TABLE OF CONTENTS

1. INTRODUCTION

In photogrammetry and remote sensing, matching can be defined as the establishment of the correspondence between various data sets. The matching problem is also referred to as the correspondence problem. The data sets can represent images, but also maps, or object models and GIS data. Many steps of the photogrammetric processing chain are linked to matching in one way or another. Examples include the reconstruction of the interior orientation: the image of a fiducial is matched with a two-dimensional model of the fiducial; relative orientation and point transfer in aerial triangulation: parts of one image are matched with parts of other images in order to generate tie points; absolute orientation: parts of the image are matched with a description of control features, mostly ground control points; generation of digital terrain models (DTM): parts of an image are matched with parts of another image in order to generate three-dimensional object points; and finally the interpretation step: parts of the image are matched with object models in order to identify and localize the depicted scene objects.

Looking at this large variety of tasks it comes as no surprise that matching has long been and still is one of the most challenging tasks in photogrammetric research and development. Here, an overview is given of a more specific class of matching algorithms usually called digital image matching, especially for DTM generation. Digital image matching automatically establishes the correspondence between primitives extracted from two or more digital images depicting at least partly the same scene. The primitives can be grey level windows or features extracted from the images. Thus, all input data sets are images or functions thereof. Objects as such don't need to be modelled explicitly. It should be remembered, however, that each algorithm uses at least an implicit model of the object surface, since it is the object surface which is depicted in the images. In this work, the aim is to describe the several matching methods used for DTM generation and the criteria that are being used in order to check and control the matching in terms of accuracy and reliability. The knowledge to be gained out of this study of literature sets the base for further investigations in matching algorithms.

2. IMAGE MATCHING BACKROUND

First solutions for image matching have been suggested already in the late fifties (Hobrough, 1959); he still used analogue images and procedures. Since then, a steady increase in the interest for image matching has occurred, and the question may be asked, why image matching has not been solved long ago. A first answer can be given by considering the information content of the most elementary primitive in the input data set, namely a pixel. An aerial image scanned with 15 μm contains approximately 235,000,000 pixels, and each gray value usually lies in the range of 0 to 255. Assuming an equal distribution of the grey values, the image contains roughly 920,000 pixels of each grey value.

This little computation demonstrates that matching on the basis of single pixels is certainly impossible. It also exemplifies two fundamental problems of image matching:

1. Ambiguous solutions may occur, if image matching is tackled using only local information
2. Good approximations are needed

Using image matching, we try to reconstruct the three-dimensional object surface from two-dimensional projections. During this projection information is lost. This is most evident in the case of occlusions. Image matching belongs to the class of so called inverse problems, which are known to be ill-posed. A problem is ill-posed, if no guarantee can be given that a solution exists, is unique, and is stable with respect to small variation in the input data. Image matching is ill-posed, because for a given point in one image, a corresponding point may not exist due to occlusion, there may be more than one possible match due to repetitive patterns or a semi-transparent object surface, and the solution may be unstable with respect to noise due to poor texture (see e.g. Zheng, 1993 for a more detailed discussion on ill-posed problems).

In order to find the solution of an ill-posed problem, one usually has to deal with an optimisation function exhibiting many local extrema, and thus a small pull-in range. Therefore, stringent requirements may exist for initial values of the unknown parameters to be determined. Moreover, usually there is a large search space for these parameters, and numerical unstabilities may arise during the computations. Ill-posed problems can be converted, at least partly, to well-posed problems by introducing additional knowledge about the problem and costraints.

3. IMAGE MATCHING ALGORITHMS

The distinction between different matching primitives is probably the most prominent difference between the various matching algorithms. The primitives fall into two broad categories: either windows composed of grey values or features extracted in each image a priori are used in the actual matching step. The resulting algorithms are usually called:

1. Area based matching (ABM), and
2. Feature based matching (FBM), respectively.

Area-based matching, sometimes called signal-based matching is considered more traditional. The cross correlation and the least squares matching are well known methods for ABM. These methods usually require very good initial values for the unknown parameters.

Feature-based matching determines the correspondence between image features and it does not require very precise initial estimates. In both cases, there is a choice between local and global support for the primitives. The terms local and global are not sharply defined. Local refers to an area seldom larger than about 15 * 15 or 21 * 21 pixels in image space, global means a larger area and can comprise the whole image. Local matching is very precise, but ambiguous and the various local features may not be consistent with each other. Global matching is more robust but not necessarily accurate in local areas. There are algorithms characterized as single point algorithms and their weakness is that they are still not capable of taking into consideration the matching results of other points around the points that are being matched separately. This means that the method does not consider the compatibility of matching results amongst neighbouring points. Obviously, the reliability of matching is closely related to the information in the match window. The less the information, the less the reliability. Some researchers think image correlation based on grey level is changing to matching based on edges recognition (Greenfeld and Schenk, 1989) . However, some images lack edge features. Under this situation, surface interpolation is often used between the edges. The result is often not satisfactory.

In general, algorithms now frequently use constraints on the primitives in order to find an optimal solution. The most common ones are:

- **Epipolar constraint**: two homologous points must respectively be located on the two epipolar lines of the respective object point. This constraint depends on the geometry of the aerial survey and is independent of the scene's content.
- **Uniqueness constraint**: the primitive of an image cannot have more than one homologous primitive in the other image. This rule often has exceptions, such as occlusions, transparent surfaces or different points that lie on a line, which is being seen as a point.
- **Surface continuity constraint**: depending on the surface we can assume e.g. planar surface.
- **Order constraint**: if p1 is on the right of q1 in image1, then p2 is also on the right of q2 in the second image. An inversion of points occurs only for objects that present a high perspective deformation on the image.

- **Photometry constraint**: based on Lambert's hypothesis, for an object point the reflected light intensity should be the same in all directions. However, many reasons may cause radiometric variations between two overlapping images.

3.1 Area-based matching techniques

In ABM, each point to be matched is the center of a small window of pixels in a reference image (template) , and this window is statistically compared with equally sized windows of pixels in other (target) images . The measure of match is either a difference metric that is minimised, such as RMS difference, or more commonly a similarity measure that is maximized, such as mean– and variance–normalized cross-correlation. ABM is usually based on local windows.

3.1.1 Cross- correlation

In order to compute the cross-correlation function of two windows, a template window is shifted pixel by pixel across a larger search window (see Fig. 1), and in each position the cross-correlation coefficient ρ between the template window and the corresponding part of the search window is computed according to Eq. (1). The maximum of the resulting cross-correlation function defines the position of the best match between the template and the search window. The cross-correlation coefficient is is a simple but widely used measure for the similarity of different image windows (see e.g. Sharp et al., 1965; Kreiling, 1976; Hannah, 1989)

$$\rho = \frac{\sum_{r=1}^{R}\sum_{c=1}^{C}(g_1(r,c)-\mu_1)(g_2(r,c)-\mu_2)}{\sqrt{\sum_{r=1}^{R}\sum_{c=1}^{C}(g_1(r,c)-\mu_1)^2\sum_{r=1}^{R}\sum_{c=1}^{C}(g_2(r,c)-\mu_2)^2}} \quad ; \quad -1 \le \rho \le 1 \qquad \textbf{(1)}$$

Where:

g_1 (r,c) individual grey values of template matrix

μ_1 average grey value of template matrix

g_2(r,c) individual grey values of corresponding part of search matrix

μ_2 average grey value of corresponding part of search matrix

R, C number of rows and columns of template matrix

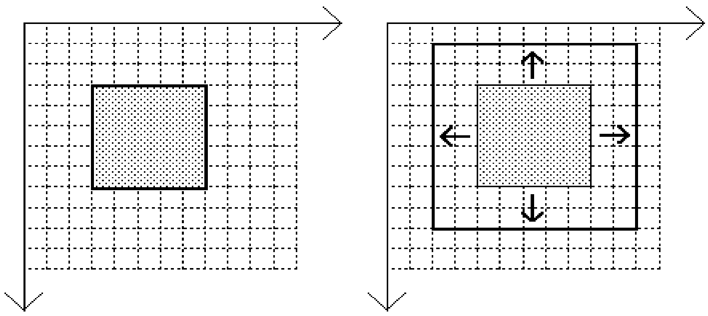


Fig. 1. Principle of cross correlation: the patch window on the right is shifted pixel by pixel across a larger search window and the maximum cross-correlation to the template on the left is found

Cross-correlation works fast and well, if patches to be matched contain enough signal without too much high frequency content (noise) and if geometrical and radiometric distortions are kept at minimum. Both conditions are often not encountered in or with aerial images.

3.1.2 Least squares matching

An approach on least squares filters for object location has been described in Foerstner (1982). Ackermann (1984) proposed least squares correlation and since then it has received a wider attention in photogrammetry. Gruen (1985) presented an extension to LSM method, which is going to be discussed in section 3.1.3. This method is based on the minimization of the squared differences of the grey values between two (or more) image patches.

