# INTEGRATION OF DIGITAL IMAGE MATCHING AND OBJECT SURFACE RECONSTRUCTION

Heinrich Ebner
Chair of Photogrammetry
Technical University of Munich
Arcistr. 21, D 8000 Munich 2
Federal Republic of Germany

Christian Heipke
Industrieanlagen - Betriebsgesellschaft (IABG)
Einsteinstr. 20, D 8012 Ottobrunn
and
Chair of Photogrammetry
Technical University of Munich
Arcistr. 21, D 8000 Munich 2
Federal Republic of Germany

Commission III

**Abstract**

A new approach is presented, integrating digital image matching, point determination and object surface reconstruction. The unknown quantities (the geometric and radiometric parameters for the approximation of the object surface and the orientation parameters of the images) are directly estimated from the given informations (the pixel density values and the control data) by a simultaneous least squares adjustment. Any desired number of images, scanned in various spectral bands can be processed at the same time.
A practical example demonstrates the efficency of the new approach.

## 0. Preface

The work for this paper was carried out within a cooperation between the Chair of Photogrammetry, Technical University Munich and the Departement of Terrain Data Processing, IABG (Industrieanlagen-Betriebsgesellschaft), Ottobrunn. It is supported by the Ministery of Research and Technology of the Federal Republic of Germany.
The IABG is a company providing technical and scientific services in the high-tech domain and employs more than two thousand people.
In the Departement of Terrain Data Processing, twenty scientists work on raster orientated methods of terrain data processing for data acquisition, storage in raster data bases, data manipulation and display.
Hardware in the form of a vector computer Siemens VP 200 (64 MB main memory, 500 MFlops), various mini computers and an image processing system is available.
Most projects require up-to-date height information of the earth's surface. Digital image matching techniques have the potential of modernizing, accelerating, and facilitating the task of data acquisition for deriving up-to-date, accurate Digital Terrain Models (DTMs) considerably.

## 1. Introduction

A requirement for photogrammetric object surface reconstruction is the identification of homologous points in different images. A human operator uses the stereo effect to find the position of these points at maximum accuracy.
Digital correlation allows for automatic determination of the homologous points. The image information has to be present in digital form, i. e. in the form of density value matrices. Correlation is commonly performed using one of the following methods:

- maximizing the correlation coefficient (referred to hereafter as digital image correlation) or
- minimizing the sum of the squared density value differences (referred to as digital image matching).

In the first case, a pattern window consisting of pixels of the first image is shifted pixel per pixel over a larger search window consisting of pixels of the second image. At each position the correlation coefficient of all common pixels of both windows is calculated. The maximum value of the coefficient corresponds to the best fit of pattern and search window.
In the second case a transformation of the pattern window onto the search window using geometric and radiometric parameters is formulated. These unknowns are computed minimizing the sum of the squared differences between the density values of the transformed pattern and search window. This method has a very high accuracy potential. Using a pixel size of approximately $20*20\ \mu m^2$ the image coordinates can be determined within $1 - 2\ \mu m$ (ACKERMANN, 1984; FÖRSTNER, 1984, 1986).
This approach can be generalized combining various pattern and search windows to form a grid. The geometric parameters

are the unknown shifts in the pattern grid points. They are determined simultaneously in a least squares adjustment (ROSENHOLM, 1986).
By connecting the density values with the object coordinates and the orientation parameters of the images it becomes, in principle, possible to compute any number of object points and orientation parameters of a photogrammetric block at the same time (GRÜN, 1985).
The present article deals with a generalization of the concept of digital image matching allowing for direct object surface reconstruction from the density values of the images. It is possible to process any desired number of images, scanned in various spectral bands, simultaneously. The principle of this method was shown in (EBNER et al., 1987).
A similar generalization was given by (HELAVA, 1987) and (WROBEL, 1987).

