MULTISTAGE IMAGE MATCHING

Luisa Maria Gomez Pereira, Branko Makarovic
International Institute for Aerospace Survey
and Earth Sciences (ITC)
350, Boulevard 1945
7500 AA Enschede
Commission II

ABSTRACT

In the design of image matching techniques, accuracy, efficiency and reliability are the most significant influencing factors. To this end, a flexible four-stage process strategy is proposed, based on the principle "from coarse to fine", with adaptive capability incorporated.

In this paper, attention is given to the two intermediate stages: the coarse matching of strings of image primitives in conjunction with a hierarchical tree matching from coarse to fine. The conjugate images can be sampled and digitised directly in epipolar lines. The software is modular and suitable for off-line preprocessing and image matching.
The development reported here is part of a project to be continued.

I. INTRODUCTION

The impact of rapidly advancing computer technology on digital photogrammetry is especially pronounced in automatic image matching and related operations. One of the most significant applications of image matching is automatic modelling of the terrain relief. Thus it is essential to devise an effective overall process strategy and to develop software for automatic measurement of the parallaxes (image disparities) in stereo-images. Because the amount of image information involved is huge, the process has to be time efficient. Moreover, to attain sufficient quality, the accuracy and reliability of image matching are of paramount significance. A flexible multistage strategy of processing "from coarse to fine", with adaptive capability incorporated, tends to provide an adequate solution of the problem.
The strategy proposed here addresses four main sequential stages. The first stage predicts image match approximately by exploiting a priori knowledge about the geometry of a photogrammetric stereo-model. The second stage proceeds with course matching by means of the dynamic programming technique applied to low frequency (spatial) image information. The corrected parallax values are entered in the next matching stage. The third stage uses as the input image segments structured in several hierarchical levels. The coarse to fine matching proceeds down the branches of the hierarchical tree. The fourth stage is intended for fine matching when the required accuracy is very high. It uses the a priori information, including that gained in the preceding matching stages, in conjunction with a set of selectable algorithms for fine matching.
The first matching stage, concerning match prediction, is rather straight-foreward and is therefore considered here only marginally. Hence, attention has been focused on the two intermediate stages (2 and 3), which provide for sufficient "pull-in" and thus reliability. The attainable accuracy will presumably meet the requirements for the common applications of digital terrain models (DTM). The development of these two stages has not yet been fully completed. The initial software has been developed and tested [2];

some optimization, however, is an issue for further development. The last stage, pertaining to fine matching, is beyond the scope of this paper.
The paper concerns the concepts for a multistage strategy and for the individual matching stages. Most of these concepts have emerged in the course the authors' work since 1975, [3], [4], [5], [6]. An outline is presented of the procedures, and of the software and its testing. Moreover, consideration is also given to further development.

II. CONCEPTS FOR MULTISTAGE MATCHING

The concepts concern the overall process strategy, the geometry of a photogrammetric model, including the epipolar geometry, and the operations involved in the four main matching stages. In the follwing a review is presented of these concepts and their interrelationships.

1) Process strategy

The concepts concern the arrangement of the four main matching stages and their interrelationships (figure 1)

Stage 1: Match prediction

Stage 2: Coarse matching

Stage 3: Coarse to fine matching

Stage 4: Fine matching

Fig. 1: Sequential matching stages

In stages 2 and 3 their main phases (preprocess, matching and postprocess) form feedback loops which provide for optimization (figure 2).

Preprocess

Matching — feedback loops for optimizing

Postprocess

Fig. 2: Optimization in stages 2 and 3

Table 1 presents a list of the concepts for each of the four matching stages. The lists are open ended and liable to alterations.

2) Geometry of photogrammetric images and stereomodels

The basic geometric concepts can be found in photogrammetric textbooks and will therefore not be considered here. Implementation of the epipolar geometry in analytical stereoplotters for digitising of stereo-images is, however, of special significance. A simple geometric situation is attained when using only the rotation parameters for the relative orientation. Hence, the base B is parallel to the X-axis in virtual model space XYZ (figure 3).