Assume two image regions are given as discrete two-dimensional functions f(x,y) , g(x,y), which might have been derived from continuous functions. f(x,y) and g(x,y) can be defined as conjugate regions of a stereopair in the ’’left’’ and the ’’right’’ image. f(x,y) is the template, g(x,y) is the patch in the other image. Correspondence is established if

$$\mathbf{f(x,y) = g(x,y)}$$

Because of random effects (noise) in both images, the above equation is not consistent. Therefore, a noise vector e(x,y) is added, resulting in

$$\mathbf{f(x,y)-e(x,y) = g(x,y)}$$

The location of the function values g (x, y) must be determined in order to provide for the match point. This is achieved by minimizing a goal function, which measures the distances between the grey levels in template and the other patch. The goal function to be minimized in this approach is the L_2-norm of the residuals of least squares estimation. The location is described by shift parameters which are estimated with respect to an initial position of g (x, y). In order to account for a variety of systematic image deformations and to obtain a better match, image shaping parameters (affine image shaping) and radiometric corrections are introduced beside the shift parameters.

3.1.3 Adaptive Least Square Matching

This method was presented by Gruen, 1985. It combines the geometrical model of LSM and the collinearity constraints and multiple use of more than two images. This combination leads to the simultaneous determination of 3-D coordinates. The designation ″adaptive″ indicates the capability of the algorithm to restrict the type and the number of modelling parameters to those, which are safely determinable (excluding those individual parameters, which are nondeterminable in order to avoid the danger of overparameterization). Since the determinability of individual parameters depends on the structure of the match signal, the aforementioned goal could be achieved by a priori analysis of the signal and by a rule-based selection of a fixed set of parameters prior to adjustment. Depending on the surface form and image texture, it can use different number of parameters in order to model the surface e.g. if the local terrain surface patch is a plane in sufficient approximation, then the corresponding image can be strictly described by a projective transformation between object and image and since it can usually be assumed that the facet image is very small with respect to the full image format this projective transformation (8 parameters) can be approximated by an affine transformation (6 parameters). This method is a widely accepted technique for matching with subpixel accuracy.

3.1.4 Overview of some additional ABM methods

Instead of using L_2-norm ($\sum v^T Pv = \min$) several investigations have shown that L_1-norm or absolute deviation (LAD) can be used to improve the accuracy of the estimation problem in the presence of outlier pixels in the data, (Calitz and Ruther, 1994). The LAD algorithm is based on the method of Barrodale and Young (1966), but this algorithm was found to be subject to numerical instabilities and tended to converge rather slowly. As an improvement to that, the Barrodale and Roberts implementation was adapted to the correlation problem and found to converge stable. In general, the LAD algorithm achieves a more accurate matching, e.g. if the central points of the template and the patch which are to be matched fall in the boundary of the largest continuous region it is proven that this method gives better results in the presence of outliers.

Gruen (1985) proposed a method of **multipatch matching** with simultaneous computation of parallaxes and object coordinates in grid meshes by forcing through fictitious, weighted observations the corner points of each mesh to have the same parallax in all neighbouring meshes.

Rosenholm (1987,1988), Rauhala (1987) and Li (1989) focused on **multipoint matching (MM)** algorithms, that are image-based and use simultaneous computation of parallaxes in grid points which are connected with bilinear finite elements describing the parallax differences. Thus, the object model used is a continuous surface with continuous first derivatives. They also use additional fictitious, weighted, continuity constraints on parallaxes to strengthen the connections between the grid points. Rosenholm uses multipoint matching to bridge areas with poor signal content. The possible constraints on the parallaxes are:

1. The second derivatives of the parallaxes are zero (minimization of the curvature). The results of Rosenholm are based on the use of this constraint.

2. The first derivatives of the parallaxes are zero (minimization of the slope). This method can be of importance at the border of the grid.

Rauhala uses more global solutions (matching of even whole images) with the to increase the accuracy and especially the reliability of matching without any success.

Li investigated the problem by including breaklines and embedded the problem in a multiresolution, multigrid approach. For the detection and localization of breaklines he first determines breakpoints by two different methods and then uses either a line-fit or a Hough transform to detect the position and orientation of the breakline. This information is used as approximation in the MM model with breaklines which localises the breaklines accurately simultaneously with the parallax determination at the grid nodes. He assumes that there is only one breakline going through the whole grid, and for the detection stage this is a straight line. Furthermore, it is assumed that the parallax along the breakline is smooth and that there is enough texture across the breakline. Through this approach, to model even the simplest case of breaklines is very complicated and time consuming.

Baltsavias (1991) focused on multiphoto geometrically constrained matching by introducing two new elements: (a) the exploitation of any a-priori known geometric information to constrain the solution and (b) the simultaneous use of more than two images. The method uses the least squares matching technique and the least squares observation equations consist of two parts: the equations that formulate the grey level matching and the equations that express the geometric conditions that must be fulfilled. The two parts must be related to each other through unknown parameters or known functions of them that appear in both sets of equations. This algorithmic approach offers the additional advantage of determining simultaneously in one step additional unknown parameters that are required, e.g. X, Y, Z object coordinates. Generally, a good texture implies the existence of edges. With this to achieve high precision and avoid multiple solutions the measurement points should lie along edges with a large enough intersection angle with the epipolar lines. In Multiphoto Geometrically Constrained Matching the optimal points are selected only in the reference image by using an interest operator that detects pixels along edges. To avoid point clusters, a thin out of the results is applied and the best point within a neighbourhood is kept. Best is the point with the highest interest value.

Methods based on dynamic programming (Bobick, 1999, Horiuchi, 1994) tries to do the following. For grey levels of two epipolar lines, a graph of costs is built for the pairs composed of one point on the reference line (manually selected) and another point one the research line. The cost function represents the similarity of their possible correspondence; it might be the difference between the grey levels of the two points or also a more complex function. A 2D array of costs is built, with the points of the left epipolar line in abscissa and points of the right epipolar line in ordinate. Then, the algorithm computes the sequence of all best corresponding pairs, according to the costs previously calculated. For each path of this graph, a total cost is given by the sum of the costs of all involved pairs. The path with the minimal total cost is computed using a dynamic programming method such as the Viterbi algorithm (Forney, 1973).

A method of least-squares matching in object space is being mentioned in Wrobel (1987), Helava (1988), and Ebner (1987). As a part of the effort to develop object-space correlation techniques, the concept of ″groundel″ of ″surfel″ has emerged. The groundel is a unit in object space similar to pixel in image space. In order to visualize this concept, the terrain is checkered into small squares, for which each center is positioned in elevation to represent the average elevation over this area. The object is modeled by two surfaces, a geometric (height) and an optical density surface and the relation between image and object space is through the known collinearity equations. These surfaces are estimated at the nodes of the regular grid and the grid spacing for the radiometric grid corresponds to the groundel. The size of the groundel should approximately be the size of the pixel times the scale factor. The groundel contains at least one density value and the density value spacing is approximately equal to the pixel size times the scale factor. The grid spacing for the geometric surface is usually larger than that of the radiometric but it can even be equal, (i.e. case where the height is estimated at each radiometric groundel. The parameters are estimated by a simultaneous least-squares adjustment. The observations are the pixel intensities. The unknown parameters are the heights and the densities at the nodes of the two grids. Other parameters like sensor parameters (interior, exterior orientation and radiometric parameters) and illumination model parameters can theoretically also be determined but in practice only the two surface grids or even only the height grid is estimated.

Schenk (1997) proposed the method of multiple patch matching in the object space, which employs matching of warped images and matching more than two images simultaneously. The mathematical model for matching multi-image patches is an extension of the classical LSM. In multiple-patch matching, with a not very good approximation of the point location in the images and approximate orientation parameters, the image patches that are centered on every point are warped, using a surface model and then matched. The image patches are warped with respect to the surface patch and the matching is performed between warped patches. In the first iteration, when no surface approximation is available a horizontal plane is assumed. In each iteration a warped image patch is computed in the object space, centered on the approximate point. Each matched point is projected back to the image space and through forward intersection a new correct point is determined. The warping is using a local regular grid. The grid points are matched using LSM and the local surface is refined through backprojection and intersection.

3.2 Feature-based matching techniques

FBM comprises two stages. Firstly the detection of interesting features and their attributes in all images and secondly, the determination of the corresponding features. The two stages are related to each other in the sense that the feature extraction and computation of their attributes must be such that the second stage of the correspondence is easy, not sensitive to errors and precise.

3.2.1 Feature extraction

In FBM, features are extracted in each image individually prior to matching them. Local features are points, edges and lines, and regions. Larger (global) features are called structures. Global features are usually composed of different local features. Besides the attributes of the local features, relations between these local features are introduced to characterize global features. These relations can be geometric such as the angle between two adjacent polygon sides or the minimum distance between two edges, radiometric such as the difference in grey value or grey value variance between two adjacent regions or topologic, such as the notion that one feature is contained in another. Matching with global features is also referred to as relational or structural matching (Shapiro and Haralick, 1987). Features should be distinct with respect to their neighbourhood, invariant with respect to geometric and radiometric influences, stable with respect to noise, and seldom with respect to other features.

