Multi-Point Matching Using the Least-Squares Technique for Evaluation of Three-Dimensional Models

Dan Rosenholm
Department of Photogrammetry, Royal Institute of Technology, S-100 44 Stockholm, Sweden

ABSTRACT: A multi-point matching method using finite elements has been formulated. Parallaxes of points in a grid connected with bilinear elements are computed simultaneously directly from image data using the least-squares technique. Constraints, such as minimizing of the curvature, are implemented as fictitious observations. Methods for inclusion of breaklines and discontinuites, where it is possible to include the position as unknowns, are also discussed. Preliminary tests of the multi-point matching method have been performed.

INTRODUCTION

ORDINARY SINGLE POINT MATCHING METHODS do not take the relation between the points into consideration. In areas with bad texture or large radiometric differences between the images it is often impossible to obtain any result. Such areas could be bridged over by using knowledge of the neighboring points in the matching algorithm. When interpolations of digital terrain models in symmetrical grids from arbitrarily chosen elevation points are performed, this is done by finite element methods (Ebner *et al.*, 1980). Similar methods can be used in digital matching. In the least-square method it is possible to use different geometrical descriptions as additional unknowns. In Figure 1 the different complexity of parallax differences, caused by different surfaces and camera geometries, is shown as surfaces. These surfaces could be connected to the use of geometrical parameters as follows:

- The first surface (Figure 1a), the horizontal, shows the geometrical model of the parallax differences without any additional geometrical parameters used in the least-square matching algorithm. Generally, this is geometrically a surface without any parallax difference within the window used for the matching.
- The next case (Figure 1b) illustrates the geometrical model when affine parameters are used as an approximation of projective parameters. If the relative orientation is known, the geometrical parameters only have to be applied in the epipolar direction. The differences in parallax within the image could, using this model, be expressed as linear functions of the coordinates.
- The multi-point matching method using bilinear finite elements is a discrete approximation of the third geometrical model (Figure 1c) with continuous changes of the parallaxes between the images.
- The fourth surface (Figure 1d) with discontinuties in the first derivative of the parallax is approximately described with a multi-point matching method with breaklines included.
- Inclusion of discontinuities in the multi-point matching method corresponds to the fifth surface (Figure 1e).

This article deals mainly with the third group, continuous changes in the parallax, but other geometric descriptions are discussed.

The least-square matching method has, since it was presented (Förstner, 1982), been experimentally and theoretically investigated in many photogrammetric research departments. From the reported projects some important conclusions can be drawn. It seems as if we have already reached an acceptable precision comparable to manual measurements, but it also seems that we still are not able to obtain results with high reliability and that we are not able to detect the gross errors sufficiently well. I will give some examples from the literature.

Pertl (1985) has reported practical results from his work with the least-square matching method on an analytical plotter. The precision of parallax measurements with this method for relative orientation was reported to be 3.8 μm. The standard error of unit weight for a bundle adjustment with this method was 3.4 μm, which should be compared to 2.8 μm using monocomparator measurements.

In a project concerning digital matching of simulated SPOT images (Rosenholm, 1985), the results pointed towards a root-mean-square error of around one-third of a pixel.

Experiments using least-squares matching for measurements of tie points were reported by Ackermann and Schneider (1986). Accuracies of the tie points after the adjustment of between 3.5 and 7 μm were reached. Only small reliability problems occurred as the process was operator controlled.

Ackermann *et al.* (1986) reported investigations concerning the number of unacceptable matches in different image material. The number of gross errors varied between more than 50 percent in some data sets down to around 5 percent in other.

Grün and Baltsavias (1986) obtained mean errors of between 2 and 3 μm in an investigation using a least-squares algorithm constrained to be one-dimensional. The reliability problem was small as the points were manually chosen and the image material was of good quality.

