Hierarchical Multiphoto Matching and D.T.M. Generation

Emmanuel P. Baltsavias
Institute of Geodesy and Photogrammetry
Swiss Federal Institute of Technology
ETH - Hoenggerberg
CH - 8093, Zürich, Switzerland
Commission III

Abstract

A hierarchical multiphoto matching was used for Digital Terrain Model (D.T.M.) generation from small scale aerial images. The procedure was performed automatically and included all stages from digitisation and camera calibration to the analysis of the results. Severe problems in the digitisation degraded the already poor quality original data and influenced all subsequent stages. The calibration of the sensor with respect to an analytical plotter proved to be more accurate than manual measurents. An image pyramid technique was used for derivation of approximate values for the D.T.M. The D.T.M. points, forming an irregular grid, were selected at the lowest pyramid level by an interest operator. The final matching was done using four images simultaneously and the collinearity conditions as constraints. At all pyramid levels a test for the elimination of blunders is applied. The matching results are compared to manual measurements. The accuracy level of both type of measurements is similar. The matching RMS, including the errors from all stages and other external factors, is 0.2 ‰ of the flying height over the terrain (hg). The reliability of the algorithm is high. Most of the blunders were detected. The percentage of undetected blunders was 1.5 % and the max absolute error 1.1‰ hg. The reliability was strongly related to the number of rays per point. The very high theoretical precision of 0.01‰ hg shows the accuracy potential of the algorithm but also the amount of problems still to be solved. The advantages of use of multiple images and geometric constraints were verified and are reflected in the reached accuracy and reliability. A feature based selection of the D.T.M. points instead of a regular grid also contributed to the results and should be prefered.

Introduction

Over the last decade extensive research has been conducted in the application of digital methods, especially image processing techniques, in photogrammetry. Various hardware and software, on-line and off-line solutions have been proposed for the automatisation of almost all standard photogrammetric tasks. A problem of central interest is the automatic acquisition of D.T.M. data. D.T.Ms gain continuously in importance and applications and the aim to automate their production is justified by the vast amount of primary data required for their production. Recently there is an increasing involvement of photogrammetry in nonconventional applications, especially in industry and medicine. In many of these applications a measurement of an object surface is required. The term D.T.M. is inappropriate in this case, so the term Digital Surface Model will be used in general for an object surface. Although there are differences between aerial and close-range applications, the heart of the problems for an automatic D.S.M. generation remains the same. A digital, automatic procedure for D.S.Ms is of interest and finds applications in different disciplines and gives photogrammetry possibilities to expand beyond the traditional applications which seem to stagnate.

Some of the commercially available solutions for automatic D.S.M. generation include the Kern DSR-11 analytical plotter (Bethel, 1986) and the Zeiss Planicomp (Pertl, 1984), both eqipped with CCD cameras for digitisation and appropriate hardware and software. Zeiss offers two packages Indusurf (Schewe, Foerstner 1986) and Toposurf for industrial and topographic applications respectively. Apart from the tendency to upgrade the existing analytical plotters there are parallel efforts to create a fully digital universal photogrammetric equipment, termed by many digital plotter. An automatic D.S.M. generation could be incorporated in both.

This paper addresses the problems of automatic D.T.M. generation from small scale aerial images and particularly the accuracy aspects. Additionally, an effort was made to face the different parts of the whole procedure from camera calibration to the analysis of the D.T.M. results in their relationships and perform it digitally. The work is based on the algorithms described in Gruen, Baltsavias, 1987a, 1987b. The Multiphoto Geometrically Constrained Matching (MGCM) was used for simultaneous matching and X,Y,Z object coordinate determination. Due to the use of more than two images and the use of the collinearity equations as constraints for the transformation of the digital images to be matched, the accuracy, success rate and especially the reliability of the matching are increased. New aspects, not reported in the above papers, include the derivation of approximate values by an image pyramid, and use of interest operators for selection of D.T.M. points in a irregular grid. The whole procedure including camera calibration, interior orientation and D.T.M. generation was performed automatically, off-line on a Sun 3 computer. By camera calibration it is meant, in this case, the geometric calibration of the sensor with respect to the analytical plotter. Only the required outer orientation was performed manually at the AC1 analytical plotter. Finally the D.T.M. points were measured manually and the two sets of measurements were compared. It must be noted here, that it was irrelevant whether the points were useful for a D.T.M. or not (points on houses, trees etc. were also measured). The purpose of this test was to check the accuracy and reliability of the algorithm and not its ability to find automatically the type of the points. This problem cannot be solved automatically at the moment and requires manual intervention.

Test data

The image data consist of a 2 x 2 block in the Swiss Alps with 60% overlap in both directions. In this 5.5 x 6.5 km region the height differences span a range of 1500m. The scale was 1:42500 and the flying height over the terrain 6500 m on the average. The camera was a Wild RC 10 with focal length 153 mm. The film was a KODAK Panatomic X, but its quality was very poor. The steep gradient of the characteristic curve of this film is ideal for signalised points but leaves out small density differences, which makes this film inappropriate for D.T.M. measurements, where the contrast of the natural points used might be low. In addition to this disadvantage the grain noise was high. From the overlap region 4 squares of varying terrain use and slope were selected and digitised in all 4 images. The 4 squares in one of the images are shown in Fig. 1. Although the image scale is quite small, the perspective differences among the images are by no means negligible. Fig. 2, with square 3 in two images, shows clearly the existing geometric and radiometric differences. The digitisation was done with an AQUA-TV HR 600 frame transfer CCD camera which was mounted at the AC1 (Fig. 3). The camera and the measuring mark are fixed and by moving the stages different parts of the film can be digitised. The objective lens is separated from the camera body. By moving both vertically, the imaging ratio can be changed from 1:0.5 to 1:1.8. Only one camera under the right stage was used. The computer pixel spacing was 10.2 μm in x and 10.96 μm in y, the size of the images 512^2 pixels, so each image covered approximately 5 x 5 mm on the film. The frame grabber of the Kontron IPS system was used. There was no attempt to calibrate radiometric and geometric distortions of the camera sensor. For more details on camera calibration and especially for this camera see Guelch 1986, Beyer 1987, Daehler 1987. Each image was averaged 32 times to reduce the noise. The camera was working under the automatic gain mode but the AC1 illumination was also controlled and corrected manually to avoid blooming and smearing and increase the image contrast. The camera and the AC1 were turned on many hours before each digitisation to avoid warm-up effects. Problems occured with dust and dirt on the AC1 plates, the film and the sensor itself. The latter is due to the separation of the lens from the camera body. The stage of digitisation and the acquisition of good quality data is very important and can adversely influence the following steps. In this test, the combination of poor quality film and the problems of the CCD-frame grabber system resulted in degraded data and definitely affected the accuracy of the automatic measurements. The errors vary considerably from image to image although the digitisation parameters, except the illumination, were kept constant. This shows that the performance of the given CCD-frame grabber was very

instable and that time varying errors, which are difficult to model, are the most important degradation sources. The noise level, especially the grain noise, was also enhanced by the small pixel size. Grain noise is multiplicative and influences both the sampling interval, which could not be altered, and the accuracy of the grey level quantisation. The data was preprocessed to solve the following problems:

a) Alternating dark and light horizontal lines due to grey level shift between the two fields of the digital image. This sort of noise seems to be additive. The grey level shift between adjacent lines is on the average 4-5 and may be due to the AC power supply - either to the camera, the frame grabber or the object illumination.
b) Alternating vertical dark and light lines. Each line has a width of 3 pixels and the dark lines are darker every 12 pixels. These lines seem to be interference patterns caused by phase differences in the 3 read out registers.
c) Partwise low contrast.