STAGE	CONCEPTS	OUTPUT
Stage 1: Match prediction	Photogrammetric model geometry Epipolar geometry	Predicted parallax
Stage 2: Coarse matching	Low pass filter Edge extraction Edge description String matching (dynamic programming) Epipolar parallax profile Other related	Coarse epipolar parallax profiles
Stage 3: Coarse to fine matching	Hierarchical structure of image information Filters Epipolar strip Hierarchical triple-tree (of epipolar lines) Layout of target segments Hierarchical bynary tree (of image segments) Search segment size Edge extraction/description (as in stage 2) Matching techniques Adjacency Epipolar parallax profile Other related	Fine (raw) epipolar parallax profiles
Stage 4: Fine matching	Analysis of parallaxes from stage 3, Joint analysis of parallaxes and intensity of the corresponding images, Matching techniques Selection of the technique Quality control Editing Other	Super-fine (edited) epipolar parallax profiles

Table 1: List of concepts for multistage matching

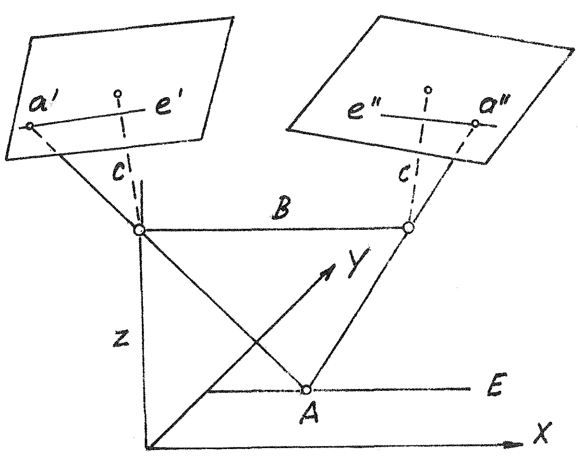


Fig. 3: Simplified epipolar geometry (X||B||E)

The epipolar lines e′ and e″ are images of the line E, which is parallel to X in the XY plane at Z=c (where c is the principal distance). The transformation of E into e′ and e″ is realised by the real-time program of the analytical stereoplotter.
For direct digitising of an image pair in epipolar format, a set of equidistant lines E is laid-out in the XY-plane such as to cover the entire model area. Then a special subroutine controls tracking of the sensor arrays along all conjugate line pairs e′ and e″. The sampled information is digitised and recorded for further use.

3) Individual matching stages

Each main stage of matching contains preprocessing, matching and eventually some postprocessing, and each of these comprises several operations. The concepts can be classified accordingly.

In stage 1 the 'base to height' ratio (B/Z) is essential. After relative orientation, the prediction of the parallax can be improved by correcting it for the effect of convergency ($Z\Delta\phi$).

In stage 2 the concepts address lowpass filtering (e.g., combined wiht Laplacian operator and by enhancing edges in the Y-direction), edge extraction by thresholding, edge description, matching of individual edge-pairs (and the acceptance criteria), matching of strings of edges, and formation of coarse epipolar parallax profiles.

In stage 3 the concepts for preprocessing concern building the hierarchical structure of image information, formation of the epipolar strips of images, lay-out of the target segments and definition of the corresponding search segments. The concepts for filtering, edge extraction and description can be as those in stage 2. The concepts for matching concern the navigation down the branches the 2D-hierarchical tree, and the technique for matching individual segment pairs. They also include the constraints emerging from the image geometry, and some neighbourhood relationships.

In stage 4 the concepts concern: a profound analysis of the a priori knowledge and of the information gained in the course of the process, the different techniques for fine matching, and the rules for selection of a specific technique (based on the outcome of the analysis). The concepts also address quality control.

III. MATCHING TECHNIQUE

Each matching stage is preceded by some preprocessing and it may be followed by some postprocessing. The latter can overlap with the preprocessing in the following stage. As stated above, the matching stages and their phases are interrelated. In the following, coinsideration is given to stages 2 and 3.