## 2. The Algorithm

The object surface can be described using a geometric model z(x,y) and a radiometric model of density values g(x,y). The geometric modelling is carried out according to the finite element approach. For the radiometric description, object surface elements are defined forming a square grid in the x,y plane which in size correspond approximately to the pixels of the digital images multiplied by the image scale factor. A density value is assigned to every resulting object surface element.
The parameters of the geometric model, the assigned density values of the object surface elements, and the orientation parameters of all images are treated as unknowns and are estimated from the pixel density values and control information.
For reasons of simplicity in this paper, the geometrical model or Digital Terrain Model (DTM) shall be restricted to bilinear surface elements, forming a square grid in x,y as used e. g. in (EBNER et al., 1980). Each of the resultant grid meshes consists of n*n object surface elements (s. fig. 1).
The elevation of an object surface element can be represented as a linear function of the z coordinates of the four neighbouring grid points.
From the x, y, z coordinates of the center of the object surface elements and the orientation parameters of the relevant images, the image coordinates $x_I$, $y_I$ can be computed (s. fig. 2). The position $x_S$, $y_S$ of each point $x_I$, $y_I$ in the scanning coordinate system is calculated using e.g. an affine transformation with the fiducials as identical points.
Because $x_S$, $y_S$ are non-integer coordinates the corresponding image density value g must be computed from the neighbouring density values by e.g. bilinear interpolation. Thus, g is a function of the unknown coordinates $\hat{z}$ of the grid points of the geometric surface model and the unknown orientation parameters $\hat{p}$ of the images.

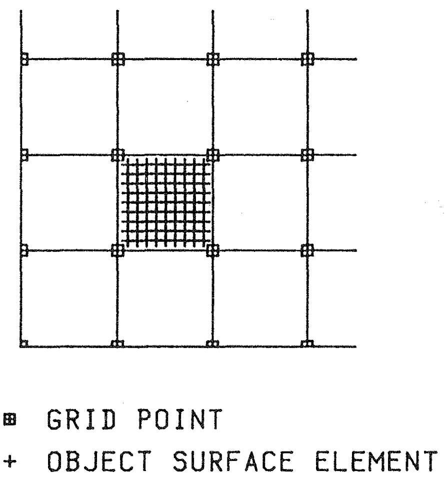


**fig. 1** : Grid points of geometric and object surface elements of radiometric model

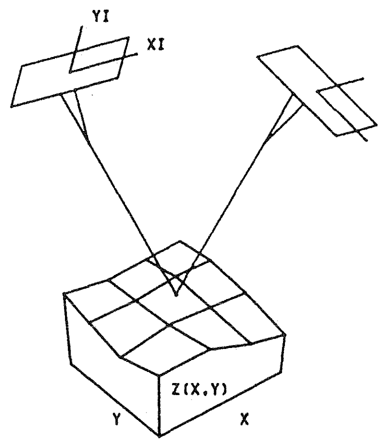


**fig. 2** : Connection between object and image coordinate system

The measurement of the fiducials in the scanning coordinate system is carried out using image matching techniques according to (ACKERMANN, 1984) and is described in more detail in (MAYR and HEIPKE, 1988).
If, for example, the considered object surface element is transformed into four images, the algorithm yields four image density values, which are then regarded as observations in a least squares adjustment. Four observation equations of the following type can be formulated :

$$\hat{v} = \hat{g} - g(\hat{z},\hat{p}) \tag{1}$$

where $\hat{g}$ is the unknown density value assigned to the object surface element and $\hat{v}$ the difference between $\hat{g}$ and the corresponding image density value $g$.
For all n*n object surface elements in every grid mesh the corresponding observation equations are formulated. The system is completed by adding the observation equations for the control information.
The equations (1) are nonlinear in $\hat{z}$ and $\hat{p}$, hence a Taylor expansion is necessary. The linearized observation equations read :

$$\hat{v} = \hat{g} - (\partial g/\partial z)_o \,\Delta\hat{z} - (\partial g/\partial p)_o \,\Delta\hat{p} - g(z^o,p^o) \tag{2}$$

where $z^o$ and $p^o$ stand for initial values and $\Delta\hat{z}$, $\Delta\hat{p}$ for the corrections of the unknowns $\hat{z}$ and $\hat{p}$ of the adjustment.
The image density values $g(z^o,p^o)$ are computed from the initial values $z^o$ and $p^o$ using the collinearity equations, an affine transformation and a given density value interpolation :