From a project concerning least-squares matching, in which the work reported in this article was a part (Rosenholm, 1986a and 1986b), root-mean-square errors of about 7 μm in the epipola direction and 5 μm in the y-direction were obtained. The reliability, after automatic detection of gross errors, varied between more than 50 percent gross errors down to some few percent. In the last case, a large number of points without gross errors were rejected because of the criteria for gross error detection used.

These results show that we are able to get parallaxes measured with high accuracy by the least-square matching method in points where we do not have gross errors. It also shows that when the points are not manually chosen, but are measured in regular grids (Rosenholm, 1985; 1986a; 1986b; Ackermann *et al.*, 1986), we get a much too poor reliability.

In this article a method which may solve the reliability problem is formulated. It has similarities to an array algebra matching method independently proposed by Rauhala (1986).

THE LEAST-SQUARES MATCHING ALGORITHM

The basic algorithm, which minimizes the squares of the grey level differences between the two images, can in two dimensions be formulated with the following generalized observation equation:

$$I_m(x_m, y_m) + n(x_m, y_m) = I_t(x_t, y_t) \tag{1}$$

where

$$x_t = f_1(x_s, y_s)$$
$$y_t = f_2(x_s, y_s)$$
$$I_t = f_3(I_s)$$

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 53, No. 6, June 1987, pp. 621–626.

0099-1112/87/5306–621$02.25/0

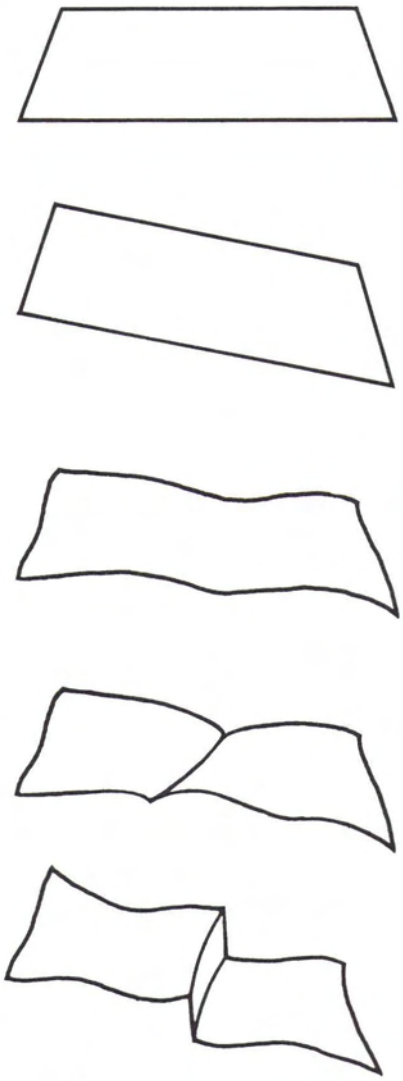

FIG. 1. Different degrees of geometric complexity shown as parallax differences in the epipolar direction. (a) A horizontal surface. (b) A plane surface with an arbitrary slope. (c) A continuous surface with continuous first derivative. (d) A continuous surface with discontinuities in the first derivative. (e) A surface with discontinuities.

and in which x_s and y_s are the coordinates of the pixel in the search window, x_m and y_m are the coordinates of the pixel in the mask window, I_s is the grey level of the pixel in the search window, and $n(x_m, y_m)$ is the additive noise.

The search and mask windows are parts of the two images. The mask window is the template against which the search window will be compared in order to find a parallax. Assuming the x-direction to be equal to the epipolar direction, the matching can be performed one dimensionally in oriented images. The simpliest one-dimensional least-square matching algorithm uses the geometric functions

$$\begin{aligned} x_t &= p_1 + x_s \\ y_t &= y_s \end{aligned} \tag{2}$$

where p_1 is the translation in x which gives the following linear observation equation:

$$\Delta p_1 \, I_x = \Delta I \tag{3}$$

where Δp_1 is the computed translation and

$$I_x = \frac{dI_s}{dx} \approx \frac{dI_m}{dx} \approx (I_m(x+1, y) - I_m(x-1, y)) / 2$$