1) Stage 2: Coarse matching

Preprocessing for coarse matching is directly linked with the hierarchical structuring of the images in stage 3. A further level (n) is created above the upper-most level (n-1) of the hierarchical tree, built in stage 3. To this end, the images are filtered as in any other transition from a given level to the next higher one in stage 3 (vide stage 3). The filter is composed of a smoothing component (modified Gaussian) and an edge enhancement component (Laplacian). In both stages the images are arranged in epipolar strips (figure 4).

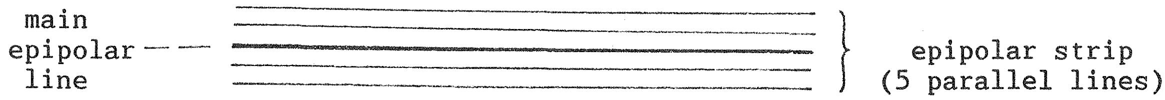


Fig. 4: Epipolar strip

In contrast to stage 3, the strips in stage 2 are not segmented (for a binary tree structure). The techniques for the extraction and description of edges (as primitives), however, are identical in both stages.
Hence, in stage 2 (level n) the images are represented by pixels

$$a_n = a_1 2^{(n-1)}$$

where a_1 is the pixel size of the original digital images in the lowest level 1. The higher levels (2,3,..., n-1) are built by applying successively the same composite filter (Laplacian of modified Gaussian), and by forming the epipolar strips composed of five (or more) parallel lines. The spacing between the adjacent lines is triplicated in each transition from a given to the next higher level. For the extraction of edges, thresholding is applied to the filtered images (in level n). During the process the threshold value can be automatically adjusted. Subsequently, a subroutine attributes to each edge a set of descriptors (vide stage 3). To limit the search range prior to matching, the maximum expected parallax correction (to the predicted one in stage 1) should be assessed. By means of a mirror stereoscope and a parallax bar, the approximate parallax difference can be measured between the highest and the lowest point in the stereo-image, and a safety margin is added.

Matching comprises two steps. First the similarity is assessed for each potentially conjugate edge-pair within the search range. To this end, the corresponding edge descriptors are compared. These descriptors can be weighted according to their impact on the assessment of similarity. Edge pairs with opposite polarity (e.g. signs of the gradient or of the second difference) are rejected. The similarity in other descriptors is tangeble and thus quantifieble by the figures of merit. A subroutine measures the degree of similarity for each descriptor separately, and converts it, via a look-up table, into the corresponding figure of merit. If the merit is less than the value of the corresponding acceptance threshold, the edge-pair is rejected. Otherwise an additional composite similarity is assessed, e.g., the weighted mean (S) of the similarity estimates for individual descriptors. These mean values (S) are compared against the corresponding acceptance threshold.
The second step concerns matching a string of potentially conjugate edge-pairs accepted in the first stage. A suitable technique for string-matching is dynamic programming with constraints [1]. To this end, the composite similarity estimates (S) are arranged in a matrix (figure 5).

The best match for a string of edge-pairs is obtained by the path of the total maximum similarity ($\Sigma S \rightarrow \max$). The computation of the potential paths is by the Viterbi algorithm [1]. The geometric constraints restrict the number of variants in each state (Sij) to the angular segment (quadrant) between the positive directions of both epipolar lines. After matching the conjugate strings of edges, the coarse epipolar parallaxes are known for all of the edge-pairs. Together they form a coarse epipolar parallax profile which serves as the input for the matching stage 3.