3.2.1.1 Point extraction

The point features are usually extracted by local operators, often called ’’interest“ operators. The attributes are computed within a rectangular or circular window, in selected or in all directions and are usually compared to a threshold to decide whether a feature is good or not. To avoid many neighbouring pixels a window to suppress the local non-maxima is used. Within this window only the feature with the best attribute values (e.g. edge pixel with the highest grey level gradient) is kept. After the features have been detected, some operators require a reduction of the selected window centers to the actual position of the feature. The computed attributes of each feature depend on the dimensionality and type of the feature and on the desired computational complexity and matching accuracy.
Detection of edge points can be performed by edge detectors mentioned in section 3.2.1.2.
Some interest operators are listed below:

1. The operator of Moravec (Moravec, 1977, 1979, 1980) detects points with high grey level variances. A uniform threshold (empirical) is used for the whole image and no points are detected in low contrast image regions.
2. Luhmann and Ehlers (1984) used a grid-like feature extraction without contrast-depending thresholding to obtain a homogeneous distribution of the features.
3. Dreschler (1981) developed an interest operator based on the principal and Gaussian curvatures and the angle of the principal direction of the intensity function. As interest value the corner, i.e. the zero crossing of Gaussian curvature between two extrema, is used.

4. The Foerstner operator (Foerstner 1986a, 1986b, 1987) selects optimal windows for template or least squares matching, for corner detection and for the extraction of centers of circularly symmetrical features. A feature is selected if the windows grey level signal ellipse is small and circular based on two perspective thresholds. In Fig.2 the localization of the point is being shown The procedure that the Foerstner operator works is as follows:

 – Compute the grey level gradients of the image g_x , g_y by using an edge operator (e.g. Sobel)
 – Compute the elements of the normal matrix N

$$N = \begin{bmatrix} \sum g_x^2 & \sum g_x g_y \\ \sum g_x g_y & \sum g_y^2 \end{bmatrix}$$

 – Estimate two criteria within a small window, e.g. 7x7
   a) weight w, relates to the size of the error ellipse, proportional to contrast

$$w = \frac{\det N}{trN} \quad \text{with}$$

   det N = determinant of N; tr N = trace of N

   b) roundness q, relates to the shape of the error ellipse, i.e. the length of the semimajor and semiminor ellipse axes

$$q = \frac{4 \det N}{tr^2 N}\ ;\ \ 0<q<1$$

 – Threshold criteria

   w (x,y) = w (x,y) if g (x,y) >=q_{min} and w (x,y)>= w_{min} otherwise

   w (x,y) = 0

 – Non-maxima suppression by using w (x,y) within window size
 – Estimation of accurate point coordinates

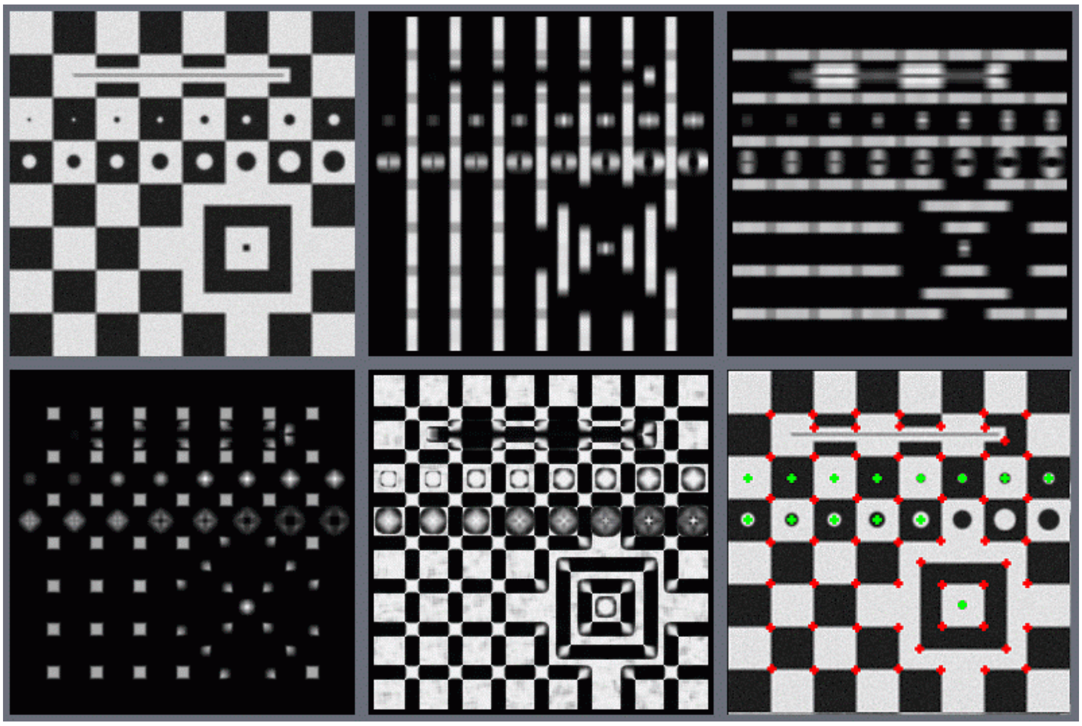


Fig.2. Rowwise from top left: 1. Original image, 2. and 3. Elements of the normal equation matrix N_{xx}, N_{yy} computed from Sobel gradients. 4. Size (area) of the error ellipse = w, 5. Roundness of the error ellipse = q, 6. Points chosen after non-maxima suppression

5. Hannah (1974) uses an operator that selects points with steep autocorrelation function of the grey levels in all directions (conceptually similar to Foerstner)

3.2.1.2 Edge extraction and matching

Edges can be extracted from operators that detect edgels. Two of the most known are:

- Gradient operators, which is based on the calculation of the local intensity gradient with the use of a mask, usually with odd dimensions. The ″Sobel operators″ and the ″Prewitt operators″ are two of the most used ones. The detected are edgels are thick.
- The zero-crossings of the Laplacian, which are inflection points of the intensity function *f(x,y)* .

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$$

In the following Figures (3,4) is being shown the result of Sobel operator and Laplacian operator on an image.

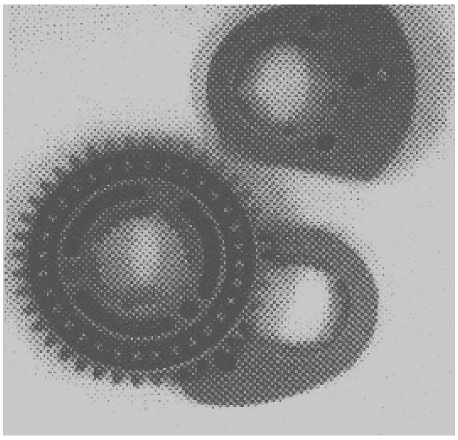
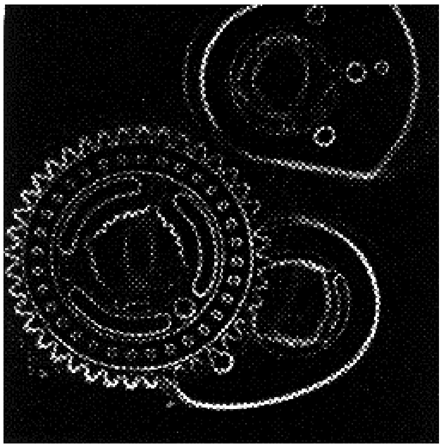



Fig.3: Left original images and right the Sobel operator has been applied and in some region the edges has been smeared out

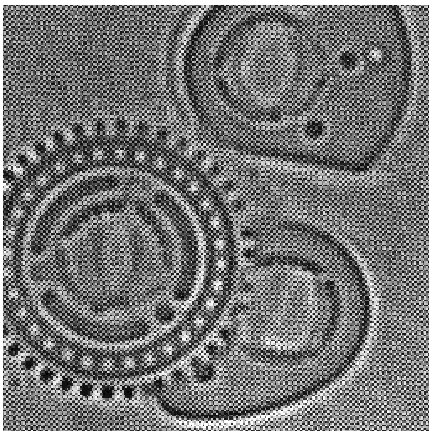
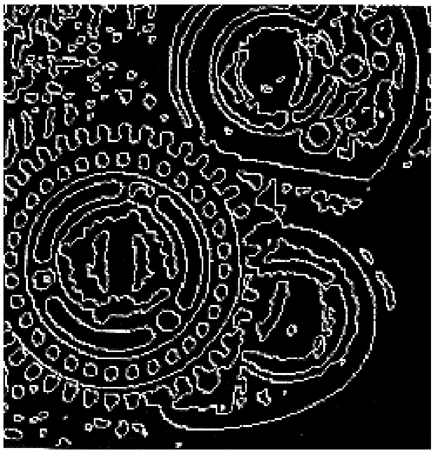



Fig.4: Left original image and right the Laplacian operator has been applied and in some regions edges are artificial

The Laplacian is based on second-order derivatives, thus is highly sensitive to noise and therefore is combined with a smoothing operation, e.g. convolution with a Gaussian filter. The use of the Laplacian has often been advocated with the argument that it yields closed contours but these closed edges are sometimes wrong or include artifacts

Furthermore, apart from the individual detection of pixels that lie on an edge, the combination of such pixels into linked lists that described edges as one-pixel thick connected lines should be taken into account. *Hough transform* allows the detection of any predefined shape. *Salient edge point linking* looks for the pixel with the highest intensity gradient and then starts growing the corresponding edge (this is done by considering all neighbouring pixels with an intensity gradient magnitude above some threshold that can slightly change). *Dynamic programming methods* can be used to connect pixels by an optimal path, optimal being defined on the basis of a cost function. A method used for edge detection and tracking (Gruen, 1994), is based on least squares matching and an edge pattern is introduced as the reference template which is to be subsequently matched with image patches containing actual edge segments.