The first two problems were faced by convolving the original image sequentially by a 3 x 3 and a 7 x 3 filter specially designed to reduce the above mentioned errors. These filters were low pass filters, so they reduced also the noise in general, reduced the contrast and smoothed the edges. The average of the noise is positive which means that the amount of noise is larger for high grey level values. Then the grey level of the images were stretched in the range 0-255. Fig. 4 shows a twofold enlargement of a part of square 2. Fig. 5, 6, 7 show the result after each of the three preprocessing steps. Fig. 8, 9, 10 show the difference of each preprocessed image from the image of the previous step. The positive differences appear black, the negative ones are wrapped around the 255 value and appear white. An interesting effect, which was observed in all images, is that the amount of noise is higher at the top rows of the image which correspond to the bottom (next to the storage area) of the integration area of the sensor chip (Fig. 11).

Geometric calibration

It consists of the determination of two transformations. Pixel to carriage coordinates and carriage to image coordinates.

Pixel to carriage coordinate transformation

For this purpose the nine engraved calibration crosses on the carriage plates of the AC1 and the measuring mark were used. To drive to the crosses we used the AC1 program CMO without any film on the stages. The advantage of this program is that it drives automatically to the center of the crosses which are very accurately measured during the plotter's calibration, so inaccuracies and bias from manual measurements are avoided. The carriage coordinates of the centers of the crosses are known. The measuring mark is in every digital image and its carriage coordinates were read in the AC1's LED display. The display shows up to the μm decimal position, so a max 0.5 μm round off error might be introduced at this stage. A shifted local carriage coordinate system with origin the measuring mark was used, since the measuring mark is the only common and known point existing in all images. A 5 x 5 grid in the digital image was simulated by using the central cross and moving the stage, so that the cross appeared each time at one of the 5 x 5 positions. The pixel coordinates of the cross center were determined automatically. Having the carriage and pixel coordinates of 25 points, a transformation from one system to the other was computed with least squares adjustment. In the first attempt the pixel coordinate system was also centered to the measuring mark and a four parameter (two scales, two shears) tranformation was used. This centering required the determination of the position of the measuring mark. This was done by automatic driving to all 9 cross centers, digitisation of the images and automatic determination of their centers. The result of this tranformation is biased because of the absence of the shifts. The sum of the x and y residuals were not zero. Thus, we transformed using a full affine transformation. The affine transformation can also be used with the original (not centered) pixel coordinate system. The results of the two approaches are equivalent. The latter has the advantage that it does not require the pixel coordinates of the measuring mark. It is suggested though, that this operation is done as it is useful to check the stability of the sensor during time. To find the coordinates of any pixel in the plotter's

carriage coordinate system, first a transformation of the pixel coordinates in the local carriage coordinate system should be done and then the carriage coordinates of the measuring mark, which can be read in the LED display, should be added.
The automatic determination of the cross center was done in two steps. First a coarse matching, which utilises the known cross geometry and grey level gradients, is used. The algorithm gives approximate values better than 0.5 pixel, does not require that the cross is entirely in the image and needs 1.3 sec CPU time on the average. The quality of the approximate values is judged by the amount they change after the fine matching. The results are listed in 1 and an example showing one cross image and the determined position is in Fig. 12. The approximate values are refined by a 2D least squares matching with a synthetically generated ideal cross. The size of the image patches to be matched were 13^2 pixels and all affine parameters were used. The results are summarised in Table 1 and an example is in Fig. 13. The top row shows the search windows with the affine transformed patches inside, the starting position (square) and the final position (cross), and the bottom row the patches to be matched. The rightmost images refer to the reference image (template).

	| x-shift | (pel)	| y-shift | (pel)	stand. dev. x-shift	stand. dev. y-shift	correlation coefficient	So	number of iterations
average	0.213	0.152	0.008	0.011	0.97	5.28	3.3
minimum	0.000	0.005	0.008	0.010	0.96	3.82	2.0
maximum	0.530	0.456	0.009	0.012	0.99	6.37	5.0

Table 1. Results of coarse and fine matching of the calibration crosses

The determination of pixel to carriage coordinate transformation should be repeated from time to time since the whole AC1/camera system is not stable, but it is not required for each image to be digitised. The digitisation in our case was done in two phases with 2.5 months difference and before each digitisation this calibration step was performed. The results are listed in Table 2.

	x-scale (µm)	y-scale (µm)	x-shear (µm)	y-shear (µm)	RMS (µm)
1 calibration	10.1917	10.9612	0.0933	0.1013	0.8
2 calibration	10.2053	10.9557	0.0885	0.1028	0.9

Table 2. Results of pixel to carriage coordinates transformation

The shift of the measuring mark's pixel coordinates is expected but its magnitude is too large, if we consider that the cameras and the AC1 were switched on days before (Table 3.). The cause might be the fact that they were switched off between the two digitisations. The scales on the other side seem to be more stable, while the shears are negligible. The changes are larger in x than in y direction which is in accordance with previous investigations showing that the CCD camera errors are larger in x. It must be noted here that distortions introduced in the optical train from carriage to the sensor were not investigated.

	x average	x stand. dev.	y average	y stan. dev.
1 calibration	238.543	0.074	229.869	0.099
2 calibration	235.600	0.070	230.491	0.081
difference (pel)	2.943	0.004	-0.622	0.018
difference (µm)	30.010		-6.820	

Table 3. Pixel coordinates of measuring mark

Carriage to image transformation

The AC1 program PMO was used to drive to and digitise the 8 fiducials of each film. The fiducials are four at the corners (group 1) and four in the middle of each film side (group 2), which are not so precisely calibrated. Their only difference in the form is the orientation of the four lines pointing to the fiducial centre. The pixel coordinates of the fiducial centers were found out automatically, and they were tranformed to carriage coordinates by using the transformation determined in the previous stage. The first stage of the calibration is performed in diapositive mode, while the second one can be either in diapositive or negative mode, so appropriate modifications should be made for each case. Then, an affine transformation from carriage to the image coordinates, known from the aerial camera calibration, was determined with least squares.

The pixel coordinates of the fiducial centres were determined again in two stages. For the coarse matching the known geometry and the grey level gradients were used. Only the circular ring was utilised in order to increase the reliability of the match and be able to process both groups of fiducials. The approximate values were less than 0.5 pixel and required 2.5-3 sec CPU time. The results are listed in Table 4 and an example is in Fig. 14. Fine matching was performed as explained before. Two synthetic templates, one for each fiducial group, were used. The patch size was 71^2 to include except from the center also a good part of the four lines pointing to it. Small dimensions are not suggested because in some cases the center is not imaged well or its form is changed due to noise, dirt etc. The results are in Table 4 and an example in Fig. 15.

	| x-shift | (pel)	| y-shift | (pel)	stand. dev. x-shift	stand. dev. y-shift	correlation coefficient	So	number of iterations
average	0.101	0.212	0.002	0.002	0.98	7.16	2.8
minimum	0.007	0.006	0.002	0.002	0.94	5.19	1.0
maximum	0.263	0.523	0.002	0.002	0.99	10.91	3.0

Table 4. Results of coarse and fine matching of the fiducials

The transformation was estimated by using all fiducials and also the fiducials of group 1 and 2 separately (versions 1,2 and 3 respectively). The inner orientation was also done manually at the AC1. The results of the transformation are listed in Table 5. The following can be concluded. The corner fiducials give more precise results than the middle ones, which therefore should not be used. The middle fiducials are also systematically worse in y direction than in x. For each version the results of all 4 images are similar, with the exception of version 2, where there are big differences between left and right images. The left images' fiducials were of poorer quality and also their manual measurements had constantly larger residuals. Another explanation can be the fact that only one CCD camera was used and the digitisation of the left image fiducials occured with a time difference from the other images by transfering the films on the right stage. For all versions in all images the same fiducials appear systematically to have the largest residuals in x or y. Thus, such a digital technique can be also used for checking the quality of the camera calibration. The accuracy of the inner orientation with only the corner fiducials is very high, especially if we disregard the systematic errors that seem to exist for the left image fiducials because of the use of only one camera. The operator measurements with an average RMS of 4.6 μm were less accurate. The camera calibration does not pose any difficulties for matching. The limiting factor to reach the theoretical precision is basically the quality of the digital data.