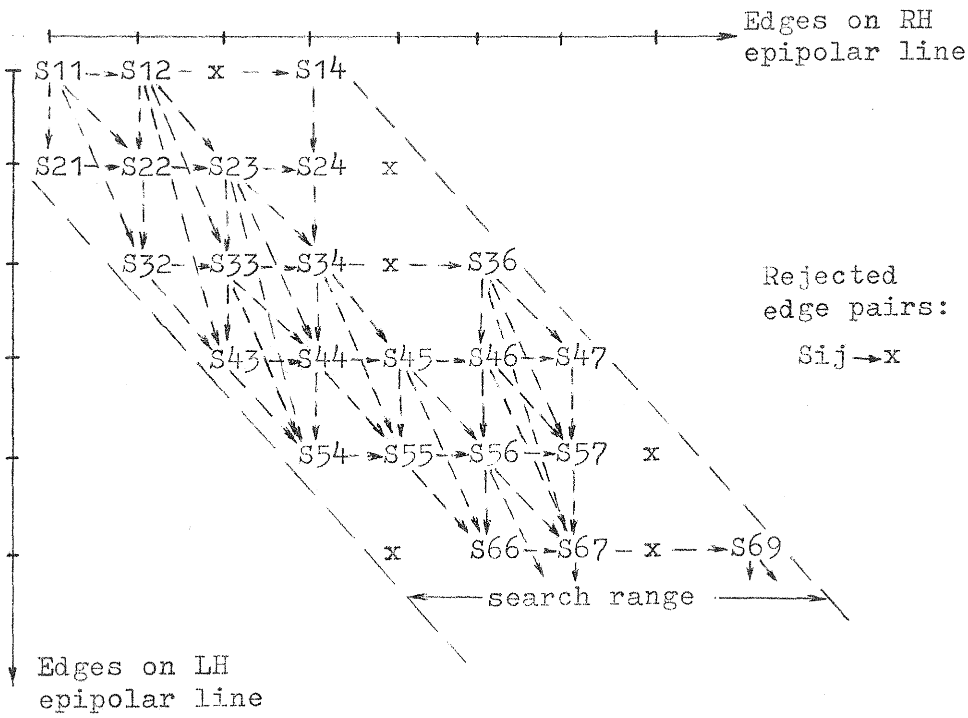
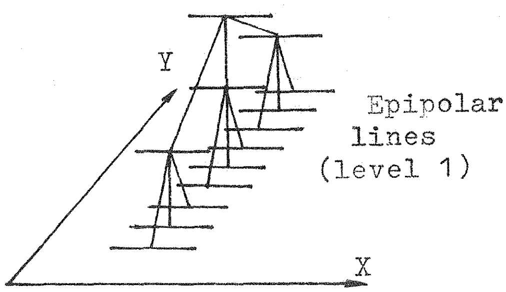


Fig. 5: Search for maximum similarity path

2) Stage 3: Coarse to fine matching

The core in this matching stage is the hierarchical tree structure of the images, which is built in the preprocessing stage. The preparation concerns the coarse parallaxes from stage 2, required for the lay-out of the search segments in the upper-most hierarchical level (n-1).

Hierarchical tree
The hierarchical tree is two-dimensional [3]. In the Y-direction the epipolar lines are arranged in a triple-tree (figure 6a).



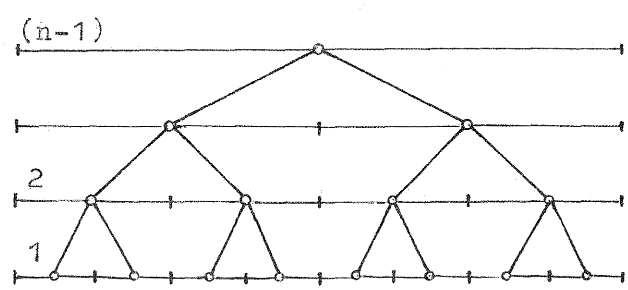


Fig. 6a: Triple tree of epipolar lines Fig. 6b: Binary tree of segments

Actually to each epipolar line a strip of four parallel lines is assigned (figure 4). The target segments along epipolar lines are arranged in a binary hierarchical tree (figure 6b). The tree of the target segments is symmetric. The conjugate search segments, however, have to be off-set from the positions of the targets for the amount of the corresponding parallax corrections (determined by matching in the next higher level). To minimize the search range in the upper-most level (n-1), the initial off-sets (coarse parallax corrections) for the midpoints of all search segments are determined by interpolation in the coarse parallax profiles from stage 2. The lenght of a search segment in the level (n-1) is equal to the length of the corresponding target segment, extended on each side for the required search range. The latter is, however, defined by the inaccuracy of the initial off-set (i.e., interpolated parallax correction from level n). If the maximum matching error in level n is one pixel (a_n), then the search range in level (n-1) should be two pixels ($a_n = 2a_{n-1}$).