The followings methods refer to edge extraction for feature matching or to both feature extraction and matching:

1. Marc and Hildreth (1980) developed an operator to detect grey level changes. It is a Laplacian of a Gaussian or approximately the difference of two Gaussians. The zero-crossings of the filter are used as features with attributes the sign of the zero-crossings and the orientation of the contour to which the zero-crossing belongs. This procedure is applied to epipolar images in a multi-resolution fashion.
2. Lemmens et al. (1988) extract lines and edges which are invariant to geometric distortions. An approach in order to produce reliable matching through a selection of well scattered windows, is to pass a statistical operator over the image and find the appropriate positions (this is a product of the image variance and the minimum of ratios of directed differences over windows of the specified size). Local peaks in the output of this operator are recorded as the preferred places to attempt the matching process. This operator penalizes windows with low information and windows whose only information is contained in strongly linear edges
3. Zhang (1989) uses as features the two edge ramp end points (upper and lower) and the zero crossings between. Windows with such a feature at the beginning and the end of the edge are established in epipolar images. The first feature of the first line is usually matched manually. The possible corresponding windows are restricted by using the attributes of the end feature, the remaining candidates are resampled to remove scale differences and the window with the maximum correlation coefficient, given that exceeds a certain threshold, is selected as the corresponding one. The corresponding windows are reduced to corresponding points by using a small window centered at the zero crossings of the corresponding end features. The correspondence is found by using resampling and the correlation coefficient as mentioned before. The edge-line following technique is applied to transfer the matched points of the previous epipolar line to the current line as control points for further matching.
4. Zhang (1990) extended the bridging-mode method from a feature oriented one to image-segment oriented one (where features are taken as borders of image segments). In this extended bridging method, the whole target window is regarded as the ″image″ segment to be matched. According to several constraints such as the similarity of image features and search range a few alternative search windows on the right epipolar line are found. Another extension of the method is from one-

dimensional to two-dimensional. The matching window in the one-dimensional bridging mode method is the segment which uses feature points as its borders, while the matching window in the two-dimensional bridging mode method is the image segment which uses edge lines as its borders.

5. Relaxation algorithms are applied in most cases in FBM and combine the advantages of parallel and sequential methods. Relaxation can take into account the results of neighbouring point in its processing. The characteristics of relaxation matching are:
   - Global application with parallel processing of spatial localized inputs
   - Iterative process with partial results
   - Multiple potential assignments
   - Attempt to optimize some function

   Assume that we have an object set O = $\{O_1, O_2,...,O_n\}$ and a class set C = $\{C_1, C_2,...., C_m\}$. The problem is to define each object O_i that belongs to a class C_j. The a priory probability P_{ij}^o of $O_i \in C_j$ can be defined and the compatibility coefficient C (i,j;h,k) of the case $(O_i \in C_j) \cap (O_h \in C_k)$. The prior P_{ij}^o is done independently from one another. Relaxation processing is a new probability estimated using the previous probability and the compatibility coefficient. This technique is being used in Global Matching with Relaxation and we assume the image point in the left image as ″object″, and the point in the right image as ″class″. As compatibility measurement the correlation coefficient between the left and right image segment is being used. After P_{ij} and C(i,j;h,k) are defined, a global image area can be matched.

6. Zhang (2000) detects edge pixels with the Canny operator and an edgel aggregation method is applied to generate contours with small gaps bridged using the criteria of proximity and collinearity. All segments are checked using their direction along their length and split at points where the change in direction exceeds a given value. For each line which satisfies the epipolar constraint, their attributes are used to compute a similarity score and through the probability relaxation method the matching is being done. Structural matching (see below) establishes a correspondence of the primitives of one structural description to the primitives of a second structural description and is conducted bidirectionally from left to right and from right to left.

3.2.1.3 Regions extraction

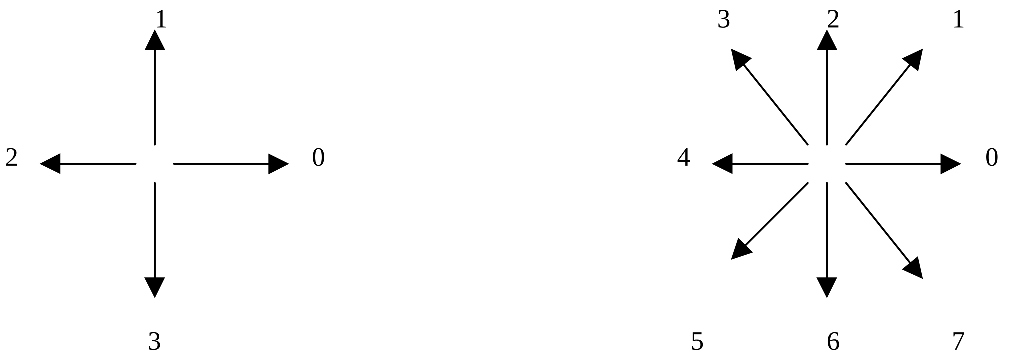
This extraction of regions starts with the detection of some very "homogeneous regions" of the image, e.g regions with intensity variations below a certain threshold, and are then maximally grown but only to the extent that they remain sufficiently homogeneous (region growing stops at pixels that violate the homogeneity criteria. Region-based methods can fail if the initial homogeneous ″seed″ regions to be grown are chosen erroneously, e.g. do not fall in all homogeneous areas.

Also a method for image registration can be used to match homologous regions in two images. Patch-based matching in images based on shape was presented from Dezfouli and Freeman (1994). In fact the word patch means region, because the shape of it can be different from one another depending on the case.

This method identifies templates (at least 3*3), in which the grey values differ by no more than a small number. The candidates in the second image are found in case they satisfy a number of criteria of similarity in comparison with the template (e.g. shape, size and relative geometry of patches). Width of the bounding rectangle, height of the bounding rectangle, and area of the patch are used as measures of size. The shape of the patch has been measured using a modified chain code method and the multi-patch process uses shape and size to reduce the number of eligible patches. The chain code is a sequence of connected straight-line segments of specified length and direction between neighbouring pixels. Each direction can be represented with a small integer Fig. 5. As measures of the shape of the patch, the frequencies of the direction segments in the chain code and the frequencies in of change in direction segments are being used. The multi-patch process based on the relative geometry works as follows.

1. Three patches are chosen in the first image.
2. A search is conducted in the candidate list of each of the three patches for a combination of candidates that form a triangle of similar size, shape and orientation to the patches in the first image.
3. If no suitable set of candidates can be found, one of the patches in step 1 is replaced, and step 2 is repeated.
4. Steps 2 and 3 are repeated until suitable triangles are formed.
5. With the first two successfully matched patches, other patches are processed in turn from the first image, with the candidate list being compared as in step 2 to find the one with the correct geometry.
6. The triangles are compared by the orientation and length of their edges, and are regarded as similar if edge orientation is approximately 5 ^{o}.

Patch –matching is more successful with large patches than with small, because in the latter case there are many more candidates. False matches occur most often when there are severall small patches relatively closed together.



(a) (b)

Fig.5 Neighbourhood numbering schemes defined using (a) 4 directions (b) 8 directions

There are several other region extraction methods, which are not presented here, as regions are almost never used as features in matching for DTM generation.

3.3 Structural matching

Also structural image matching is beeing used (Wang, 1998). Structural matching is sometimes referred to as relational matching. It establishes a correspondence from the primitives of one structural description to the primitives of a second structural description. A structural description is defined by a set of primitives (features) and their interrelationships. Since structural matching techniques utilize not only image features but also topological and geometrical relations among the features to determine the correspondence, the image matching tasks can be fully automated without any initial estimates or very coarse ones. Each primitive and relation can be decribed by several attributes. For example a point primitive P_I can be described as P_I = { coordinates = (x,y), greyvalue = g, gradient = t} and a relation r_j between two lines as : r_j = { lines l_m, l_n,(cross = yes, angle = a)}. The goal is to find a correspondence or the best match between the primitives and relations of two structural descriptions. An evaluation function is being applied. According to the maximum likelihood estimation, the match of two descriptions (D_L, D_R of the first and second image) h_b should have the maximal conditional probability among all possible matches $h_1, h_2, \ldots, h_n$.

$h_b = \max_I P(h_i/D_L, D_R)$ and using Bayes' formula $P(h_i/D_L, D_R) = P(D_L, D_R/h_i)\,P(h_i)/P(D_L, D_R)$

Secondly, an informed tree search method is employed. It utilizes the similarity of the tree nodes to guide the search.