	RMS (μm)			average |x residuals|	average |y residuals|
	aver.	min	max		
vers. 1	4.9	4.1	5.7	4.2	4.5
vers. 2 right	1.4	1.2	1.5	1.2	1.5
vers. 2 left	4.7	4.1	5.2	4.9	4.3
vers. 2 total	3.0	1.2	5.2	3.0	2.9
vers. 3	4.9	4.3	5.2	3.6	5.8

Table 5. Carriage to image coordinate transformation

Outer orientation

The outer orientation for the two image pairs was done manually. The control points were rather small and imaged very poorly, so larger scale images had to be used for their identification. The results are listed in Table 6. The outer orientation was used as a fixed known quantity in the MGCM algorithm to determine the D.T.M. point heights automatically. In one oriented model the same points were also measured manually, so the outer orientation was the same for both methods. Its accuracy, which is rather low, influences both measuring techniques. An additional negative effect for the MGCM comes from the fact that it uses all four images simultaneously, whereby the orientation was determined separately for each image pair. Thus, if the orientation is weak, the two pairs might not be consistent among each other, the four rays might not intersect exactly in the object space and this

	control points	RMS from control points (m)			stand. dev. of outer orientation (m , grad)					
		X	Y	Z	Xo	Yo	Zo	omega	phi	kappa
Model 1	9	0.251	0.167	0.158	0.84	0.84	0.52	0.008	0.008	0.004
Model 2	12	0.171	0.199	0.245	0.83	0.59	0.41	0.006	0.009	0.004

Table 6. Outer orientation results (the standard deviations are the average of both images)

would degrade the accuracy of the automatic matching.

Automatic D.T.M. measurements

Computation of approximate values

The method of image pyramids was used for derivation of approximate values. The pyramid levels required depend on the estimated maximum parallaxes and they were determined empirically to be 2 to 3. The decimation step was two and the generating kernel was a 13 x 13 optimal approximation of the ideal kernel which is an infinite sequence of sin terms. The kernel is symmetric and was modified to fulfil normalisation and equal contribution of each pixel to the next level, but it is not unimodal. It is computationally expensive but its results are excellent, as there is a minimum loss of information at each higher level. The fact, that it can be separated into two equal 1D kernels with half band property (every coefficient with distance a multiple of the decimation step from the central coefficient is zero, so there are only 49 nonzero coefficients out of the 169), can be exploited for faster implementation. At the borders of the image wrap around was used. Actually this is not justified and creates the effects shown in Fig. 16 (levels 1-3 of square 4), but it did not influence the matching since the border regions were not used. For more details on image pyramids see Burt, 1983, Meer et al, 1987.

The original image is termed level 0. First, one of the four images was selected to serve as template. The selection criteria were the amount of noise and the quality of the contrast. Then a second image was chosen. If we think the 2 x 2 block as a 2 x 2 matrix, either the vertical or the horizontal neighbour was chosen, since the diagonal would have parallaxes in both x and y directions. We built the image pyramid only for those two images and used the MGCM in all levels from top up to and including level 1 to derive X, Y, Z approximations. At level 0, image and pixel coordinates in all four images could be computed by projection of the approximate X, Y, Z in the image space.

In each of the levels from top up to level 1 the matching points were defined on a regular grid with equal x, y spacing. The use of only two shifts instead of an affine transformation generally suffices, since the accuracy sought for at this stage is not the ultimate one and the shears and scales, due to the small image scale, appear to be small. Only for square 3 an x-scale and shear was used to account for the geometric distortions, which are especially visible at the lower right of the image. In all levels of the pyramid radiometric corrections were applied, so that the grey level average and standard deviation of each patch becomes equal to those of the template. At each successive lower level the grid was gradually densified to model better the surface and the patch size was reduced to get more accurate height approximations. At each level each point was subject to a quality test. The test is meant to detect blunders which are excluded from the interpolation of the approximate values for the next level. An example of the results of level 1 of square 2 is shown in Fig. 17. The larger points represent the points which are detected as blunders. The size of the parallaxes that were recovered by this technique are listed in Table 8. At the final level instead of a regular grid, irregularly distributed points, selected by an interest operator, were used. The method of using a regular grid up to level 1 required the determination of much more points than actually used in level 0. The points selected by the interest operator could have been used directly in all upper levels to derive the approximate values. This approach was not followed since for this test we wanted to determine all selected points and some of them might have been rejected in one of the upper levels because of the quality test.

Feature based selection of D.T.M. points

The decision to select the D.T.M. points by an interest operator was made, based on the following thoughts. Enforcement to match on a regular grid blindly, independently of the image texture and contrast, leads to an increase of false results in the D.T.M. data. These errors will either be not discovered or will require an enormous amount of time for manual editing and corrections. One might raise the question of what happens with the distribution and the density of the data, factors that influence a lot the accuracy of the end product. As far as the distribution is concerned, there is no evidence that the terrain heights vary in regular spacings. The density should be a function of the terrain variation. This was proposed first by Makarovic and is implemented today in most commercially available D.T.M. programs. This procedure can be followed with feature based selected points too. The density of the points can be increased by decreasing the dimensions of the window to thin out clusters of selected points and/or by lowering the threshold values. If it is wished, the image can be divided into cells centered at the nodes of a hypothetical regular grid and the best point within each cell can be selected, in the worst case even independently of thresholds. This would definitely lead to a choice of better quality points than those selected rigidly on a regular grid. As far as the D.T.M. data structure is concerned, there exist programs which can handle irregularly distributed data directly. Otherwise an interpolation of a regular grid can be made. On the other hand the interest operator based selection offers the following advantage. Very often the topographic edges coincide with the photometric ones, so a great deal of the densification points required and of the characteristic lines are already determined during this stage, which leads to a further reduction of the manual editing time. Small scale images are particularly suitable for such a procedure because the "allowable" search space for a point extends over a larger area and photometric edges are more probable to be encountered.

So, the strategy followed was selection of points with good contrast in the template, interpolation of their approximate values from the regular grid of level 1 and use of all images simultaneously to determine X, Y, Z. The points were extracted by an edge detector. The MGCM algorithm forces the patches to be transformed along the epipolar lines, so the edges should be vertical to those lines. Their direction for the given geometric configuration was approximately one horizontal, one vertical and one diagonal (see Fig. 18). The points were selected, if the grey level gradients were larger than a threshold in any of these three directions. The mask size was 3 in width and 5-7 in length, where length is the dimension along the edge. The longer length was used to exclude detection of edges due to noise, which might have steep gradients but their length is rather small. The gradient threshold was chosen between 4.5 and 5.5. A relative low threshold was chosen, so that points that were characteristic for the terrain but not of particularly good texture would be selected. This was necessary because, for a reliable test, many points were aimed to within an area than was too small. With a realistic point density one could search in neighbouring digital images for such characteristic points without sacrificing the quality of the detected points. To avoid clusters of points, only the best point within a predefined window of size 21-29 pixels was chosen. Half the dimension of the window is the minimum distance between two neighbouring points, so it is related to the grid spacing of a regular grid. The points of the template image of each square selected by the interest operator are shown in Fig. 19-22. The image scale of each square and the range in object space that is spanned by the selected points are listed in Table 7.

	max ΔX (m)	max ΔY (m)	max ΔZ (m)	scale factor
square 1	167	179	57	40800
square 2	173	172	49	38600
square 3	208	148	61	45800
square 4	174	147	47	45000

Table 7. Planimetric and height differences of feature based selected points and image scale

