level 7 the number of pixels and of DTM meshes remained constant, but the size of the object surface elements was doubled in each direction. The same was done again for level 8. The details of the pyramid structure can be seen in table 4.1.

4.3 Results and conclusions

Matching was performed through the pyramid as described. The elements of exterior orientation (straight flight path and disturbed flight path respectively) were introduced as constant values in all cases. For the initial height values a horizontal plane was used. Starting heights between 1500 and 4000 m all yielded the same result. This shows that the convergence radius, which amounts to only a few pixels in image space for least squares matching without image pyramids, can be extended nearly arbitrarily.

Table 4.2 shows the results in detail. In the ideal case (1) a standard deviation of 1.22 m between the known and the derived DTM was obtained. This corresponds to 0.1 pixel (1 μm in image space) or 50 μs in the image strip. It is caused by errors during the generation of the semi-synthetic images. It can also be seen that the radiometric noise does not influence the matching results to a large extend. The noise of well calibrated CCD sensors, case (2), can be neglected. Even in case (3) the results are still acceptable.

Errors in exterior orientation can be directly related to locations in image space. In case (4) and (5) they induce an error of about 0.5 pixel, in (6) and (7) about 1 pixel and in (8) and (9) about 2 pixels standard deviation. The derived errors in height show, how sensitive the approach is towards errors in the exterior orientation. Therefore in any practical application it should be determined simultaneously with the DTM heights. The additional introduction of grey value noise has only minor effects onto the results.

The algorithm was further tested with a very rough but exactly known flight path. In this flight path deviations of up to 10 m in position and 5 mgrad in attitude within 10 image lines (= 31.25 m in object space) occurred. New image strips were generated using this flight path. The matching results were exactly the same as for the straight flight path, case (1). Thus a rough flight path does not pose a problem as long as it is exactly known.

	Grey value noise	Flight path errors		Results: known DTM - derived DTM	
	σ_r [grey levels]	σ_p [m]	σ_a [mgrad]	stand. dev. [m]	mean [m]
(1)	0.0	0.0	0.0	1.22	0.09
(2)	3.0	0.0	0.0	1.50	0.10
(3)	12.0	0.0	0.0	3.50	0.39
(4)	0.0	0.4	0.08	2.51	0.15
(5)	12.0	0.4	0.08	4.69	0.17
(6)	0.0	1.0	0.2	6.89	0.81
(7)	12.0	1.0	0.2	8.01	1.63
(8)	0.0	2.0	0.4	19.86	2.33
(9)	12.0	2.0	0.4	19.09	1.07

Table 4.2: Numerical results of simulation

5. FUTURE ASPECTS

There exist various possible extensions of our approach. The introduction of a non-regular grid DTM adapted according to the roughness of the terrain, and the consideration of breaklines and discontinuities within the DTM can be achieved by altering the geometric object model. The simultaneous extraction of breaklines and discontinuities is more complicated to realize.

A special feature of 3-line cameras is a different resolutions in each channel. For instance the MOMS camera has a resolution of 4.5 m for the nadir, 13.5 m for the other stereo channels. The matching process can be adapted to different resolutions. It should yield significantly better results than those obtained from the coarser level only.

The simultaneous estimation of the parameters of exterior orientation in the least squares adjustment is already incorporated in the presented approach. As shown above this it is necessary for any practical application. Such investigations have still to be carried out.

6. REFERENCES

Ackermann F., Bodechtel J., Lanzl F., Meissner D., Seige P., 1989: MOMS-02 ein multispektrales Stereo Bildaufnahmesystem für die zweite deutsche Spacelab Mission D2, GIS (2) 3, 5-11.

Albertz J., Scholten F., Ebner H., Heipke C., Neukum G., 1992: The camera experiments HRSC and WAOSS on the Mars 94 mission, IntArchPhRS (29) 1.

Barnard S. T., Thompson W. B., 1980: Disparity analysis of images, IEEE-PAMI (2) 4, 333-340.