$$\Delta I = I_m(x, y) - I_t(x, y)$$

This model does not compensate for any geometrical differences between the images. An alternative geometrical model for epipolar matching, compensating for linear geometrical differences in the epipolar direction, is

$$\begin{aligned} x_t &= p_1 + p_2 \, x_s + p_3 \, y_s \\ y_t &= y_s \end{aligned} \tag{4}$$

from which we get the following linear observation equation:

$$\Delta p_1 \, I_x + \Delta p_2 \, I_x \, x + \Delta p_3 \, I_x y = \Delta I. \tag{5}$$

One of the main advantages of the least-square method is the possibility of making a direct estimation of the accuracy of the parallax. The error of the parameters can be computed by error propagation from the standard error of unit weight.

MULTI-POINT MATCHING

A limitation, especially with automatic parallax measurements performed with image processing methods, but also with manual measurements, is the difficulty of making measurements in areas with low signal content. Common matching methods, measuring one point at a time, are incapable of computing a parallax in such an area, while humans have abilities beyond the performance of automatic matching methods of today. What is needed is a method which can bridge areas with low signal content. I will here formulate a method which makes it possible to simultaneously compute the stereo parallaxes in grid points, which are connected with bilinear finite elements (Figure 2).

Using the generalized formulation of the observations above (Equation 1), we can formulate the method by the following geometric functions:

$$\begin{aligned} x_t = x_s + [&p_{ij}(x_{i+1} - x_s)(y_{j+1} - y_s) \\ &+ p_{i+1,j}(x_s - x_i)(y_{j+1} - y_s) \\ &+ p_{i,j+1}(x_{i+1} - x_s)(y_s - y_j) \\ &+ p_{i+1,j+1}(x_s - x_i)(y_s - y_j)] \\ &/ [(x_{i+1} - x_i)(y_{j+1} - y_j)] \end{aligned} \tag{6}$$

for

$$x_i < x_s < x_{i+1}, \; y_j < y_s < y_{j+1},$$

and

$$y_t = y_s$$

where x_i is the x-coordinate of the points in the ith column, y_j is the y coordinate of the points in the jth row, and $p_{i,j}$ is the unknown parallax in the point i,j.

The points are connected by bilinear functions describing the parallax differences in the epipolar direction (the x-direction). We can now formulate the observation equations as

$$b_{i,j}\Delta p_{i,j} + b_{i+1,j}\Delta p_{i+1,j} + b_{i,j+1}\Delta p_{i,j+1} + b_{i+1,j+1}\Delta p_{i+1,j+1} = \Delta I \tag{7}$$

where

$$b_{i,j} = \frac{I_x \, (x_{i+1} - x_s)(y_{j+1} - y_s)}{(x_{i+1} - x_i)(y_{j+1} - y_j)}$$

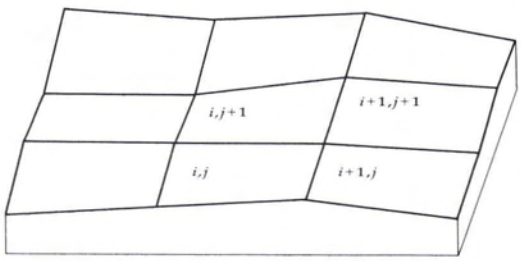

FIG. 2. An example of a bilinear surface in a rectangular grid.

$$b_{i+1,j} = \frac{I_x (x_s - x_i)(y_{j+1} - y_s)}{(x_{i+1} - x_i)(y_{j+1} - y_j)}$$

$$b_{i,j+1} = \frac{I_x (x_{i+1} - x_s)(y_s - y_j)}{(x_{i+1} - x_i)(y_{j+1} - y_j)}$$

$$b_{i+1,j+1} = \frac{I_x (x_s - x_i)(y_s - y_j)}{(x_{i+1} - x_i)(y_{j+1} - y_j)}$$

and

$\Delta p_{i,j} \ldots \Delta p_{i+1,j+1}$ = the corrections to the unknown parallaxes.