Fine matching using the MGCM algorithm

At these points the approximations of level 1 were improved by an intermediate matching

with a typical patch size of 31^2 pixels, which was between the effective patch size of approx. 61^2 of the level 1 and the patch size used at the final run (typically 13^2). Only two shifts were used at the final run. The use of all affine parameters for so small patches would cause instabilities and is not necessary for a patch size corresponding to 5.5 m on the ground with an average distance of 6,500 m from the sensor. Fig. 19-22 show the final results in the template image. The crosses signalise the detected blunders. It should be mentioned that the quality test is applied for each ray of each point separately. It means that some of the rays might be rejected but others not, as for example in case of occlusions, which "appear" only in some of the images due to different perspective view. In this test a point was accepted only if at least three rays were correct. The error detection is of great importance for the derivation of reliable approximate values and the accuracy of the final results. Our aim was to detect only gross errors. For this purpose 8 criteria were used with thresholds derived from the neighbourhood of each point. Most of the rejected points were blunders. There were few correct points rejected and some blunders left in the data, though. A summary of the results of the final run are listed in Table 8.

	stand. dev. x-shift (pel)	stand. dev. y-shift (pel)	correl. coeffi.	So	number of iterations	maximum |x parallax| (pel)	maximum |y parallax| (pel)
square 1	0.097	0.072	0.87	12.2	7.2	63.5	33.0
square 2	0.087	0.059	0.82	7.0	6.6	66.5	42.5
square 3	0.066	0.048	0.89	9.2	6.9	49.4	36.6
square 4	0.062	0.045	0.88	9.4	7.0	44.1	18.7
average	0.078	0.056	0.87	9.5	6.9		

Table 8. Results of the fine matching at the lowest pyramid level

As it can be seen in Table 9, the theoretical precision of the object coordinates is very high. From a point of view it is too optimistic, but it must be understood that these values are functionally related, due to the geometric constraints, to the precision of the x,y shifts, which for ideal targets can be lower than 0.01 of a pixel. This shows the great accuracy potential of this algorithm, which can only be achieved, if the limiting errors are accounted for. The theoretical precision was also determined by a bundle adjustment without additional parameters. For well defined points with four rays at the same image positions as the squares, image scale 1:42,500, the standard deviations were as listed in Table 9. These are computed , assuming a pointing accuracy, for natural points and the specific image material, of σ_o equal to 7.5 μm.

		standard deviation X		standard deviation Y		standard deviation Z	
		(cm)	(per mil hg)	(cm)	(per mil hg)	(cm)	(per mil hg)
matching	square 1	1.5	0.002	0.8	0.001	6.4	0.010
	square 2	3.1	0.005	1.4	0.002	6.1	0.010
	square 3	5.0	0.007	0.7	0.001	7.8	0.011
	square 4	4.0	0.006	1.8	0.003	6.5	0.009
	average	3.4	0.005	1.1	0.002	6.7	0.010
bundle adjustment		23.2	0.036	17.6	0.027	48.0	0.074

Table 9. Theoretical precision of object coordinates from matching and bundle adjustment

Manual measurements and accuracy test

To check the accuracy of the matching all points, that were not rejected by the quality test, were measured manually at the AC1. Only one model was used for the measurements. More than one operator measured the points. For the measurements program PMO-single point measurement was used, whereby the X, Y, Z coordinates of the matching were used to drive automatically to the same points, so the starting point for the manual measurements was the X, Y, Z computed by matching and the operator had to set only the correct height at the same

planimetric position. At this stage an error in the AC1 software had as a result, that the image coordinates of the points, that were driven to, were different than those of the matching, although the object coordinates and the outer orientation were the same. The average shift of the image coordinates is listed in Table 10. This shift changed both planimetric and height position of the points. At flat terrain, the planimetric shift would not cause much difference in Z. But quite some points, especially since they were chosen by an interest operator and were lying at hypsometric edges, when they were shifted in planimetry, they had another height. The operator measured the height at the shifted positions and this introduced a systematic difference between manual and automatic measurements.

		Δx (μm)	Δy (μm)
square 1	left	9.5	-24.2
	right	5.2	2.9
square 2	left	24.4	-24.8
	right	10.4	2.2
square 3	left	-16.2	-21.2
	right	-3.9	2.5
square 4	left	-3.2	5.8
	right	-14.2	-12.4

Table 10. Image coordinate errors introduced by the AC1

For the comparison between the two sets of measurements four versions were used. Version 1 without bias and without blunders, version 2 without bias and with blunders, version 3 with bias and without blunders,version 4 with bias and with blunders. Blunders were considered observations beyond the range average - 3 * standard deviation, average + 3 * standard deviation. After each blunder was removed, the average and standard deviation were recomputed to check for new blunders. This was iterated until no blunder was found. The bias (mean difference between the two type of measurements) should be subtracted, in the opinion of the author, since each operator has always a tendency to measure higher or lower. To verify that, square 2 was measured by three operators and by two of them twice. All the pairwise differences among these sets of measurements and the matching results were computed for all four versions (total of 60 pairs). The average of the differences with sign and the average of their absolute values were very close, i.e. the differences were one-directional. Each operator seemed to measure at a more or less constant level. The difference of the levels was considerable, partly because of the poor image quality. The set of measurements that was further than the other ones belonged to an operator. The remaining three sets had a relative good closure among them. The set of measurements that had the least blunders when compared with the other ones belonged to matching. The operator's measurements had at some points jumps, while matching measures in a more systematic way, it has an internal consistency. Furthermore the model was oriented by another person and it was impossible to adjust for the personal index of each operator by reorientation of the model, because of the poor quality of the control points. Another source of bias comes from the way the operator treated the different terrain type. In cases of points near trees and buildings' edges, the operator often tended to measure the height by using the nearby ground, even if the measuring mark was sitting correctly on the surface. This is the case especially in square 3 and less in square 4. From a D.T.M. point of view the operator handled correctly but it is to be reminded that the objective was the test of the measurement accuracy, independent of the use of the object. So, the results that more realistically reflect the accuracy of the automatic matching are those of version 2. Blunders included in this version were undetected and should be taken into account when computing the accuracy measures. Bias should not be considered for the reasons explained above.

The analytical results are listed in Table 11. The results of the versions without bias are

	RMS		maximum | ΔZ |		average | ΔZ |		stand. dev. | ΔZ |	
	(m)	(per mil hg)	(m)	(per mil hg)	(m)	(per mil hg)	(m)	(per mil hg)
version 1	1.33	0.20	3.80	0.59	1.04	0.16	0.82	0.13
version 2	1.56	0.23	7.16	1.09	1.12	0.18	1.03	0.16
version 3	1.85	0.28	4.89	0.75	1.50	0.23	1.08	0.17
version 4	2.11	0.33	8.26	1.26	1.58	0.24	1.26	0.20

Table 11. Matching heights versus manual measurements - accuracy measures

similar for all four squares. Even with bias the differences among the images are relatively small. The ratio worst to best result (based on the RMS) is 1.24 for version 1, 1.3 for version 2, 1.96 for version 3 and 1.94 for version 4, although the terrain use and partly also the terrain slope and image scale were different. Apart from the homogeneity, the matching results show also high accuracy, and with moderate to steep terrain slope indeed. Without bias the average RMS is 0.20-0.23 ‰ hg, the average of absolute errors 0.16-0.18 ‰ hg and the maximum absolute error 0.59-1.09 ‰ hg. It must be considered here that the measurement errors listed above are the result of the combination of the errors of all stages starting from calibration. The poor quality of the digital data and the weak outer orientation contributed the most. The accuracy was lowered also by the fact that many points on trees, houses etc. were measured. These difficult cases, which are anyway useless for a D.T.M., can be excluded from matching with manual editing. The conclusion is that the automatic measurements are at the same accuracy level as the manual ones. The RMS listed in Table 11 are within the fluctuations of the operator measurements. That is indicated by the differences between the two measurements that each operator made, as mentioned above, which had an RMS of 0.9 m and 2.4 m. If it is required that the RMS should be 1/3 of the contour interval and the max error one contour interval, then the results are accurate enough for contours with 5-10 m interval.