As an example for the extraction of structural descriptions, the Foerstner operator is applied for the point extraction to increase matching accuracy (Wang, 1994). Image edges are extracted by means of one-dimensional operator and the lines are extracted from the edges with the method of mathematical morphological transformation (Wang, 1994). A boundary line based region growing has been developed for the region extraction (Wang, 1992, 1994). The flowchart of structural matching of two images is shown in Fig. 6

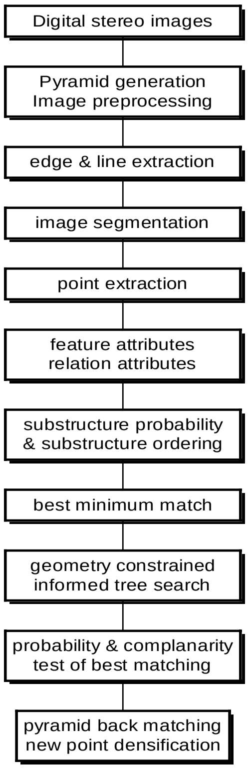


Fig.6. Flowchart of structural matching for two images.

4. Commercial Systems

1. The MATCH-T program (Ackermann and Krzystek, 1995) operates with pairs of digital images the orientation of which is supposed to have been established. The basic procedural steps are quickly listed: Firstly, the photo-pairs are normalized (on the basis of the given orientation parameters), in order to subsequently exploit the epipolar geometry. Image pyramids are formed, with 8 or 9 levels, together with complete feature extraction in all pyramid levels. The basic feature extraction on all levels utilizes the Förstner operator. On each level, feature matching is performed, made efficient by epipolar constraints and bounded search spaces. The matched image points are processed analytically to 3D model (terrain) points, from which a finite element fitting is derived, in tiles, representing the respective DTM. Each level of the pyramid gives a better approximation of the DTM. As initial approximation, a horizontal plane is sufficient to start from. The final DTM is generated as a rectangular grid, defined by the finite elements. Normally, an interactive editing phase concludes the DTM generation. The method is originally based on the assumption of smooth terrain. Refinements are introduced by additional break-lines (and excluded areas) which are pre-determined interactively. The feature matching can, optionally, be supplemented by least squares matching. A characteristic item of the system is high redundancy. Normally, much more terrain points are captured (e.g. 500,00 per model) than in conventional analytical DTM capture.
2. In the Socet Set software of Helava Associates Inc., the method for determing automatically the elevations in an elevation model is based on hierarchical correlation. The basic premise of this method is to use a hierarchy of image resolutions along with the hierarchy of elevation density measurements to arrive at a final grid of elevation points. In the new method that is being used in the Socet set is called AATE (Adaptive Automatic Terrain Extraction). Until now with the non adaptive ATE method, the user had to set the correlation parameters and create extraction subregions and in addition to that it was required that the terrain type does not change within each subregion, that a unique set of correlation parameters can be applied. So a stereo model should be partitioned into subregions and these subregions then be merged into a single DTM afterwards. In the AATE method allows to produce the total DTM in one execution and plus it can support more than two images.That means that in AATE optimal pairs are selected for imager correlation for different areas. The best pair for correlation is defined as the image pair which has the least image relief distortions and occlusions (the best 2 images based on terrain slope normal angle for each DTM). At these subregions uses an intial pass to collect stats but its results are not used and the strategy varies according to local terrain type, but it is not clear within what neighborhood it is estimated . After the selection of optimal image pairs performs epipolar resampling on the fly and as final result it can generate both grid and TIN DTMs. The program interface does not let the user

to define all the parameters for matching and unfortunately not all of them that are being used in the strategy files are explained clear

5. Sampling modes - Different point distributions

In matching, either predefined points (e.g. signalized points) or freely defined points must be measured. Optimal are the points which permit a precise and reliable matching. In surface measurement a requirement for optimality is the selection of points which represent well the surface to be measured, e.g. characteristic points. The best case is to measure just these points that are sufficient to describe the surface. For general, free form surfaces it is very difficult to automatically find such points just from the image content.

In the photogrammetric practice, the following sampling modes have been used:

- Grid sampling

  Usually, a regular grid is defined in object space, i.e. the planimetric X,Y object coordinates are fixed and at these positions the Z is measured. A variation of this sampling mode is mode is to define a grid in the digital image space (in one of the images) and determine the corresponding points in the remaining images.

- Progressive sampling (Makarovic, 1973)

  This is a variation of the grid sampling. Starting from a fundamental grid spacing, the grid is sequentially densified (usually by a factor of two) at the positions where the object surface deviates from a plane. To decide whether a grid mesh deviates from a plane, the already calculated heights at the neighbourhood of the mesh are used. This is essentially a hierarchical approach. Starting with the fundamental grid at the top level of the pyramid, the grid densification can be applied either in the (0^{th}) level or in intermediate levels of the pyramid and the approximations for these new grid points can be derived either by starting from the top level (lowest resolution) or from a previous level with simultaneous exploitation of the already known heights at neighbouring points. This grid densification is being used in Match-T.

- Composite sampling (Makarovic, 1979)

  This a two stage approach. First, progressive sampling is applied and then characteristic lines (breaklines etc.) and discrete points are measured. This is the most common approach in manual photogrammetric practice, whereby the progressive sampling stage is usually replaced by a regular grid sampling or profiling.

- Arbitrarily distributed points

  The measured points do not lie on a grid but are selected such that the automatic measurement procedure is expected to succeed. Such methods are described in Baltsavias (1988), Hannah (1988), Hahn and Foerstner (1988).

Commercial systems select match points in an image grid, object grid, or point features (small image patches) with good texture. The first two approaches may lead to matching even at areas with no texture, obviously leading to wrong results or less dense match points, and thus worse surface modelling.

6. Ways of providing approximate values

6.1 Hierarchical techniques

Good approximations are necessary to **reduce the number of false matches and multiple solutions**. To overcome this problem different strategies can be used like the multiresolution approach of Hannah (1988). Different levels of resolution of the image are being created and the match is usually applied in the highest level of the pyramid in order to provide a initial approximation. This technique can be found under the terms "hierarchical matching", "image pyramids", "multiresolution representation", "frequency analysis", "image compression and coding", "image decomposition", "image transmission", "scale-space representation", "multigrid methods". The pyramid levels are constructed by reducing the resolution. The reduction of the number of pixels is justified because the resolution reduction is based on low-pass filtering. The more coarsely resolving pixels of each level are computed from a small neighbourhood (window) of the previous (lower) level, whereby the pixels within the neighbourhood may get different weights. The goal of any pyramid structure is the ordering of the image features according to their importance (size, length, contrast, etc.). Important features should appear at the top levels of the pyramid where the amount of data is small. The match features in each pyramid level are either enforced to be the same or are independently selected. Image pyramids although have the main disadvantage that texture may disappear in upper levels.

But unfortunately constraints imposed by results obtained at the lower resolution can also lead to generation of erroneous elevation data in some circumstances. For example if the terrain is spotted with individual trees and, buildings or other structures, the correlation process, when performed on the low resolution images, will include approximations to the heights of these features in the resulting elevation data. The influence of these erroneous elevations will be carried from one step to the next in the hierarchical procedure. The result of this “noisy‘‘ elevation data can be minimized by using a “full resolution ‘‘correlation or the “Iterative Orthophoto Refinement‘‘ (Norvelle, 1993).

– If a correlation method allows it, the user can construct ‘‘starting profiles‘‘ in a full resolution stereo model to initially constrain the correlation process. One profile would, for example, be drawn across the top of the stereo model and one along the left edge. They would be drawn through the undesirable vertical features and thus constrain the correlation process to treatment of terrain points only.
– The IOR method is based on the premise, that given an accurate DEM and known orientation and position data for a stereo pair, orthophoto images made of both the left and right image of the

original stereo images should be identical. If they are not, the mismatches are due to errors in the DEM. These mismatches can be converted to equivalent elevation errors. The elevation errors are then subtracted from the current DEM. The initial DEM can be derived in the 0^{th} level of the pyramid and because less information is available this DEM would be coarse enough. The improved DTM is used to generate new orthophoto images and the correction process is repeated. The IOR procedure is iterated until mismatches are eliminated and the DTM is corrected in the process.

6.2 Epipolar Geometry

Approximate values in the matching of features can be successfully derived with epipolar images. Epipolar images are usually referred to as normalized images. The key idea is to transform the original image pair into new digital images in which the rows become epipolar lines. Any matching procedure for determining homologous points can be restricted to the rows of the images which implies that the required matching operations and computation times are significantly reduced.

Zhang (1986) in the system SODAMS uses the results of a previous epipolar-line correlation to predict in the next epipolar-line where the matching points would be and furthermore a group of matching results can be used to predict the position of the matching points. Therefore, the results of this method rely on the processing results.

Furthermore, the edge matching method of Zhang (2000) uses the epipolar geometry and reduces the search space more by defining an epipolar band with the following way: The two end points of the line segment in one image generates two epipolar lines in the other image and with the approximate height information derived from DSM data, an epipolar band is defined. Any line included in this band is a possible candidate, if it intersects the two epipolar lines within this band.