When solving this system with the least-squares method, we will get a banded matrix. When radiometric parameters are also added, it will become banded-bordered.

With the multi-point matching method, the parallaxes of the points are connected by bilinear functions. By using additional constraints, formulated as fictitious observations, we are capable of strengthening the connections between the points. Especially when dense grids are used, areas with a low signal content have to be stabilized. The use of the discrete second derivative in the two directions will minimize the curvature; i.e.,

$$\begin{aligned} 2p_{i,j} - p_{i+1,j} - p_{i-1,j} &= 0 \\ 2p_{i,j} - p_{i,j+1} - p_{i,j-1} &= 0 \end{aligned} \tag{8}$$

An example of another possible constraint (which is not used in this investigation) is the first derivative, which can be of importance in the border of the grid. Notice also that the structure of the normal equation system is not affected by either of the two mentioned constraints.

Some questions have to be asked in connection with the practical use of this method. The first question concerns the density of the grid. Dense grids can model the ground surface effectively, but can also be expected to be less stable, while a more sparse grid gives a less detailed model of the object. In the experiments performed, the distance between the grid points was 8 pixels, corresponding to 200 μm in the image. Totally 11 by 11 points were used in the 80 by 80 pixels mask window.

A related question is how to choose a correct *a priori* weight for the constraints. The weights should be large enough to have a stabilizing effect on the system, but still small enough to just have minor effects in areas with good signal content. Methods for making this choice have to be investigated. The weights 100 and 400, on the constraints minimizing the curvature, were used in the experiments discussed in this article.

POSSIBLE EXPANSIONS

A method commonly used in connection with measurements of digital elevation models is local densification of the grid (Makarovic, 1973). Such methods are sufficient when areas with a more fluctuating surface than normal could be identified. The interpolation procedure within such an area can be formulated analogous for multi-point matching as for interpolation in digital elevation models (Figure 3). However, the simple banded bordered structure of the normal equation system has to be changed when new elements are introduced into the interpolation model.

Another possible improvement is the inclusion of breaklines in the original formulation. In, for example, the HIFI program system (Height Interpolation by Finite Elements), which interpolates heights in grid points from arbitrarily distributed height-data points, breaklines have been included in the version using bilinear elements (Ebner *et al.*, 1980). In a formulation of breaklines in multi-point matching (Figure 4), it is possible to have the position of the breaklines included as unknowns, although the existence of a breakline has to be known previously. An example of such a formulation can be found in Rosenholm (1986a).

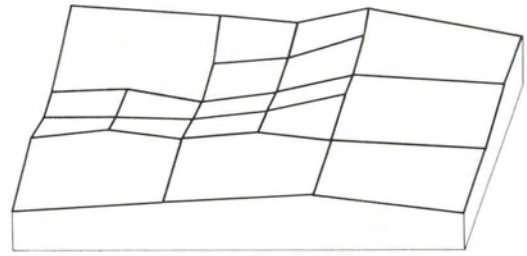

FIG. 3. An example of a locally densified grid.

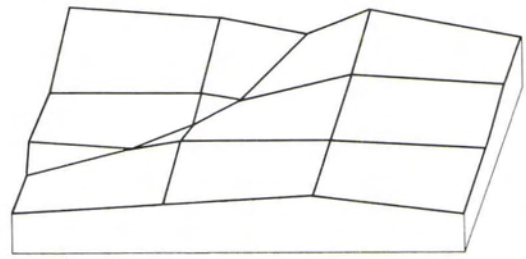

FIG. 4. A Bilinear surface with a breakline.