Blunders

The number and percentage of the blunders are listed in Table 12. They are similar, apart from square 2, where stricter criteria were used. The amount of blunders is much smaller, if we compare it with results from other investigations (Mengxiang, 1987). The reliability is particularly high, if the small patch size is also considered. Factors that influence beneficially these results are:

- The use of multiple images
- The geometric constraints which force the matching positions to be along the epipolar lines and with a given ratio among them
- The use of feature based selected points with generally good contrast
- The approximate values supplied by the image pyramid

	total number of points	detected blunders		undetected blunders	
		number	%	number	%
square 1	161	43	26.7	1	0.6
square 2	185	21	11.4	2	1.1
square 3	221	25	11.3	5	2.3
square 4	246	29	11.8	4	1.6
total	813	118	14.5	12	1.5

Table 12. Detected and undetected blunders

The first two factors make also the detection of blunders much easier. With an improvement of the quality test the few undetected blunders can be eliminated. Almost all blunders (including the automatically detected ones) occured either in regions of relatively poor contrast (since the threshold of the interest operator was intentionally lowered) or near edges where discontinuities, occlusions and geometric distortions occur. The undetected blunders are more in squares 3 and 4, which had many discontinuities, and are partly due to the tendency of the operator to measure the ground height and the introduced planimetric shift by the AC1 software error. It is interesting to note that out of the 12 undetected blunders, 7 had only three rays and four were flagged as candidates for rejection. If in all points four rays were used, as it should actually be the case, then the max absolute error without bias would not exceed 4.7 m with the exception of one point. The influence of the number of rays should be further investigated, but it can already be stated that the theoretical expectations are verified. Although the number of points with three rays are only 15.3% of the total, if the 15 worst points from each square are chosen, their percentage is 40%. If the 15 best are chosen, it is 8.3% and within the 50 best, it is 12%. It is clear that with less rays the measurements are less accurate and the most important are much less reliable.

Conclusions

The main purpose of this test was to develop and test a procedure for automatic derivation of D.T.M. data from aerial images. The procedure was performed digitally in all stages from camera calibration to the analysis of the results with the exception of the outer orientation. The camera calibration was performed digitally more accurately than with manual measurements. The central point was to check the accuracy of the MGCM algorithm as compared to manual measurements. The advantages of using more than two images and geometric constraints were verified. It came out that the algorithm results lie more or less in the accuracy range of the operator measurements. And that, in spite of the fact that the matching results were degraded by the errors in all stages of the procedure and were influenced by external factors. The theoretical precision is much lower though, and improvements in the algorithm should be made to try to reach it. The aspects that mostly influenced the reached accuracy level and thus should be further investigated are:

- the quality of the digitisation, the radiometric and geometric calibration of the sensor
- the influence of a weak outer orientation
- the influence of the number of rays on accuracy and reliability
- the improvement of the quality test to detect errors
- the type of selected points

The reliability and success rate were improved comparing to other investigations. The number of blunders is drastically cut down, but the operator is still necessary to make a final check of the results. The problem of approximate values was solved sufficiently by the image pyramid. Use of selected good contrast points leads to more accurate and reliable results than a regular grid. The algorithm cannot find automatically whether a point is useful for a D.T.M. or not. This problem can be handled today by interactive graphic editing by the operator, before and after the matching. Before, the operator can view the images and define closed polygons of useless areas (forests, lakes, etc.) which can be clipped off. This would reduce time for matching and decrease the editing afterwards. For manual editing the following methods can be used:

- stereo superimposition of the results in an oriented model. Especially good method and permits also simultaneous addition of new points.
- stereoscopic view on a monitor screen.
- stereoscopic view of an orthophoto pair produced with the D.T.M. data and comparison with the D.T.M. heights superimposed on the images.
- creation of contours, 3D views, shaded reliefs.
- sample checking of profiles, points etc. and statistical comparisons with measurements of higher accuracy.

Time consumed by matching was not measured but it was definitely more than what the operator needed. The time aspect is not a problem though, with careful coding and particularly use of dedicated hardware. Today there exist moderately priced coprocessor boards, which can be also programmed in C and Fortran and offer a 20 MFlops performance.

Acknowledgements

The author would like to thank Dipl. Ing. Z. Parsic, Dr. G. Toz and Professor A. Gruen for making the manual measurements.

References

Bethel , J., 1986: The DSR11 Image Correlator. Technical Papers 1986 ACSM-ASPRS Annual Convention, Washington D.C., USA, Vol. 4, pp. 44-49.

Beyer, H., 1987: Some Aspects of the Geometric Calibration of CCD-Cameras. Proceedings Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, pp. 68 - 81.

Burt , P., 1983: The Laplacian Pyramid as a Compact Image Code. IEEE Transactions on Communications. Vol. 31, No 4, pp. 532-540.

Daehler , J., 1987: Problems in Digital Image Acquisition with CCD-Cameras. Proceedings Intercommission Conference on Fast Processing of Photogrammetric Data,

Interlaken, Switzerland, pp. 68 - 81.

Gruen , A., Baltsavias , E., 1987a: High-Precision Image Matching for Digital Terrain Model Generation. Photogrammetria, Vol. 42, No 3, pp.97-112.

Gruen , A. , Baltsavias , E., 1987b: Geometrically Constrained Multiphoto Matching. Proceedings Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, pp. 204-230.

Guelch , E., 1986: Instrumental Realisation and Calibration of Digital Correlation with the Planicomp. Proceedings of the 40th Photogrammetric Week, Stuttgart University, pp. 91-108.

Meer, P., Baugher, E., Rosenfeld, A., 1987: Frequency Domain Analysis and Synthesis of Image Pyramid Generating Kernels. IEEE Transactions on PAMI, Vol. 9, No 4, pp. 512-522.

Mengxiang , Li, 1987: A Comparative Test of Digital Elevation Model Measurement using the Kern DSR-11 Analytical Plotter. Photogrammetric Reports, No 53, Royal Institute of Technology, Department of Photogrammetry, Stockholm, Sweden.

Pertl , A., 1984: Digital Image Correlation with the Analytical Plotter Planicomp C100. International Archives of Photogrammetry and Remote Sensing, Vol. 25/A3b, Rio de Janeiro, Brazil, pp. 874-882.

Schewe , H., Foerstner , W., 1986: The Program PALM for Automatic Line and Surface Measurement Using Image Matching Techniques. International Archives of Photogrammetry and Remote Sensing, Vol. 26-3/2 , Rovaniemi, Finland, pp. 608-622.