Another idea, which combines multiple images and their epipolar lines, was presented by Maas (1991) and was applied in the determination of TPCTF. This application is based on discrete points but can reduce the number of multiple solutions also in aerial images. The whole technique is based on that the homologous points of the first image lie along the epipolar lines of the second image and also along the epipolar line of the third image. While the strict solution with three or more images is based on a combinatorics algorithm, the reduced solution is based on probability measures for potential correspondences, which are defined by the number of images the candidate can be traced through. This is implemented in a recursive manner in a way that the longest traces (i.e. the points which can be matched successfully in the largest number of images) are accepted first and in case of ambiguities only a candidate with a trace, which is significantly longer than the traces of the other candidates, is accepted. The flow chart in Fig.7 elucidates this principle.

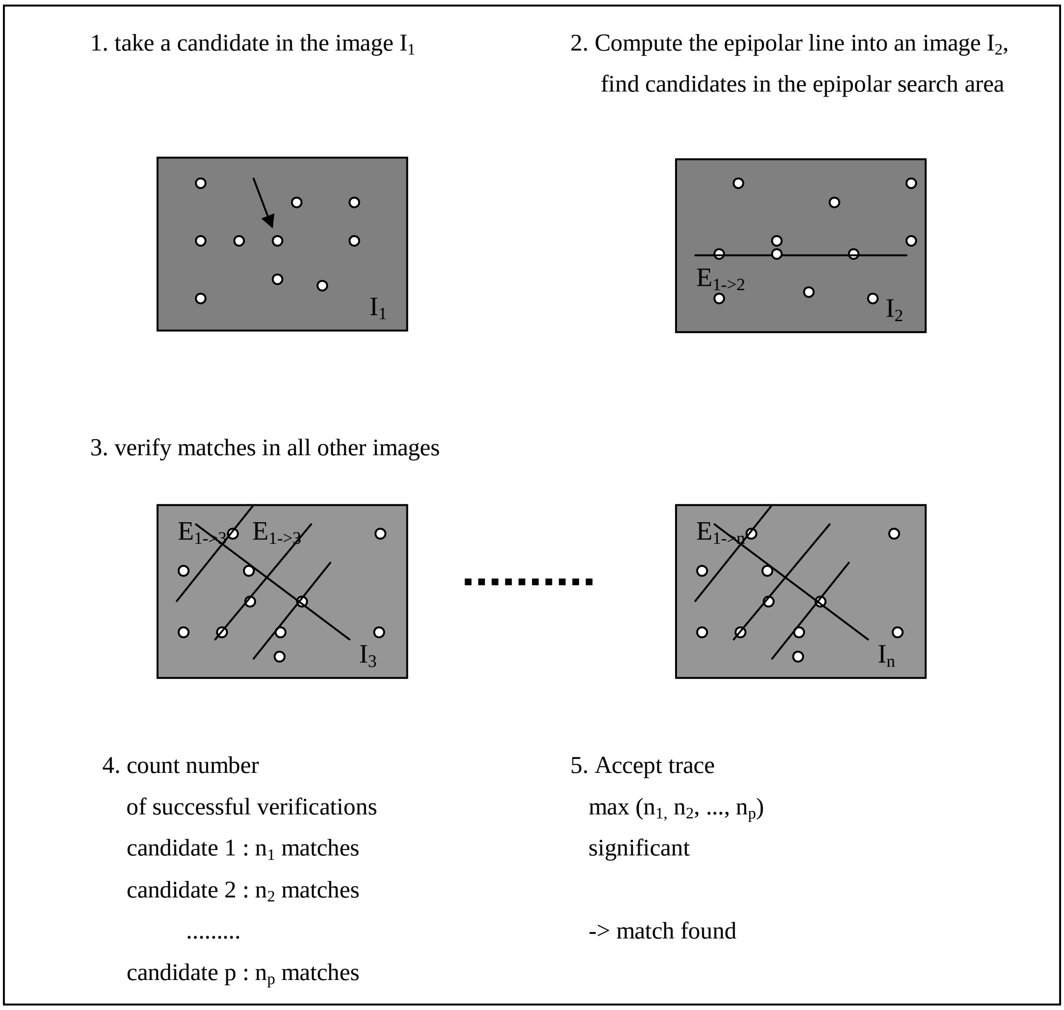


Fig. 7. Computation scheme for the automatic establishment of correspondences via epipolar line intersection

For a point P‘ in the first image the epipolar line in the second image is computed. Then, for all candidates on this epipolar line, the assumption of a correct match is checked by intersections with the epipolar line of P‘ in all other images. Finally, the number of successful verifications is counted and the match is accepted if the number of successful matches for a candidate is large enough and significantly larger than the number for all other candidates. In this way, the search space can be reduced through a multiphoto implementation. In cases of course of occlusions where points can not be defined in one or more pictures the technique does not perform equally well or may fail.

6.3 Neighbouring and seed points

Derivation of approximate values can be achieved by using neighbouring points. This method is based upon the following principle. For a few starting points approximations are either known or can be measured manually. The known points can be control points. These points have known object, image and pixel coordinates. Then points close to the known ones can get approximations from their known neighbors and matching results of these points can be used to approximate other points. Assumptions about the surface form are generally made: a) the surface has locally the same height, b) the surface height changes linearly in the neighbourhood of a point, c) the surface height variation in the neighborhood of a point can be modeled by a higher order polynomial. It is clear that any of these assumptions can be valid in a rather small neighbourhood of a point. Therefore this approximation method requires points that are dense and evenly distributed.

Approximations for a few corresponding points can be manually selected and a least squares algorithm is applied to find their exact location in the images. To define the regions between the different seed points a Voronoi tesselation is being done in the template image. The search strategy is as following: the process starts from one seed point, makes a horizontal shift in the template and in the search image and the LSM is applied in the shifted location. If the quality of the match is good the process continues until the boundaries are reached, if not satisfactory the algorithm works adaptively by changing the parameters (e.g. smaller shift, bigger size of patch). The criterion for a match whether is good or not., depends on the σ a priori.

Summarizing the requirements for the above methods to be successful are: dense points, points with good signal content, smooth variation of the object surface. And since these requirements are rarely met, this approximation method fails in a lot of cases.

6.5 Methods in Structural Matching

Especially in structural matching (Wang 1998), in order to reduce the search time further measures are integrated into the search method. They include substructure concept, primitive ordering, best minimum match and geometrical constraints. With the substructure concept the primitives and relations of a structural description are integrated based on their connectivity so that the primitive volume can be decreased significantly. In the case of structural matching of two images, the structural description can be reorganized. This is possible because in an image all the points, lines and regions are interconnected with certain relations and the attributes of a point primitive can be calculated more accurately and reliably than those of a line or region. Each point with its associated lines and regions forms a substructure and the search for matched primitives becomes a search for matched substructures.

In primitive ordering the primitives are reordered according to their type, weights and similarity so that the primitives, which are more likely to lead to a solution, are tried earlier than the others by the search method. Since the goal in structural matching is to find the best match as quickly as possible, primitive ordering is very useful in getting a best minimum match quickly.

A best minimum match refers to a part of matched features, which can be used to derive some geometrical relationships of two structural descriptions. If certain mathematical relations exist between two structural descriptions (i.e. affine transformation), they can be used as geometric constraints to guide the search paths.

7. Ways of providing reliable and quantitative quality criteria for the matching results and detection of blunders

7.1 Quality criteria

Quality control refers to measures that permit on one hand, the evaluation of the performance of the matching algorithm and on the other hand the detection and exclusion of errors due to insufficient physical modelling or signal content.

The main classes of errors are:

- Weak texture
- Multiple solutions
- Occlusions
- Surface discontinuities
- Large perspective differences
- Radiometric distortions (noise)
- Global noise from high frequences
- Error in the sensor orientation
- Errors in the determination of the pixel to image coordinates

The quality can be described by local and global measures listed in Table 1. The standard deviations of the estimated shifts are the decisive measures for high precision. As a criterion for the acceptance of a match it only has a value if the result is insensitive to undetected or unmodelled systematic effects, such as radiometric or geometric distortions. In the case of multiple solutions from the correlation, the best way to check whether the matching is accurate or not, is to match the template with patches in more than two images. The success rate is increased because of the mutual relation and the support among the image patches.

Table 1. Measures that describe the quality.

N^0	Name	Measure	Type	Causing effects
1	Precision	Standard deviation	Local	Random errors
2	Sensitivity	Bias	Local	Systematic errors
3	Convergence	Pull-in range, rate of convergence	Local, economic	Approximate values
4	Reliability	Probability of false match	Global	Approximate values

Precision

Precision of object location is determined by three parameters: the number of pixels, the variance and the covariance of the image gradient. The variance of estimated shifts can be found from the solution of normal equations, using least squares techniques after linearizing an equation for a two-dimensional model (Foerstner, 1982). Generally the precision is indicated by the covariance matrix and the RMS from comparison of the matching results with known. Apart from that how the precision in image space translates into object space precision depends on the imaging configuration (image scale, base to height ratio, rotation angles) and the precision of the interior and exterior orientation.

Sensitivity

The effect of unmodelled geometric or radiometric distortions in the correlation may cause systematic effects and reduce the total accuracy of the match.