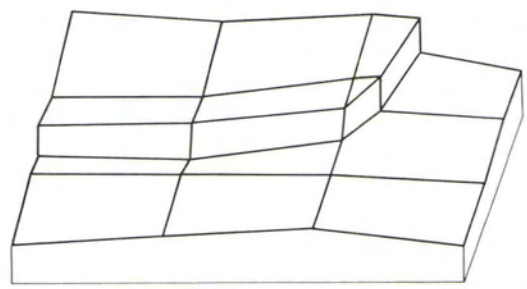

FIG. 5. A bilinear surface with a discontinuity.

The inclusion of a discontinuity into the formulation (Figure 5) is another possible expansion of the multi-point matching method. The interpolation function can be formulated almost identically as to that used for break-lines (Rosenholm, 1986a), the difference being that two different unknown parallaxes are used at each discontinuity point, one on each side of the discontinuity, instead of using the same parallax on both sides.

INVESTIGATION PROCEDURE

The photographs in the investigation were chosen from two examples of practical applications in photogrammetry. They have been chosen to represent two different photogrammetric applications. A large scale aerial photograph was used as well as an example from close-range photogrammetry, a rock tunnel wall (see Figures 6 and 7). The check points are marked with crosses in the image.

The stereo pair was manually oriented in a Wild STK1 comparator in order to get the stereo direction in the area used equal to the x-direction in the comparator and, thus, locally get close to epipolar geometry. Sixty-four transformation marks, surrounding the check points, were measured in the comparator. The 64 check points were situated in an 8 by 8 grid with 2-mm spacing. Each check point was measured six times, of which three were measured with the images interchanged in the eyepieces. Four of the measurements were made with small symmetrical displacements of the measuring mark (50 to 200 μm) from the central point in each of the four main directions.

In the Osiris interactive scanner (Aslund *et al.*, 1980), a 1024- by 1024-pixel image was scanned with a pixel size of 25 by 25

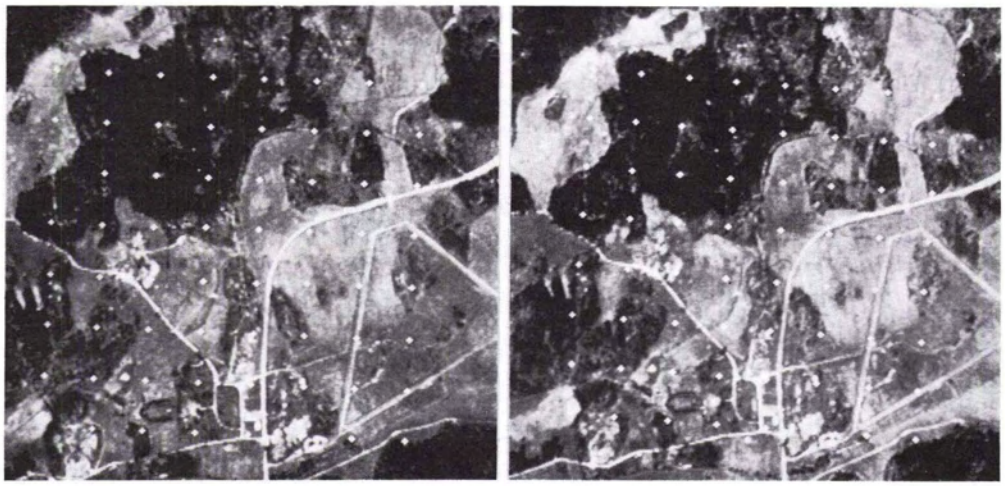

FIG. 6. A stereo pair of aerial photographs at a scale of 1:50,000 from an agricultural and forest area north of Stockholm, with an overlap of about 30 percent and a base-height ratio of approximately one.