$$\begin{aligned} x_I &= x_I(z^o,p^o) \qquad \text{collinearity equations} \\ y_I &= y_I(z^o,p^o) \end{aligned}$$

$$x_s = x_s(x_I, y_I) \qquad \text{affine transformation}$$
$$y_s = y_s(x_I, y_I) \tag{3}$$

$$g(x_s, y_s) = \sum_{kl}\sum \alpha_{k,l}\, g(x_k, y_l) \;;\quad \sum_{kl}\sum \alpha_{k,l} = 1$$

$$x_k, y_l \in N \qquad \text{density value interpolation}$$

thus

$$g(x_s, y_s) = g(x_I, y_I) = g(z^o, p^o)$$

Expanding the coefficients $(\partial g/\partial z)_o$ and $(\partial g/\partial p)_o$ of the design matrix yields

$$(\partial g/\partial z)_o = (\partial g/\partial x_s \cdot \partial x_s/\partial z)_o + (\partial g/\partial y_s \cdot \partial y_s/\partial z)_o$$
$$(\partial g/\partial p)_o = (\partial g/\partial x_s \cdot \partial x_s/\partial p)_o + (\partial g/\partial y_s \cdot \partial y_s/\partial p)_o \tag{4}$$

$\partial g/\partial x_s$ and $\partial g/\partial y_s$ represent density value gradients of the digital image at the location $x_s(z^o, p^o)$; $y_s(z^o, p^o)$.
In the most simple case the weight matrix of the density value observations is represented by the identity matrix:

$$P = I \tag{5}$$

The solution of the adjustment is found iteratively applying the least squares method. The standard deviations for the unknowns are computed, too.
The structure of the matrix of normal equations is given in fig. 3. It becomes evident, that the unknown density values assigned to the object surface elements can easily be eliminated before the inversion of the matrix of normal equations, because the relevant submatrix is of a diagonal structure.

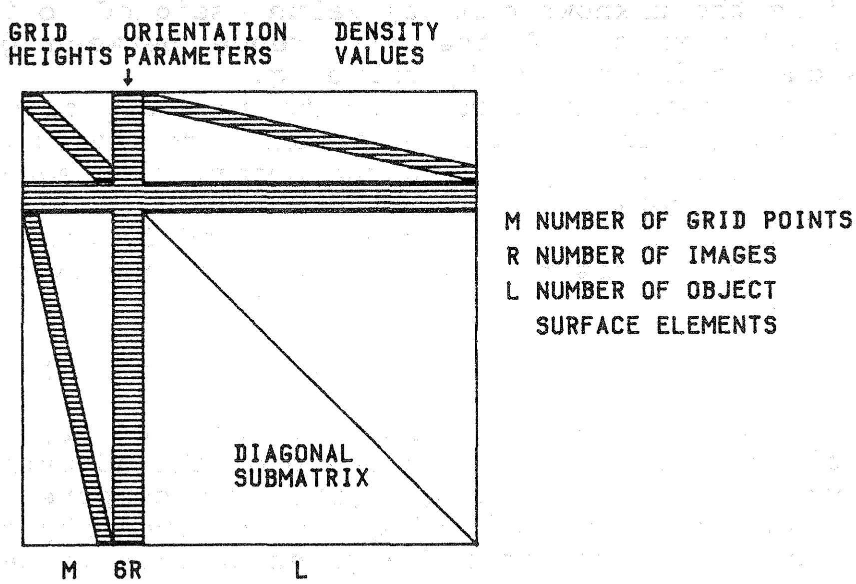


**fig. 3** : Structure of the matrix of normal equations

Radiometric differences between the images can be taken into account using linear polynomials. The coefficients of these polynomials are to be determined before or simultaneously with the actual adjustment. However, in one image they have to be set constant. The number of grid meshes, that each set of coefficients is valid for, also has to be selected beforehand.
The terms "observation" and "iteration" deserve a closer look. The observations are calculated from the initial values $z^o$ and $p^o$, which change in each iteration. Therefore, the observations also change in each iteration. This is in disagreement with the classical concept of adjustment. Thus, each iteration has to be considered as an independant least squares adjustment.
The explained approach combines digital image matching and object surface reconstruction to form a general model of digital photogrammetry. The pixel density values substitute the image coordinates as observations and the object surface is reconstructed as a whole. Digital point positioning can be formulated as a special case of this general method (EBNER et al., 1987).