Boochs F., 1984: Ein Verfahren zur Herstellung digitaler Höhenmodelle aus photogrammetrischen Stereomodellen mit Hilfe der flächenhaften Korrelation in digitalen Bildern, DGK-C 299.

Burt P.J., and E. H. Adelson, 1983: The Laplacian pyramid as a compact image code, IEEE-Transactions on communications (31) 4, 532-540.

Diehl H., 1990: Radiometric noise in digitized photographs, IntArchPhRS (28) 5/2, 974-983 .

Ebner H., D. Fritsch, W. Gillessen, and C. Heipke, 1987: Integration von Bildzuordnung und Objektrekonstruktion innerhalb der digitalen Photogrammetrie, BuL (55) 5, 194-203.

Ebner H., Müller F., 1986: Processing of digital three-line imagery using a generalized model for combined point determination,IntArchPhRS (26) 3/1, 212-222.

Förstner W., 1982: On the geometric precision of digital correlation, IntArchPhRS (24) 3, 176-189.

Förstner W., 1986: A feature based correspondence algorithm for image matching, IntArchPhRS (26) 3/3, 150-166.

Grün A., 1985: Adaptive least squares correlation: A powerful image matching technique, South African Journal of Photogrammetry, Remote Sensing and Cartography 14 (3), 175- 187.

Hahn M., Förstner W., 1988: The applicability of a feature based and a least squares matching algorithm

for DEM acquisition, IntArchPhRS (27) B9, 137-150.

Hannah M. J., 1988: Digital stereo image matching techniques, IntArchPhRS (27) B3, 280-293.

Heipke C., 1990: Integration von digitaler Bildzuordnung, Punktbestimmung, Oberflächenrekonstruktion und Orthoprojektion in der digitalen Photogrammetrie, DGK-C 366.

Heipke C., Kornus W., Gill R., Lehner M., 1990: Mapping technology based on 3-line-camera imagery, IntArchPhRS (28) 4, 314-323.

Helava U.V., 1976: Digital correlation in Photogrammetric Instruments, IntArchPhRS (21) 2.

Helava U.V., 1988: Object-space least-squares correlation, PE&RS (54) 6, 711-714.

Hofmann O., Navé P., Ebner H., 1982: DPS - A digital photogrammetric system for producing digital elevation models (DEM) and orthophotos by means of linear array scanner imagery, IntArchPhRS (24) 3, 216-227.

Li M., 1989: Detection and location of breaklines and discontinuities in stereo image matching, Phot. Reports 54, Royal Institute of Technology, Stockholm.

Marr D., Poggio T., 1979: A computational theory of human stereo vision, Proceedings of the Royal Society of London B. 204, 301-328.

Mayr W., Heipke C., 1988: A contribution to digital orthophoto generation, IntArchPhRS (27) B11, IV-430-439.

Müller F., 1991: Photogrammetrische Punktbestimmung mit Bilddaten digitaler Dreizeilenkameras, DGK-C 372.

Panton D. J., 1978: A flexible approach to digital stereo mapping, PE&RS (44) 12, 1499-1512.

Rentsch M., 1992: Ein Farborthobild vom Vernagtferner, Master thesis, Chair for Photogrammetry and Remote Sensing, Technical University Munich (unpublished).

Rosenfeld A. (Ed.), 1984: Multiresolution image processing and analysis, Springer Verlag, Berlin.

Rosenholm D., 1986: Accuracy improvement of digital matching for evaluation of digital terrain models, IntArchPhRS 3/2, 573-587.

Schenk T., Toth C., Li J.C., 1990: Zur automatischen Orientierung von digitalen Bildpaaren, ZPF (58) 6, 182-189.

Sharp J.V., R. L. Christensen, W. L. Gilman, and F. D. Schulman, 1965: Automatic map compilation using digital techniques, PE&RS (31) 3, 223-239.

Wrobel B., 1987: Digitale Bildzuordnung durch Facetten mit Hilfe von Objektraummodellen, BuL (55) 3, 93-101.