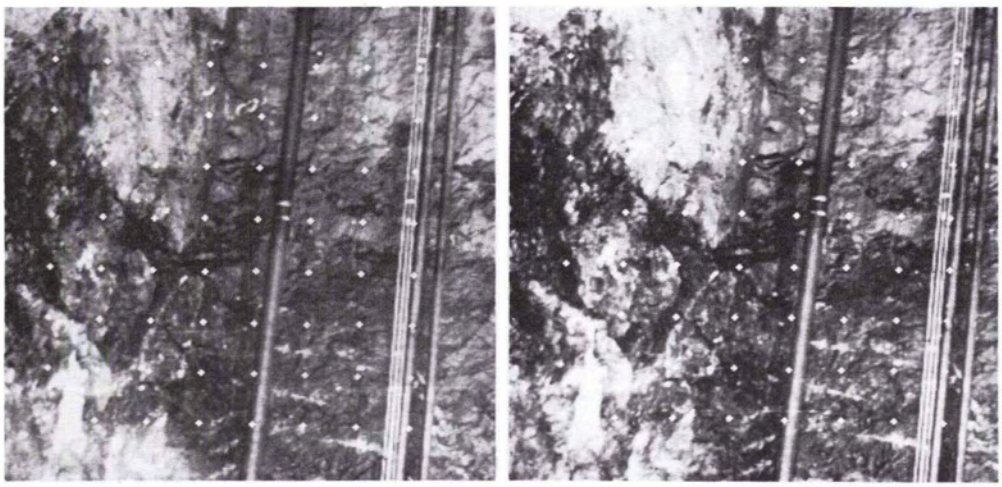

FIG. 7. A stereo pair of a rock tunnel wall exposed with a Zeiss SMK 120 camera.

μm, corresponding to a 25.6-25.6-mm area in the image. The measured comparator coordinates were transformed to the image coordinate system using the 64 transformation marks for an affine transformation.

The matching was performed with the initial value chosen to have a shift of 2 pixels in the epipolar direction (the x-direction). In the single point matching methods, which were used for comparision with the multi-point matching method, one multiplicative radiometric parameter was used as unknown. In the multi-point matching method one additive and one multiplicative radiometric parameter were used as unknowns. The iterations continued until either convergence was obtained (the computed translation was smaller than 1 μm), the maximum number of iterations was reached (16 for the single point methods and 10 for the multi-point method), or the translation exceeded 3 pixels. The results were printed individually for each matching with information about estimated errors, deviation between the matching and the manual measurement, number of iterations, size, standard deviation of computed parameters, etc.

RESULTS

The method presented for multi-point matching has only briefly been experimentally investigated. The experimental results will here be compared with the two epipolar matching methods shown above, with and without additional parameters for compensation of perspective differences between the images.

In Table 1 the precision is shown in μm. It has been computed as a root-mean-square deviation between manual measurements of check points and matchings. Points without gross errors have been chosen manually. When the methods have been compared, the same check points have been used in the different methods.

In Table 2 the reliability is shown in percent. The term reliability is used as a measure of the lack of gross errors. The

TABLE 1. ROOT-MEAN-SQUARE DEVIATIONS BETWEEN MATCHINGS AND MANUAL PARALLAX MEASUREMENTS (IN μM). THE LEAST-SQUARE MATCHING IS PERFORMED IN ONE DIRECTION (CLOSE TO THE EPIPOLAR DIRECTION) WITHOUT GEOMETRIC PARAMETERS (NO GEOM.), WITH LINEAR PERSPECTIVE GEOMETRIC PARAMETERS (PERSP.), AND WITH THE MULTI-POINT MATCHING METHOD WITH THE WEIGHTS 400 AND 100 ON THE CONSTRAINT MINIMIZING THE CURVATURE (MP400 AND MP100, RESPECTIVELY). THE WINDOW SIZE IS 20 BY 20 PIXELS FOR THE TWO SINGLE-POINT MATCHING METHODS

Data Set	Number of points	Method			
		No Geom.	Persp.	MP400	MP100
Rock	44	6.7	7.0	7.4	9.6
High1	12	10.6	9.2	10.9	—

TABLE 2. RELIABILITY IN % FOR THE SAME METHODS AS IN TABLE 1.