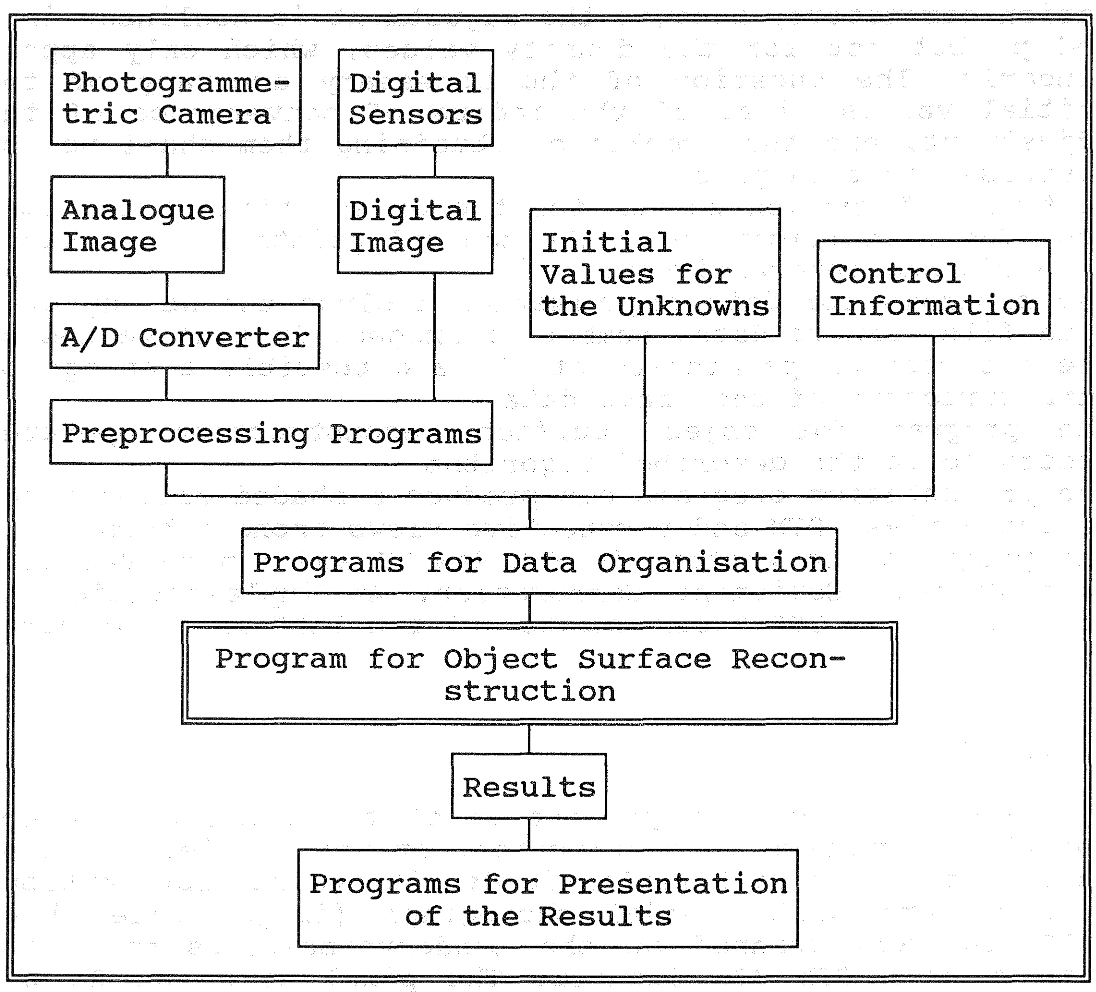


**fig. 4** : Principle data flow

## 3. The Data Flow

The principle data flow from the analogue or digital images to the resulting DTM is shown in fig. 4.
Analogue images taken with photogrammetric cameras have to be converted into a digital raster format. The resolution should be less than about 30 $\mu$m (ALMROTH, 1986; TRINDER, 1986), the depth of each pixel is typically 8 bits. The geometric stability of the A/D converter can pose a problem (BOOCHS, 1984).
Digital sensors, e.g. SPOT or an opto-electronic three line camera yield digital image data by their very nature.
Preprocessing programs include the definition and measurement of the fiducials in the digital image, an improvement of the density values (e.g. filtering, interactive removal of scratches by assigning new values to the effected pixels), and data reduction by cutting out the regions of interest.
The unknowns consist of three groups: heights for the grid points, orientation parameters of the images and density values assigned to the object surface elements. Initial values have to be introduced for the heights and the orientation parameters, because the adjustment is nonlinear in $\hat{z}$ and $\hat{p}$, but not for the density values, which only appear linearly. The question of the necessary accuracy of the initial values, i.e. of the radius of convergence of the adjustment, and the problem of obtaining them shall not be discussed in this paper.
Control information serves for the orientation of the images. The identification of the control points in the images is performed interactively.
The programs for data organisation include the set up of a data file (camera data, number of images, initial values of the orientation parameters etc.), and possibly a change of data structure of the image data.
The program for object surface reconstruction operates according to the described algorithm.
The presentation programs can produce a shaded relief image of the derived DTM and perspective views among others.