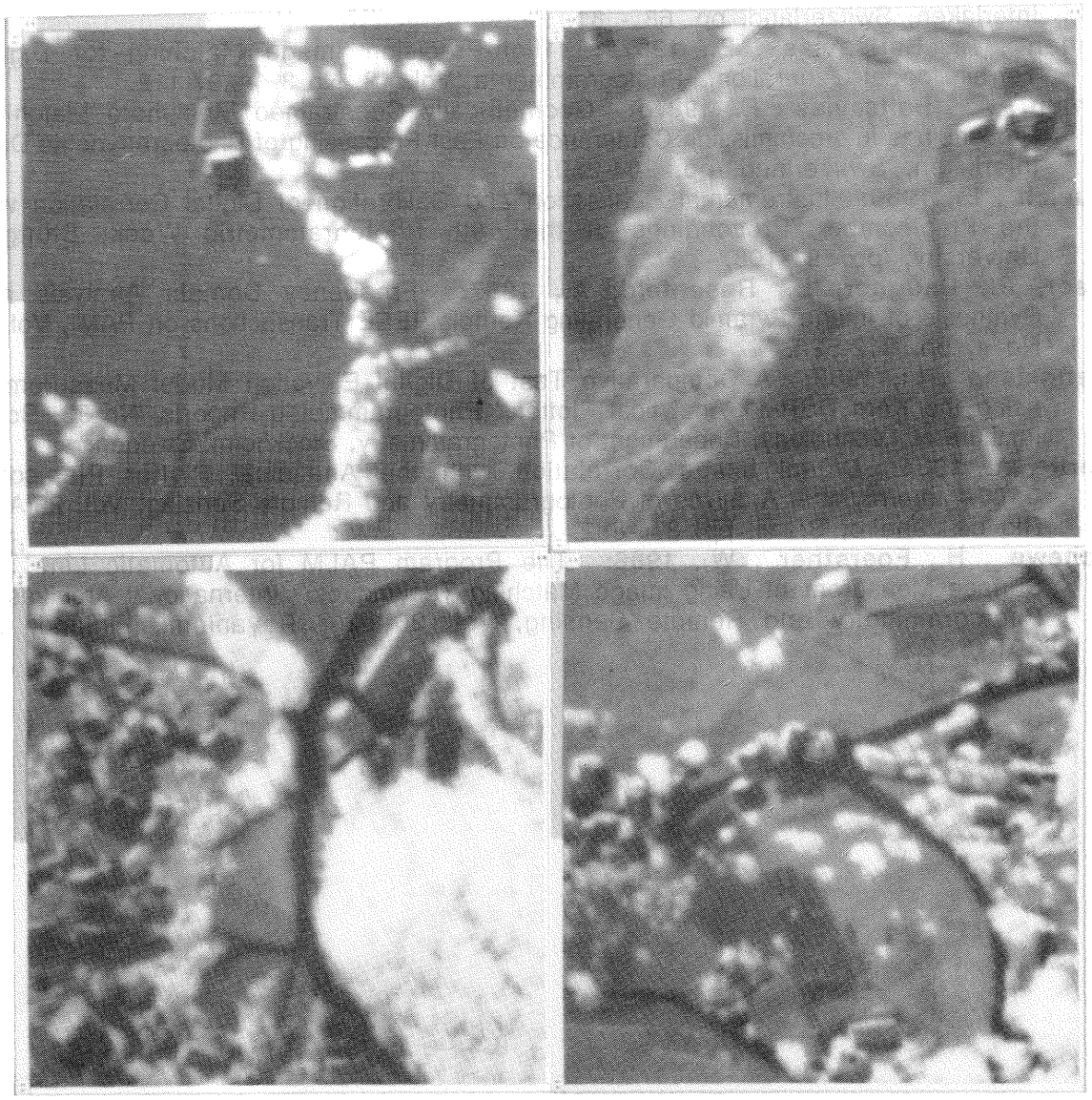


Fig. 1. Starting from top left rowwise, squares 1, 2, 3, 4.

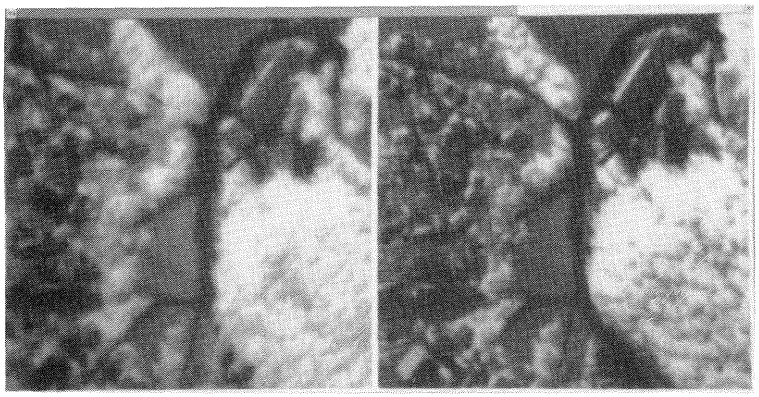


Fig. 2. Radiometric, geometric differences between two images of square 3

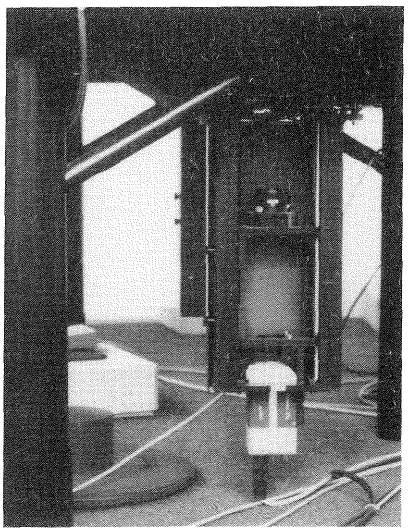


Fig. 3. The HR 600 CCD camera with its lens, mounted below the stages of the WILD AC1

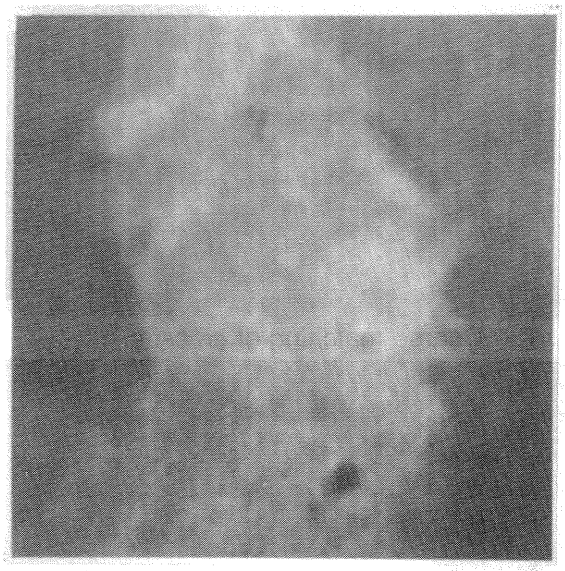


Fig. 4. Original degraded data (2 x magnified)

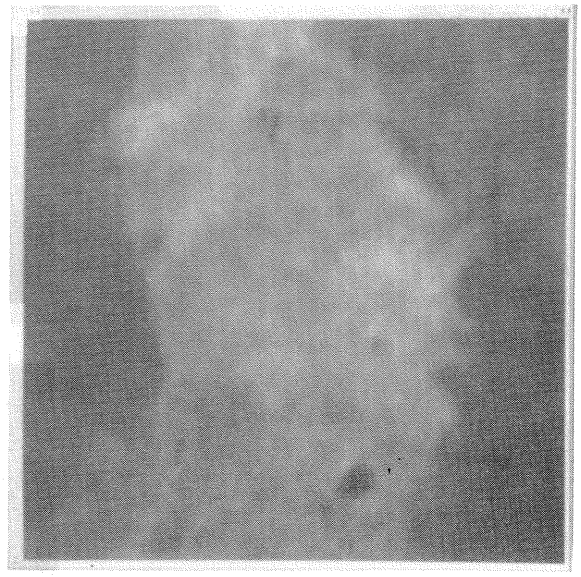


Fig. 5. Filtered to remove horizontal lines

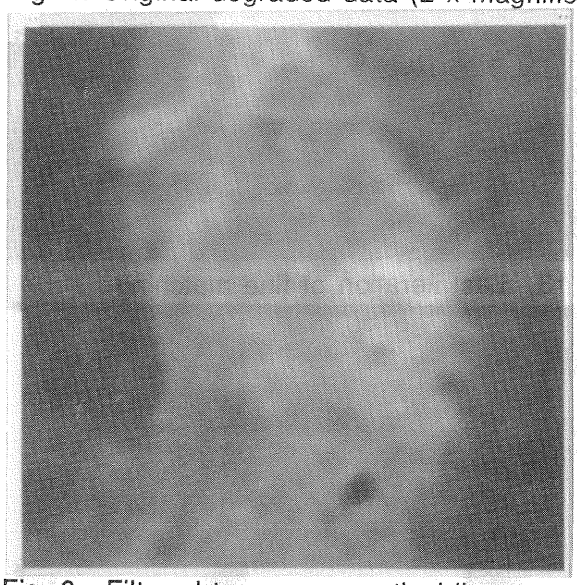


Fig. 6. Filtered to remove vertical lines

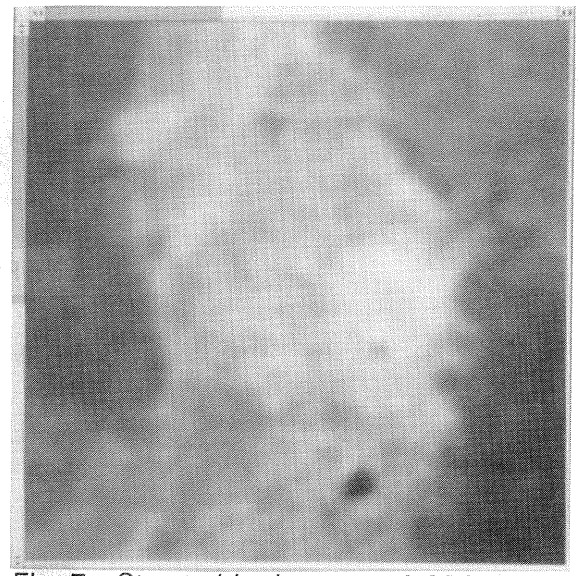


Fig. 7. Strected in the range 0-255.