Data Set	Number of points	Method			
		No Geom.	Persp.	MP400	MP100
Rock	46	83%	89%	85%	88%
High1	23	52%	64%	70%	—

number of check points with deviations below the three thresholds 10, 15, and 20 μm were averaged to get the measure of reliability. As only ten iterations were allowed in the multi-point matching method while 16 iterations were allowed in the single point methods, only check points which have not reached the maximum number of iterations in any of the matchings with the compared methods are used.

DISCUSSION

The multi-point matching method is intended for stereo parallax measurements in absolutely oriented, or at least relatively oriented, stereo images. Only the parallax in the epipolar direction is used as unknown. In this investigation a square grid in the mask window was used with 8 pixels spacing between the grid points. Two linear radiometric parameters together with 11 by 11 grid points resulted in altogether 123 unknowns. For the computation of the parallax of the points in the border of the grid, only the area inside the border points was used, making the matching window size 80 by 80 pixels. A condition minimizing the curvature was also used, formulated as a fictitious observation of the second derivative of the measured parallaxes, using three nearby points in each observation equation which created totally 220 fictitious observations.

How is the information used by the multi-point matching method? Each grid point (except those in the borders) uses data from an area of 16 by 16 pixels for the computation of each individual parallax. The weight for each pixel will decrease linearly with the distance from the grid point in the x- and y-directions. However, all points will be connected to each other, through the bilinear surfaces and strengthened by the fictitious observations minimizing the curvature. In practice, it is difficult to make a direct comparision between multi-point and single point matching methods concerning how the algorithms use the image data.

What conclusions can be drawn, regarding precision, from Table 1? The multi-point method did not increase the precision. Instead, a small decrease could be noticed. However, we know that the choice of the spacing between the grid points was done without previous knowledge about the method. Neither was the choice of the weight for the fictitious observations of the second derivatives optimized. The change of the weight from 100 to 400 had a great effect in the rock wall photographs. The influence area for each parallax in the multi-point matching window is probably too small. The spacing between the grid points, or alternatively the weight of the fictitious observation minimizing the curvature, was too small. The precision result of the multi-point matching method probably corresponds to that which would have been obtained with small window sizes, lower than 20 by 20 pixels, with single point methods. The experience I have had with single point methods shows that the optimum window size concerning precision is around 30 by 30 pixels or larger with this particular image material (Rosenholm, 1986a; 1986b).

The multi-point matching method gives in the Table 2 a clear improvement of the reliability in the aerial photographs but not in the rock tunnel photographs. We should ask ourselves why.

The answer is based on knowledge about the image material. The photographs of the rock wall are of such a good quality that already with single point methods we get a high reliability. The large window size (80 by 80 pixels) used by the multi-point matching method cannot improve the reliability further. Where there are extreme geometrical shapes (tubes outside the wall) it can even be a disadvantage. The aerial photographs, on the other hand, have no discontinuities or other such extreme geometric shapes. The image quality is low and there are large radiometric differences between the images.

The result indicates that multi-point methods may have advantages concerning the reliability. If the radiometric differences are sufficiently modeled and if the multi-point interpolation method is able to model the surface, then large windows should be a pure advantage, in terms of reliability. Extremely large windows—I am thinking of 1000 by 1000 pixels and larger—may then be possible to use.

In the multi-point method the geometry can be well modeled by points with unknown parallaxes; in some cases breaklines or discontinuities will be needed (such as the tubes outside the rock wall). Concerning the radiometric corrections, further research has to be undertaken in order to model the radiometric differences between extremely large windows.

CONCLUSIONS

The multi-point matching methods have a potential for being superior to single point matching methods insofar as reliability is concerned. The multi-point method has to be further developed in order to make it operative.

Putting together the experimental result from the investigation and the discussion, we are able to form the following two possible hypothesis.

The first hypothesis is that precision, with regard to the influence of the algorithm, is a function of the grid spacing and the strength of the additional constraints used. The change of precision resulting from changing the grid spacing corresponds to the precision changes obtained when the window size is changed in single-point matching methods. However, precision seems to a large extent to be a function of image quality, and can only to a limited extent be improved by changes in the algorithms.