In the lower hierarchical levels (n-2, ..., 1) the arrangement of the search segments is similar. Laying-out search segments and matching takes place alternatively from higher to lower levels. After completing matching in level i, the off-sets of the search segments are updated. Thus the midpoints of the search segments in level i-1 are defined, and the segments (half length of those in level i) are set out with a half length to the right and a half length to the left of the corresponding midpoints. The uniform structure of the segments and of the pixels inside the segments permits application of common algorithms for processing in all levels.

Matching

Matching proceeds in the up-down sequence in the hierarchical tree. Within each level the sequence is consistent, e.g., from upper-left corner along the epipolar strips, to the lower-right corner. Targets which do not contain significant edges are bypassed. The correponding parallax correction values are assessed from the neighbourhood (by interpolation) before matching proceeds in the next lower level. As in stage 2, matching concerns first the single, potentially conjugate edge-pairs, and then the strings of the accepted pairs. This is applied to each pair of the target and search segments in an ordered sequence. Subsequenctly, the two conjugate hierarchical trees should be matched as a whole. This is attained by accumulating the parallax corrections (off-sets, from coarse to fine), down the branches of the tree, including the lowest level. In the following, coinsideration is given to both matching single pairs of the target and search segments, and to the hierarchical tree matching.

a) Matching segment pairs

In each hierarchical level, the conjugate pairs of image segments are matched in an orderly sequence and by using uniform algorithms. The latter should be robust, i.e., insensitive to small discrepancies in geometry and intensity of conjugate images. The required accuracy is not extremely high. Thus matching of edges individually and of their strings seems feasible, though some other algorithms would also suit. The technique is as in stage 2 with the difference that strings are fragmented into segments. Figure 7 shows the segments for a pair of conjugate epipolar lines and the corresponding search range.

After a string is matched, the mean parallax correction value is computed for all matched edges within the segment. This mean value is then used further for the hierarchical tree matching [3].