Convergence

The convergence is also a local measure, which might be described by the pull-in range or the rate of the convergence. These quality measures mainly depend on the texture of the object and are discussed in detail using a joint mathematical model. Problems in convergence can reflect the quality of the approximations. Problematic behavior is expressed as slow convergence, oscillations or divergence. Apart from the quality of the approximations the convergence behavior depends also on the signal content of the patches , grey level derivatives and the application of filtering. Cases of weak or impossible correlation show up in a very slow convergence of the solution vector. Here the problem stems from the fact that the data sets do not have sufficient signal content to allow for a reasonable accurate correlation.

Reliability

In order to decrease the number of false matches the probability of correct matches can be increased. Based on the above, it has been shown that for low signal to noise ratios in the least squares matching (Foerstner, 1984), the reliability of correlation can be increased significantly by sharpening the peak of the correlation function. Also Baarda's method of data snooping can be applied. Local radiometric and geometric distortions between template and patch result systematic errors, which can be detected with this technique (see 6.2 below).

Another method to check the correctness of the match is by backmatching (starting from the target point that had been determined and finding the best match for it in the template). If the difference in the result from the backmatching is more than one pixel, that match is discarded as being unreliable (Hannah, 1988).

As mentioned above the multiple solutions can be reduced to one by implementing the matching procedure simultaneously in more than two images.

Occlusions increase with increasing base between the sensor stations, especially when the surface discontinuities are vertical to the base. Their effect could be decreased by choosing large dimensions for the template, but this violate the planar surface assumption especially at the position of surface discontinuities. Big occlusions can be detected by a blunder test in the upper levels of the image pyramids and smaller occlusions in lower levels.

7.2 Detection of Blunders

This is the most important aspect of the quality control and determines to a large extent whether automatic image matching algorithms can successfully replace manual measurements. Tests for detection of blunders use different criteria or combinations thereof and compare them against some critical values. The criteria that are being used for blunder detection are:

a) Correlation coefficient of the grey level matrices. It can be small even for the correct match points because of high noise level, occlusions, perspective differences etc.
b) A posteriori standard deviation of unit weight σ_o in methods that are using least squares.
c) Average strength of the x,y grey level derivatives
d) Standard deviation of x and y shifts
e) Number of iterations
f) Alterations of the x,y pixel coordinates from their approximate values
g) x, y parallax (these values should be similar in neighbouring points)
h) standard deviations of the object coordinates

Generally the aim is to detect corresponding pixels with very different grey levels and exclude them from the solution. If the pixels are really corresponding, their grey level difference must be due to noise or occlusions (unmodelled errors). These effects causing radiometric differences are usually local. Data snooping is a statistical test procedure, which uses the criteria

$$w_i = \frac{|v_i|}{\hat{\sigma}_{v_i}}$$, i = 1,...,M number of grey level observation

The test criteria w_i are compared to a critical value c, which can be taken from the Tau distribution or for the Student's distribution tables. If a test criterion exceeds the critical value, the corresponding observation is considered as gross erroneous and can be excluded (Gruen, 1978). If an observation is excluded, then the residuals of the remaining observations to be tested , have to be recomputed. The new a posteriori variance can be computed by

$\tilde{\hat{\sigma}}_o{}^2 = \frac{r - w_i{}^2}{r-1} \hat{\sigma}_o{}^2$, where r the redundancy and i index of deleted observation.

Data snooping is applied after a first convergence of the iterations, and then the ’‘clean‘‘ data set is iterated again, if errors were detected. At the end of each iteration the test is applied again. This method works nicely when the local distortions are small. If the distortions cover a larger part of the patch or their size (grey value differences) is big, then they will cause an inflation of the standard error and no observation might be rejected.

Another technique to overcome the problem of blunders in the data is to remove a point from the data if the residual of it was outside three standard deviations of the mean.

Robust estimators are being used widely because they are less sensitive to the presence of gross errors and outliers. Research relating to robust estimation has been carried out primarily by statisticians (Hampel, 1973, Hogg, 1974, Huber, 1964, 1972, 1981) and photogrammetrists (Kubik et al., 1985, 1987, 1988, Stefanovic, 1985, Veress and Huang, 1987)
A possible numerical solution is achieved by an iterative least squares weighted solution. Instead of the normal LS solution,

$$\sum_i r_i{}^2 = \min$$

the minimization is made using an alternative function of the residuals:

$$\sum_i fn(r_i) = \min$$

The estimation procedure is designed to provide a screening among the observations, taking a priori into account that not all of them should be given the same role in determining the solution. This does not happen to normal least squares matching, where all observations equally contribute, on the basis of their a-priori variance, to the solution.. The amount of the weight change is generally determined on the basis of the residual of the observation. To help decrease the influence of gross errors, if a point is identified as a gross error at iteration a, it is not used for statistical calculation necessary for the robust weighting of iteration (a+1). But the gross errors at iteration i will be reexamined using the resulting statistical information. That is, if no statistical information is known prior to the matching process, gross errors may influence the first iteration (Pilgrim, 1996).
The following example is based on the above

Initial weights:
$P_{i,o} = 1$, i= 1...n index of observations

Weighted least squares solution in iteration i

$$\sum (P_a r^2) \rightarrow \min, a = 0,1,2...$$

Computation of new weights:

$$P_{i,a} = \frac{\varphi(r_i)}{r_i^{\ 2} + \varepsilon}, \varepsilon \text{ small constant}$$

For the LS method, $\sum_i |r_i| = \min$, the weights are equal to $P = \frac{1}{|r| + \varepsilon}$. The iterations stop when no significant changes in the results are observed.

Especially in structural matching, the correlation coefficient is an important measure for quality evaluation of the matching of two point primitives. Because no orientation information is available for structural matching, the correlation method must be rotation-independent. The correlation coefficient is based on the gradient direction. Furthermore, the correctness of the match can be checked by using the magnitude of its probability and the geometric conditions. For two perspective images there is a coplanarity condition. After enforcing the coplanarity condition onto the matched points, the parameters of the coplanar equation and a variance factor can be determined. This variance can be used to judge the correctness of the matching result because the a-priori variance of the points is available from the point extraction algorithm.

8. Conclusions

There are so many problems associated with image matching in digital photogrammetry:

1. Radiometric problems: resolution, reflectance, illumination, lab processing noise, digital camera noise.
2. Geometric problems: relief displacement and occluded areas, projective deformation, scale variation
3. Textural problems: featureless surface, repetitive texture, ambiguous levels such as tree top and ground below them, thin objects.

And until nowadays no algorithm, no matching system has the capabilities to overcome the matching difficulties presented above. Each algorithm can work and give good results in some areas, but in general there is a lack of adaptive algorithms. The matching results do not mainly depend on the main matching algorithm, but are enforced from the several strategies that are being used to derive good approximations, to detect and eliminate the wrong matches even in difficult cases. A combination of simple matching components can achieve good results depending on the case.

Also in commercial systems this “adaptivity” is misused. Numerous parameters exist in the program codes without fully being explained and without logical base. The user many times does not have a clear idea, what each parameter does and the use of it. And then for sure is more difficult to run the program and produce reliable results.

In an optimal image matching technique all individual cases must be taken into account, researched, and strategies depending on this pre-processing should be implemented. Then we could start not from ‘‘chaos’’ but from a more limited area. We hope that investigations in the future would try not to focus their interest in a new matching algorithm but in a sophisticated technique of combining several methods under one umbrella and characterized as adaptive.

References

Abbasi-Dezfouli, M., Freeman, T.G., 1994: Patch matching in stereo images based on shape. Proceedings of ISPRS Commission III Symposium on Spatial Information from Digital Photogrammetry and Computer Vision. September 5-9, 1994. Vol. 30 :1-7

Ackermann, F., 1984: Digital image correlation- performance and potential application in photogrammetry. Photogrammetric Record 11: (64) 429-439 1984

Bailard, C., et al, 1998: Extraction and textural characterization of above ground areas from aerial stereo pairs: a quality assessment. ISPRS Journal of Photogrammetry and Remote Sensing 53: 130-141.

Baltsavias, E., 1991: Multiphoto Geometrically Constrained Matching. Phd. Thesis,Mitteilungen No. 49, Institut fuer Geodaesie und Photogrammetrie, ETHZ, Switzerland

Barrodale, I., Young, A., 1966: Algorithms for best L_1 and L_∞ linear approximations on a discrete set. Numerische Mathematik, 8: 295-306

Bellone, T., et al, 1996: Robust procedures for data preprocessing, testing and archiving. International Archives of Photogrammetry and Remote Sensing. Vol. XXXI, Part B1, Vienna 1996

Bobick, A.F., 1999: Large occlusion stereo. International Journal of Computer Vision. 33 (3) 181-200 Sep 1999.

Boesemann, W., 1994: Geometric Modells in Object based Multi Image Matching. Proceedings of ISPRS Commission III Symposium on Spatial Information from Digital Photogrammetry and Computer Vision. September 5-9, 1994. Vol. 30 :61-68.

Brunn, A. Et al , 1998: A hybrid concept for 3D building acquisition. ISPRS Journal of Photogrammetry and Remote Sensing 53: 119-129.

Calitz, M.-F., Ruether, H., 1996: Least absolute deviation (LAD) image matching. ISPRS Journal of Photogrammetry and Remote Sensing 52: 160-168.