Fig. 8. Image difference of fig. 4 - 5

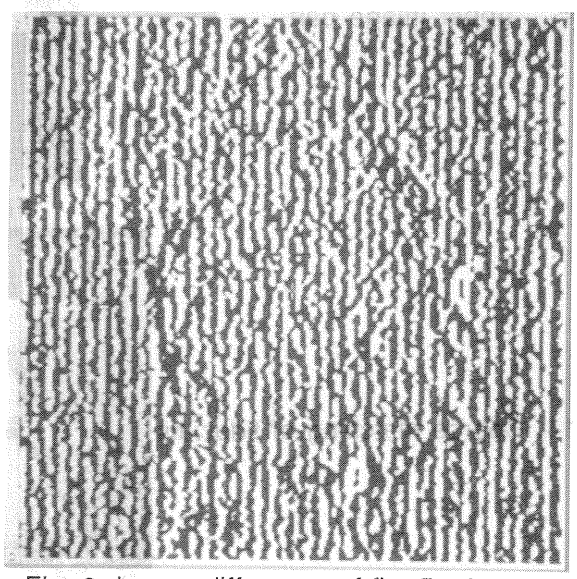


Fig. 9. Image difference of fig. 5 - 6

Fig. 10. Image difference of Fig. 6 - 7

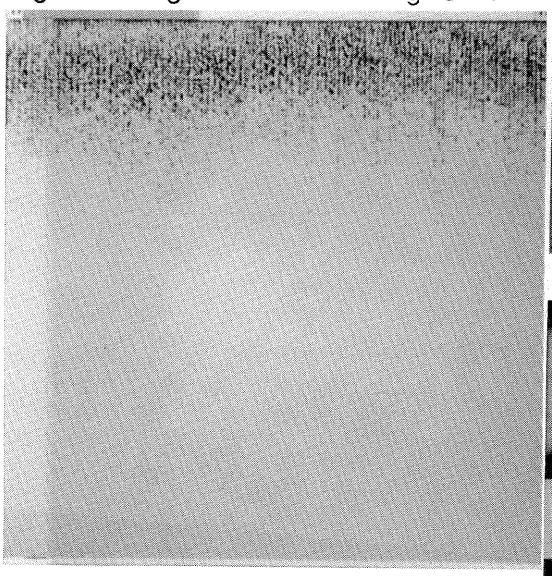


Fig. 11. Filtered sq. 3. More noise on upper image part

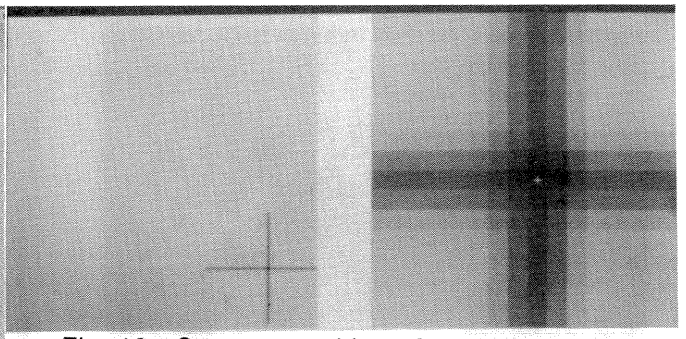


Fig. 12. Coarse matching of crosses

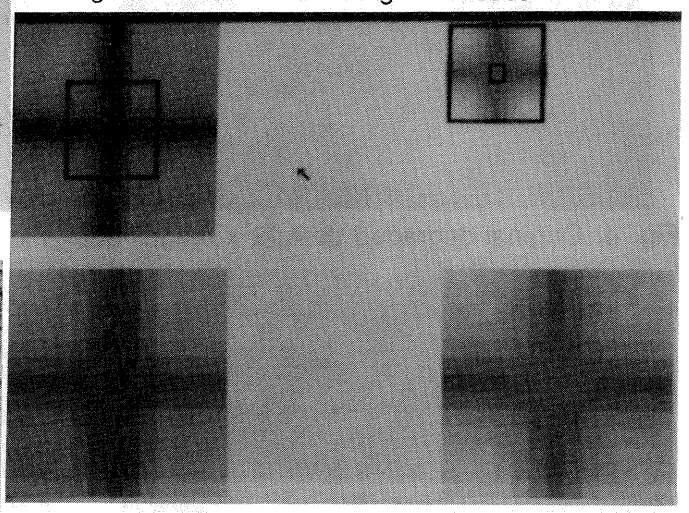


Fig. 13. Last iteration of fine matching

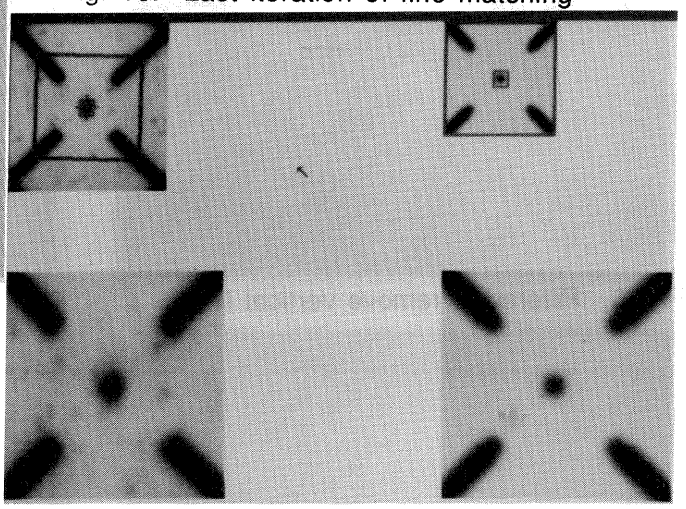


Fig. 15. Last iteration of fine matching

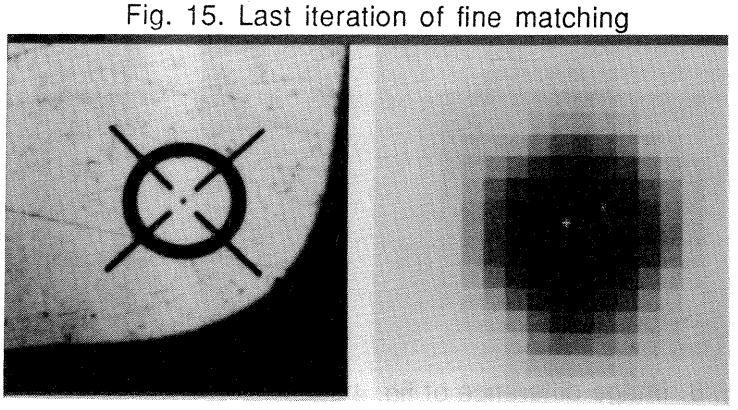


Fig.14. Coarse matching of fiducials

Fig. 16a. Pyramid level 1 of sq. 5 (256^2)

Fig. 16b. Pyramid level 2 (128^2)