The programs are written in FORTRAN 77 and run on VAX 8300 from Digital Equipment Corporation. An implementation on the Siemens VP 200 Supercomputer of the IABG is in preparation.

## 4. The Test Data

In order to study the properties of the described method for object surface reconstruction, it was applied to simulated image data with realistic density value distribution. A black and white aerial photograph (image scale about 1:30.000) was scanned at the Landesvermessungsamt Munich with a Hell CTX 330 Scanner. The pixel size was 25 $\mu$m, corresponding to approximately 0.75 m at the ground, the depth per pixel was 7 bit (which is the maximum density value resolution of this particular scanner).
In the object space, a 20 m grid was defined, each mesh containing 21*21 object surface elements. Thus, the size of

one pixel at ground level approximately corresponds to the size of an object surface element.
The relationship between the scanning coordinate system and the object coordinate system was established by an affine transformation and predefined orientation parameters for the image.
Using an existing DTM for the defined grid, density values were assigned to the object surface elements by orthophoto projection.
This orthophoto was projected back into image space using four different sets of affine and orientation parameters.
In each of the resultant images a pixel matrix of density values with 25 $\mu$m pixel size was derived by nearest neighbourhood interpolation and a filling algorithm where necessary. Thus, four simulated images were generated forming two stereo models, one in each of two crossing flight paths, with a longitudinal overlap of 60 % and a scale of about 1:30.000.
As the system is consistent, except for the effects resulting from the computation of the pixel matrix of density values in each image, the computed heights must be almost identical to the used DTM heights.
The difference to real image data is, that there is no influence from possibly wrong orientation parameters and