The second hypothesis is that reliability is related to the total window size. As long as the radiometric and the geometric differences between the images are well modeled, an increasing window size will correspond to a reliability increase. With regard to reliability, we are able to make large improvements at the algorithmic level.

THE FUTURE

The most important investigations and improvements of the multi-point matching method are

- investigation of the relation between grid spacing, strength of additional constraints, and the obtained accuracy, in order to find the optimal parameter choice and/or methods to find the optimal parameter choice in different image materials;
- investigation of methods for computation of breaklines and discontinuities (Rosenholm, 1986a);
- investigation of the possibility of using matching results, especially computed breaklines and discontinuities, as a part of object identification;
- investigation of the use of data snooping for exclusion of pixels (observations) in connection with the multi-point matching method; and
- development of suitable reflection models and methods for radiometric compensation of the differences between the images
- development of multi-point methods for surfaces with continuous first derivatives (such as bicubic polynomials)

ACKNOWLEDGMENTS

This investigation was made possible by financial support from the Swedish Board for Technical Development. I would like to thank those of the Department of Photogrammetry at the Royal Institute of Technology in Stockholm (KTH), who in one way or another have made contributions to this project. I especially want to thank Prof. Torlegård for his concern as su-

pervisor. I also want to thank the Physics IV Department of KTH Prof. Nils Aslund and co-workers for making their equipment, especially the OSIRIS-digitizer, available to me.

REFERENCES

Ackerman, F., and W. Schneider, 1986. High Precision Aerial Triangulation with Point Transfer by Digital Image Correlation. *International Archives of Photogrammetry*, Vol. XXVI, Comm. III, pp. 18–27. Rovaniemi.

Ackerman, F., W. Schneider, and G. Vosselman, 1986. Empirical Investigation Into the Precision of Digital Correlation. *International Archives of Photogrammetry*, Vol. XXVI, Comm. III, 21 p.

Ebner, H., B. Hofmann-Wellenhof, P. Reiss, and F. Steidler, 1980. HIFI A Minicomputer Program Package for Height Interpolation by Finite Elements. *International Archives of Photogrammetry*, Vol. XXIII, Comm. IV, pp. 202–215.

Förstner, W., 1982. On the Geometric Precision of Digital Correlation. *International Archives of Photogrammetry*, Vol. XXIV, Comm. III, pp. 176–189.

Grün, A., and E. Baltsavias, 1986. High Precision Image Matching for Digital Terrain Model Generation. *International Archives of Photogrammetry*, Vol. XXVI, Comm. III, pp. 284–296.

Makarovic, B., 1973. Progressive Sampling for Digital Terrain Models. *ITC-Journal*, pp. 397–416.

Pertl, A., 1985. Digital Image Correlation with an Analytical Plotter. *Photogrammetria*, 40, pp. 9–19.

Rauhala, U., 1986. Compiler Positioning of Array Algebra Technology. *International Archives of Photogrammetry*, Vol. XXVI, Comm. III, 26 p.

Rosenholm, D., 1985. Digital Matching of Simulated SPOT-images. *Fotogrammetriska Meddelanden* 2:50. Department of Photogrammetry, Royal Institute of Technology, Stocholm.

———, 1986a. Accuracy Improvement in Digital Matching. *Fotogrammetriska Meddelanden* 2:52. Department of Photogrammetry, Royal Institute of Technology, Stocholm.

———, 1986b. Accuracy Improvement of Digital Matching for Evaluation of Digital Terrain Models. *International Archives of Photogrammetry*, Vol. XXVI, Comm. III, pp. 573–587.

Aslund *et al.*, 1980. Isis and Osiris, Some Experiences of a Software Package and an Image Scanner Designed for Interactive Image Analysis. *Proceedings of the First Scandinavian Conference on Image Analysis*, Linköping, pp. 157–162.

(Received 29 September 1986; accepted 26 November 1986)