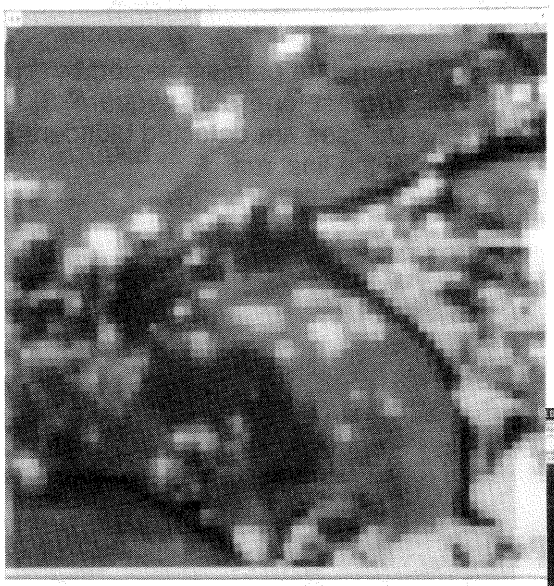
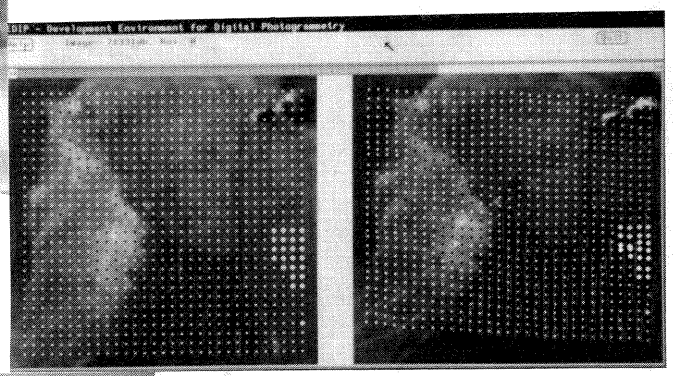


Fig. 16c. Pyramid level 3 = top level (64^2)

Fig.17. Results of level 1 of sq. 3. Points on a regular grid supply the approximations for level 0



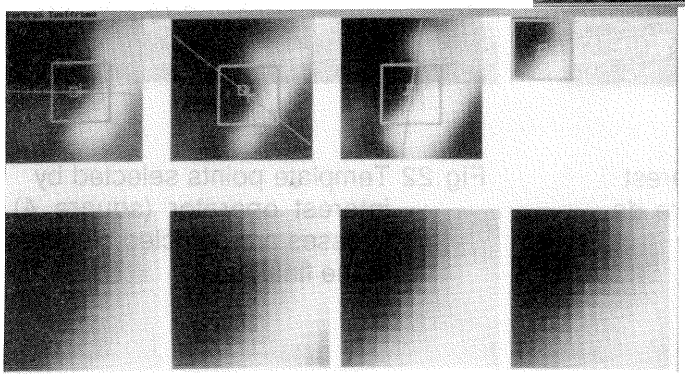


Fig. 18. An example of multiphoto matching at level 0, at a point selected by an edge operator (last iteration). The epipolar lines are also shown

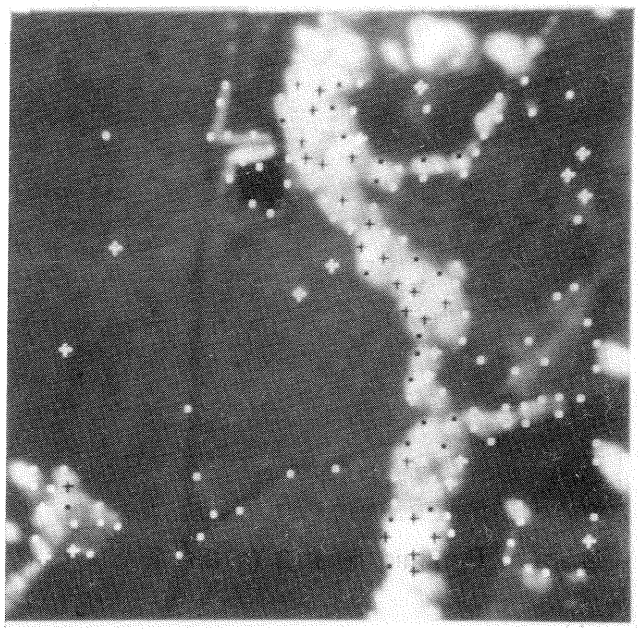


Fig.19. Template points selected by interest operator (square 1). Crosses are detected blunders of the final run

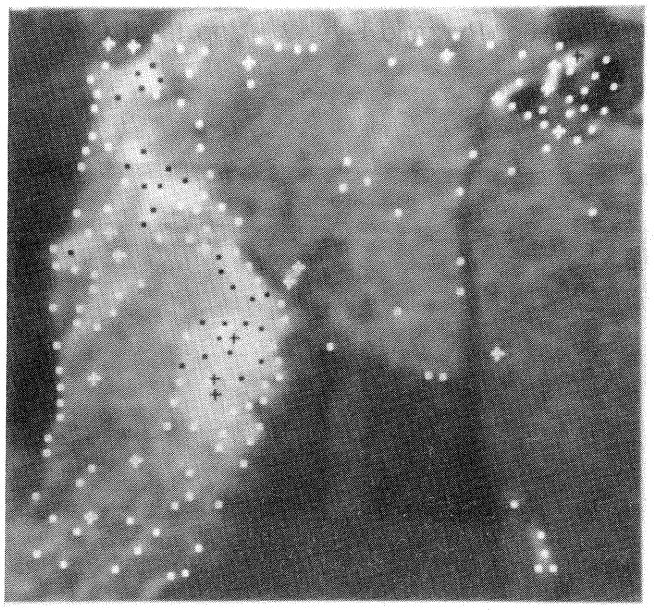


Fig.20. Template points selected by interest operator (square 2). Crosses are detected blunders of the final run

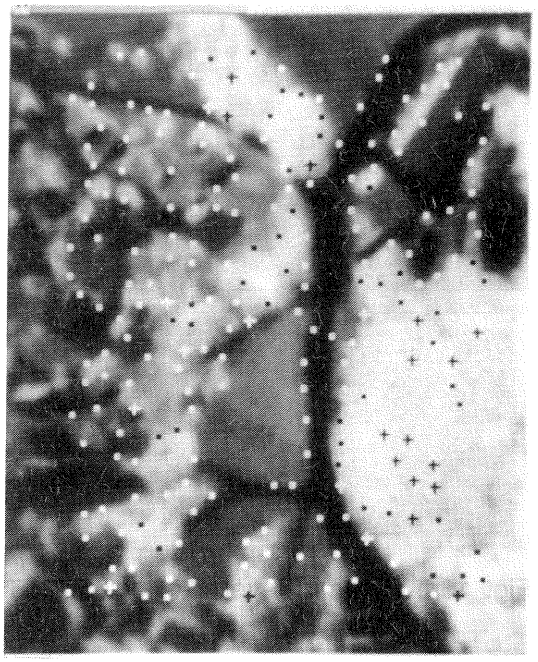


Fig.21. Template points selected by interest operator (square 3). Crosses are detected blunders of the final run

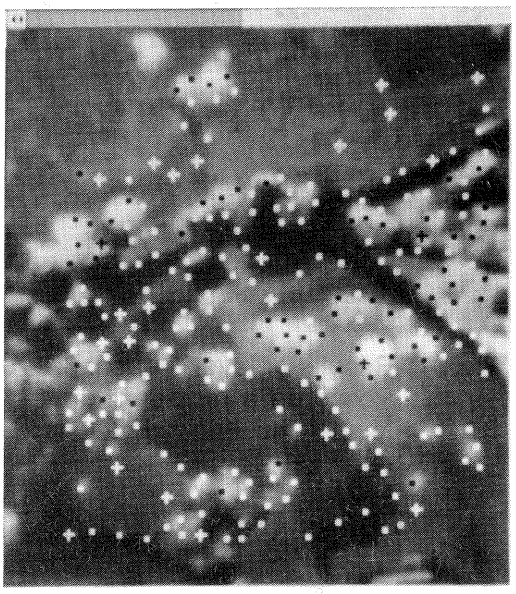


Fig.22 Template points selected by interest operator (square 4). Crosses are detected blunders of the final run