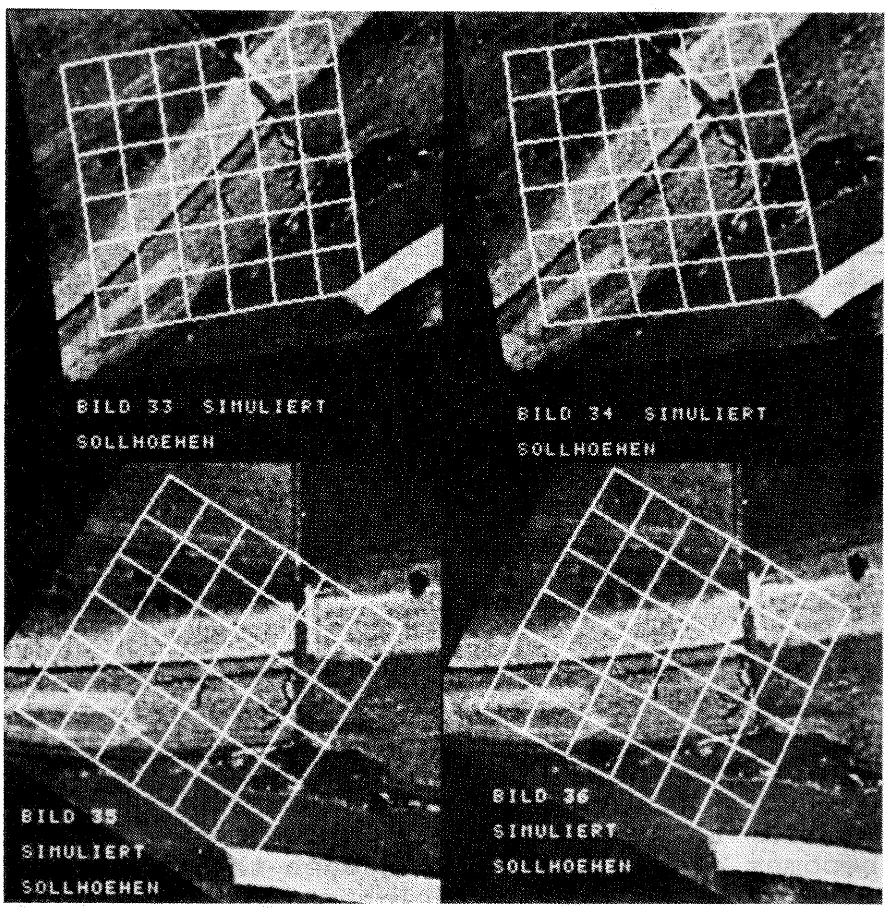


**fig. 5** : Parts of the four simulated images

from different photographs (involving occlusion, image noise, different light reflexion etc.).
The four simulated images can partly be seen in fig. 5. The DTM grid is projected onto the images for 6*6 meshes. The corresponding meshes in the individual images contain the same density information.

## 5. The Results

Various tests were performed with the simulated data. In all cases the orientation parameters of the images were helt constant. Thus, no control information was necessary. In the tests the initial values for the unknown heights and the number of grid meshes processed simultaneously were varied.
The reference DTM was defined as the DTM resulting from the 6*6 grid meshes and the four simulated images.
Because of effects due to the derivation of the pixel matrices of density values in the image space, the original DTM could not be used as reference DTM.
The initial values for the DTM were varied in two different ways. Either a constant shift was applied to all grid heights (labeled CON), or a random generator was used producing height shifts with mean zero and a predefined standard deviation (labeled SIG).
As convergence criterion a maximum height change per iteration of 0.01 m was used.
The results for 6*6 grid meshes can be found in table 1.
As can be seen, many iterations are necessary for convergence. Also, convergence occurs with wrong results (CON 10).

| height shift [m] | corresponding displacement [pixel] | Iterations | max. abs. diff. to reference DTM [m] |
|---|---|---|---|
| CON 3 | about 1 | 9 | 0.13 |
| CON 5 | about 2 | 20 | 0.18 |
| CON 10 | about 4 | 80 | 19.79 ! |
| SIG 3 | about 1 | 10 | 0.16 |
| SIG 5 | about 2 | 14 | 0.12 |
| SIG 10 | about 4 | 50 | 0.13 |

**table 1** : Results of object surface reconstruction for 6*6 grid meshes

In fig. 6 the derived DTMs and the difference of each height to the reference DTM are shown for CON 10 and SIG 10. It becomes evident, that even when the results in table 1 indicate, that no correct heights have been calculated, only the values at the grid border are wrong. Since these heights should normally be discarded after computation, the overall result is still correct.