Chung, KL.,1997: A fast algorithm for stereo matching. Information Processing Letters 63: (2) 57-61 JUL 28 1997

Cochran, S.-D., 1995: Adaptive vergence for the stereo matching of oblique imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 50(4):21-28

D'Apuzzo N., 1998: Automated Photogrammetric Measurement of Human Faces. International Archives of Photogrammetry and Remote Sensing, Hakodate, Japan, 1998,Vol. 32 (B5), pp 402-407.

Ebner, H., Fritsch, D., Gillessen, W., Heipke, C., 1987: Integration von Bildzuordnung und Objectrekonstruktion innerhalb der Digitalen Photogrammetrie. Bildmessung und Luftbildwesen: Zeitschrift fuer Photogrammetrie und Fernerkundung, 55 (5):194-203

Foerstner, W., 1982: On the Geometric Precision of digital Correlation, International archives of Photogrammetry, Vol. 24 , Helsinki 1982:176 – 189.

Fiset, R., Cavayas, F., Mouchot, M. et al., 1998: Map-image matching using a multi-layer perception: the case of road network. ISPRS Journal of Photogrammetry and Remote Sensing 53 (1998) 76-84.

Gruen, A., 1985: Adaptive Least Square Correlation : A powerful image matching technique. S Afr J of Photogrammetry, Remote Sensing and Cartography 14 (3), 1985 : 175-87.

Gruen, A., Agouris, P., 1994: Linear extraction by Least Squares template Matching Constrained By Internal Forces. Proceedings of ISPRS Commission III Symposium on Spatial Information from Digital Photogrammetry and Computer Vision. September 5-9, 1994. Vol. 30 :316-323

Hampel, F.R., 1973:Robust Estimation: a condensed partial survey. Z. Wahrscheinlichkeitstheorie Verw. Geb., 27: 87-104

Hannah, M.,J.,1988: Digital Stereo Image Matching Techniques. Proceedings of 16th ISPRSC. In ASPRS Vol.27/B3:280-293

Hannah, M.J., 1989: A system for digital stereo image matching., PERS (55) 12, 1765- 1770.

Heipke, C., 1996: Overview of Image Matching Techniques. Proceedings of the OEEPE- Workshop on Application of Digital Photogrammetric Workstations. Lausanne 4-6 March 1996.

Helava, U.V.,1988: Object-Space Least-Squares Correlation.PERS June 1988, 6(1)

Hogg, R.V., 1974: Adaptive robust procedures: a prtial review and some suggestions for future applications and theory. Journal Am. Stat. Assoc., 69 (348): 909-923

Horiuchi, T., Toraichi, K., 1994: Extended Relaxation matching method, that includes DP matching method. Systems and Computers in Japan. 25: (5) 21-27 May 1994

Huber, P.J., 1964: Robust estmation of a location parameter. Ann. Math. Stat., 36: 1753 – 1758.

Huber, P.J., 1972: Robust statistics: a review. Ann. Math. Stat., 43: 1041- 1067.

Huber, P.J., 1981: Robust statistics. Wiley, New York.

Hung, YP., et al., 1998: Multipass hierarchical stereo matching for generation of digital terrain models from aerial images. Machine Vision and Applications 10: (5-6) 280-291 Apr. 1998

Jokinen, O., Haggren, H., 1998: Statistical analysis of two 3-D registration and modeling strategies. ISPRS Journal of Photogrammetry and Remote Sensing 53 (1998) 320-341.

Kreiling, W., 1976: Automatische Erstellung von Hohenmodellen und Orthophotos durch digitale Korrelation, Dissertation, Institut fuer Photogrammetrie, Universitaet Karlsruhe.

Krupnik, A.,Schenk, T.,1997: Experiments with matching in the object space for triangulation. ISPRS Journal of Photogrammetry and Remote Sensing 52(1997) 160-168.

Kubik, K., Weng, W. And Frederiksen, P., 1985: Oh! Grosserrors! Aust. J. Geodesy Photogramm. Surveying, 42:1-18.

Kubik, K., Merchant, D., Schenk, T., 1987:Robust estimation in photogrammetry. PERS 53 (2): 167-169.

Krzystek ,P., 1992: Experimental accuracy analysis of automatically measured digital terrain models. Robust Computer Vision:quality of vision algorithms/ Förstner; Ruwiedel(ed.).- Karlsruhe: Wichmann, 1992.

Laing, R., 1995 : Aquisition Methods for 3D- Modelling. Second course in Digital photogrammetry-Institut fuer Photogrammetrie , Universitaet Bonn , February 6-10, 1995.

Lee, JD., et al.,1998: A new algorithm for two –dimensional object inspection using string matching. Mathematical and Computer Modeling 27 (1) 101-116 Jan 1998

Lemmens, M., J., P., M., 1988: A Survey on Stereo Matching Techniques. Proceedings of 16th ISPRSC. In ASPRS Vol.27/B8:11-23

Li, M.X.,1989: Hierarchical multipoint matching with simultaneous detection and location of breaklines. Ph.D thesis, Royal Institut of Technology, Department of Photogrammetry, Stockholm.

Maas, H.-G., 1996: Automatic DEM generation by multi-image feature based matching. International Archives of Photogrammetry and Remote Sensing. Vol. XXXI, Part B3, Vienna 1996.

Moravec, H.P., 1977: Towards Automatic Visual Obstacle Avoidance. Proc. of 5th International Joint Conference of Artificial Intelligence. MIT, Cambridge, MA, August 1977: 584.

Moravec, H.P., 1979: Visual mapping by a Robot Rover. . Proc. of 6th International Joint Conference of Artificial Intelligence. Tokyo, Japan, 1979: 598-600.

Nevatia, R., 1996: Matching in 2-D and 3-D. International Archives of Photogrammetry and Remote Sensing. Vol. XXXI, Part B3, Vienna 1996

Norvelle, F.R., 1992: Stereo Correlation: Window Shaping and DEM Corrections. PERS, Vol 58, No.1, January 1992:111-115.

Norvelle, F.R., 1993: Alternatives to hierarchical techniques in stereo correlation. Proceedings of SPIE, 14-15 April 1993, Orlando, Florida, Vol. 1944 :176-184

Pilgrim, L., 1996: Robust estimation applied to surface matching. ISPRS Journal of Photogrammetry and Remote Sensing 51(1996) 243-257.

Rauhala, U.A., 1987: Fast compiler positioning algorithms and techniques of array algebra in analytical and digital photogrammetry. Proceedings of the Intercommission Conference on Fast-Proccessing of Photogrammetric Data. Interlaken, June 2-4,1987,pp.156-178

Rosenholm, D., 1987: Multi-point matching using Least- Squares technique for evaluation of three-dimensional models. PERS, 53(6): 621-626

Rosenholm, D., 1988: Multipoint Matching along vertical lines in SPOT images. International Journal of Remote Sensing 9: (10-11) 1687-1703 Oct-Nov 1988

Schmutter, B., Doytsher, Y., 1990: DTM – Accuracy Estimates. 4th International Symposium on Spatial Data Handling 1990, Vol.2, Zurich, Switzerland: ,263-272

Scharp, J.V., Christensen, R.L., Gilman, W.L., Schulman, F.D., 1965: Automatic map compilation using digital techniques. PERS (31) 3, 223-239.

Schenk, T., Greenfeld, JS.,1989: Experiments with edge-based stereo matching. Photogrammetric Engineering and Remote Sensing. 55: (12) 1771-1777 Dec 1989

Wang, Y., 1998: Principles and applications of structural image matching. ISPRS Journal of Photogrammetry and Remote Sensing 53:154-165.

Wrobel, BP.,1987: Digitale Bildzuordnung durch facetten mit hilfe von objectraummodellen. Bildmessung und Luftbildwesen. Vol. 55, No 3,:93-101.

Wrobel, BP., 1991: Least-squares methods for surface reconstruction from images. ISPRS Journal of Photogrammetry and Remote Sensing 46 (2) 67-84 b.

Zhang, Z.,1989: A new approach of Epipolar-line Image Matching-Bridging mode, ACTA Geodetica et Cartographica Sinica (English version).

Zhang, Z.,1990: Extension of the concept of feature bridging mode method and global image matching. Proceedings of the Symposium Commission III of ISPRS, Vol. 28 ,3(2), May 20-25 1990:1106-1120.

Zhang, Z., Zhang, J., Wu, X., Zhang, H., 1992: Global Image Matching with Relaxation method, Proceedings of the International Colloqium on Photogrammetry, Remote Sensing and Geographic Information System,11-14 May, Wuhan, China:175-188

Zhang, Z., Zhang, J., Mingsheng, L., Zhang, L.,2000: Automatic Registration of multi-source Imagery based on Global Image Matching. PERS,Vol.66 No.5, May 2000: 625-629.

Zhang, C., Batsavias, E., 2000: Knowledge-based image analysis for 3D edge extraction and road reconstruction. IAPRS, Vol. XXXIII, Amsterdam 2000 (in print).

Zheng, Y.J.,1993: Inverse und schlecht gestellte Probleme in der photogrammetrischen Objectrekonstruction, DGK-C 390.