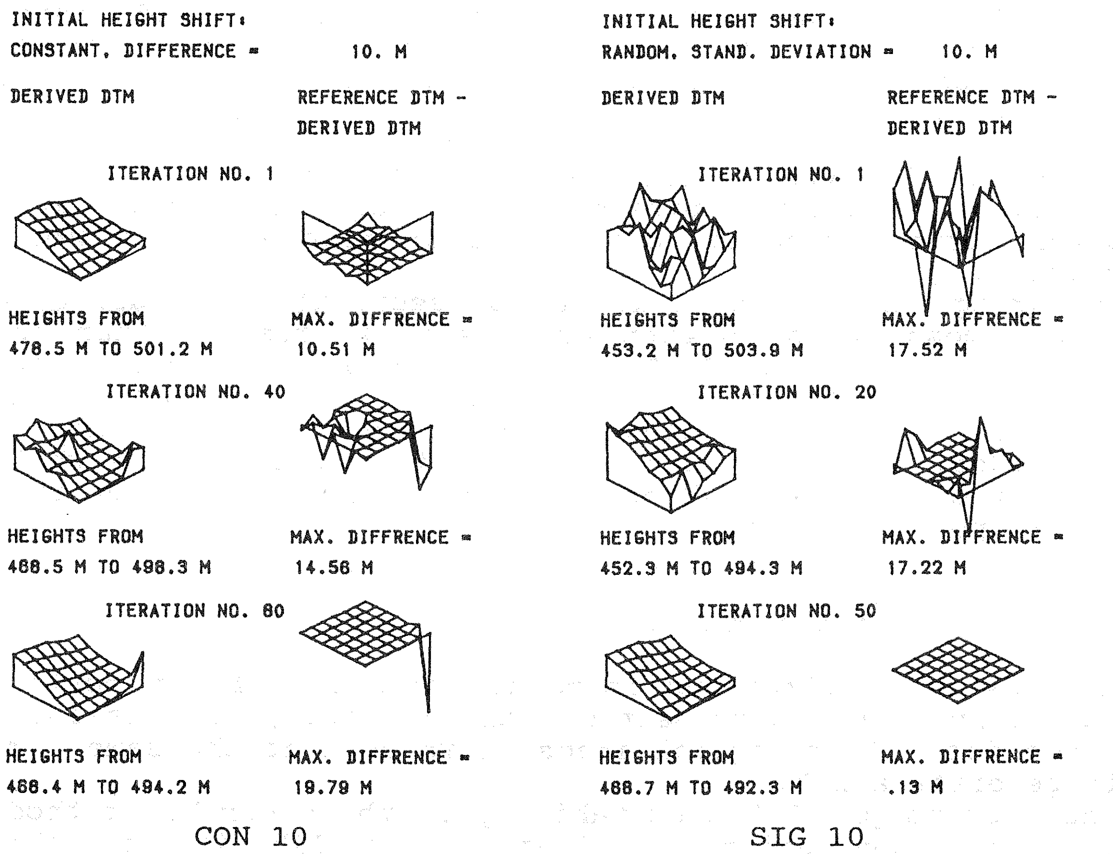


**fig 6** : Derived DTMs and differences to reference DTM

| height shift [m] | corresponding displacement [pixel] | Iterations | max. abs. diff. to reference DTM [m] |
|---|---|---|---|
| "Good" grid mesh | | | |
| CON 3 | about 1 | 9 | 0.22 |
| CON 5 | about 2 | 15 | 0.22 |
| CON 10 | about 4 | 32 | 0.22 |
| SIG 3 | about 1 | 9 | 0.29 |
| SIG 5 | about 2 | 7 | 0.29 |
| SIG 10 | about 4 | 28 | 0.22 |
| "Bad" grid mesh | | | |
| CON 3 | about 1 | 8 | 0.24 |
| CON 5 | about 2 | 19 | 0.21 |
| CON 10 | about 4 | 21 | 10.82 ! |
| SIG 3 | about 1 | 9 | 0.20 |
| SIG 5 | about 2 | 25 | 0.26 |
| SIG 10 | about 4 | 18 | 9.40 ! |

**table 2** : Results of object surface reconstruction for a "good" and a "bad" grid mesh

Table 2 shows the results for two individual grid meshes and four images. The two meshes were selected by visual inspection as one "good" and one "bad" mesh. "Good" means a possibly good matching result, "bad" a poor one.
The "good" mesh indeed shows a better performance than the "bad" one. At CON 10 and SIG 10, corresponding to a displacement of about 4 pixels in the image, the "good" mesh gives correct results, but the "bad" mesh leads to wrong results.
Comparing table 1 and 2 it can be seen, that the use of a "good" mesh leads to even more reliable results than the use of the whole grid. Taking into account the difficulty of deciding between "good" and "bad", and the fact, that the grid heights are correct, when only the grid border is discarded, it can be concluded, that processing the whole grid simultaneously leads to more reliable results than single point matching.

## 6. Conclusions and Outlook

A study on simultaneous image matching and object surface reconstruction was presented with simulated, but realistic image data. Four digital images were used at the same time to generate a DTM.
The results proof the suitability of the described method. The main conclusion is, that the realibility of this approach, a crucial point in image matching, is superior to the reliability of single point matching. The radius of convergence was found to be about four pixels.
Since there are significant differences between "good" and "bad" meshes, a pre-processing step will be designed to decide on the quality of each grid mesh for the matching procedure. Weights will be assigned to the corresponding observations accordingly. Thus, the advantages of feature based matching and object surface reconstruction as a whole can be combined.

## References

Ackermann, F. (1984), High Precision Digital Image Correlation, Inst. of Photogrammetry, Vol. 9, p. 231-243, University of Stuttgart.
Almroth, U. (1986), Digital Photogrammetry: Pixel Size, Image Quality and Noise Considerations, Photogrammetric Reports No. 2:51, "Digital Matching of Stereo Images", The Royal Institute of Technology, Stockholm.
Boochs, F. (1984), Zum Einsatz der Flächenkorrelation für die off-line Auswertung photogr. Modelle, Int. Arch. Phot., 25, Part A3a, p. 49-58, Rio de Jainero, (in German).

Ebner, H./Hofmann-Wellenhof, B./Reiß, P./Steidler,F.(1980), HIFI - A Minicomputer Program Package for Height Interpolation by Finite Elements, Int. Arch. Phot., 23, Part B4, p. 202-215, Hamburg.
Ebner, H./Fritsch, D./Gillessen, W./Heipke, C. (1987), Integration von Bildzuordnung und Objektrekonstruktion innerhalb der Digitalen Photogrammetrie, Bildmessung und Luftbildwesen, p. 194-203 (in German).
Förstner, W. (1984), Quality Assessment of Object Location and Point Transfer using Digital Image Correlation Techniques, Int. Arch. Phot., 25, Part A3a, p. 197-219, Rio de Jainero.
Förstner, W. (1986), Digital Image Matching Techniques for Standard Photogrammetric Applications, Technical Papers, ACSM-ASP Convention, Vol. 4, p. 210-219, Washington.
Grün, A. W. (1985), Adaptive Least Squares Correlation - A Powerful Image Matching Technique, South African Journal of Photogrammetry, Remote Sensing and Cartography, p. 175 -187.
Rosenholm, D. (1986), Accuracy Improvements in Digital Matching, Phot. Reports No. 2:52, Dep. of Photogrammetry, The Royal Institute of Technology, Stockholm.
Helava, U. V. (1987), Digital Comparator Correlation System. Proceedings, Intercomm. Conference of ISPRS on Fast Processing of Phot. Data, p. 404-418, Interlaken.
Mayr, W. and Heipke, C. (1988), A Contribution to Digital Orthophoto Generation, Presented Paper, Commission IV, ISPRS Congress, Kyoto.
Trinder, J. C. (1986), Geometric aspects of digitized photography. Int. Arch. Phot., 26, Part 1, p. 471-475, Stuttgart.
Wrobel, B. (1987), Facets Stereo Vision (FAST Vision) - A new Approach to Computer Stereo Vision and to Digial Photogrammetry. Proceedings, Intercomm. Conference of ISPRS on Fast Processing of Phot. Data, p.231-258, Interlaken.