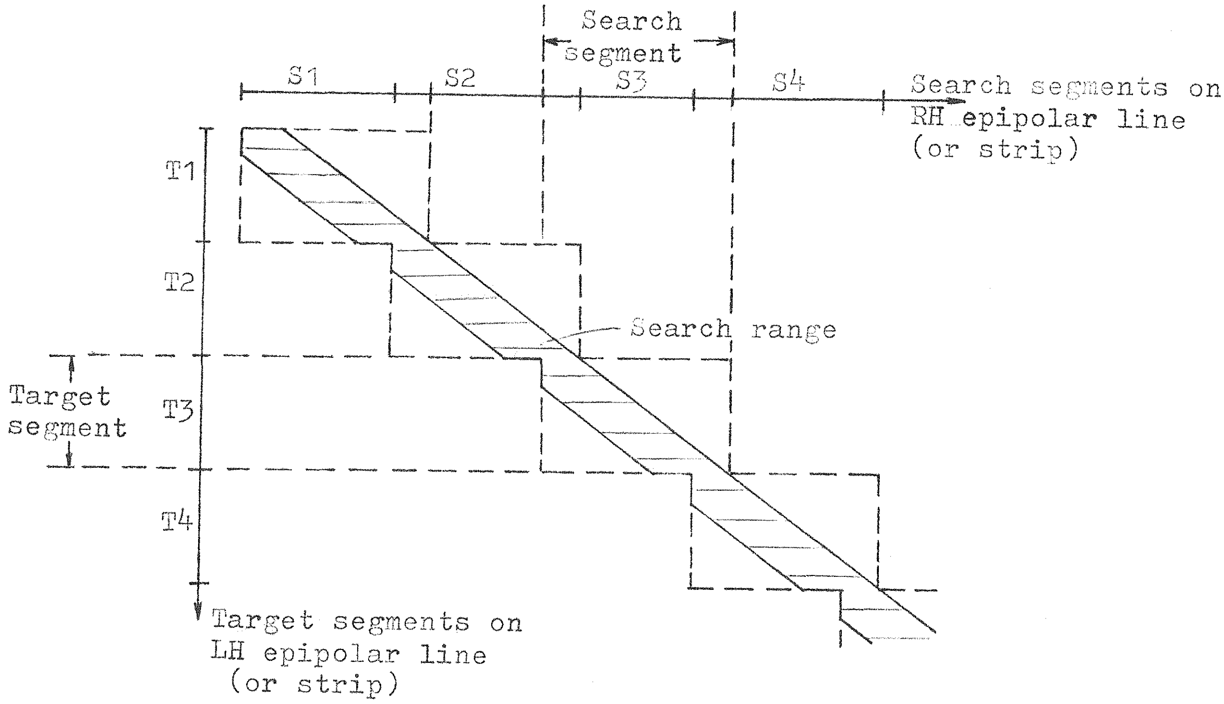


Fig. 7: Segmented string matching

b) Hierarchical tree matching

The tree matching uses as input the mean parallax correction values (offsets) for all conjugate segment pairs contained in the tree. Exempted are the excluded anomalous regions (e.g. clouds, water, etc.). For other failed matches, the parallax correction values are approximated by interpolation from the neighbourhood.
The two conjugate trees of the target and search segments are matched gradually from the upper-most level (n-1) down to the lowest level (1). A special subroutine navigates the process of accumulation of the parallax correction values down the branches of the hierarchical tree. Hence, to each target segment in the lowest level (1) the accumulated value of the parallax correction is assigned. These values, orderly arranged, along the epipolar lines, from together the resulting raw epipolar parallax profiles. For their further use in matching stage 4, the adjacent profiles should be verified and conditioned collectively by means of special subroutines.
Stage 4, addressing fine matching, is beyond the scope of this paper. For the common DTM applications, however, addtional fine matching may not be necessary

IV INITIAL TESTS

To date, the tests covered mainly the subroutines for image preprocessing and matching of edge-pairs. To this end, digital images were produced for conjugate sections of a representative image-pair using the Kern DSR1-Video system. First the two video-cameras had to be calibrated in both geometric and intensity domains. The geometric calibration was by means of the VIDICAL routine and corresponding test targets, both developed by the Kern Co. The program provides an affine transformation of the 2D-sensor array into the photo-stage coordinate system. The calibration in the intensity domain was necessary for tuning the illumination intensity to the dynamic range of the images. This is to provide optimal sampling and digitising. Both the images and a gray wedge were scanned by a microdensitometer. Subsequently, the illumination intensity could be tuned such that the level 255 corresponded with the highest intensity in the image and the level 0 with the lowest one.

Some further tests concerned warming up and stability of signal output in time, repeatibility, a comparison of both video cameras, etc.

Image sampling and digitising were done directly in the epipolar format. Positioning was controlled by a subroutine (in Pascal II); for image sampling the central column of the sensor array was used, with its mid-pixel positioned on the specified epipolar line. While moving the sensor column along the epipolar line, a strip of 369 parallel lines was sampled and digitised.

Further processing was carried out on a VAX-780 computer. After specifying the parameter values for preprocessing for stages 2 and 3 (filter kernels, thresholds, etc.), the software modules were tested in a logical sequence. These modules address building of the hierarchical levels, image segmentation, edge extraction and ege description.

Similarly, prior to testing the modules for matching, the values of the corresponding process parameters were specified for stages 1, 2 and 3. These values concern the base-to-height ratio, the convergency ($\Delta\phi$), the search range(s), the algorithms and measures for similarity assessment, and the acceptance thresholds. Then the individual modules were tested one by one. These tests and the associated optimization, however, have not yet been completed, and are an issue of further development.

V. FINAL CONSIDERATIONS

The effort in development of a flexible multistage image matching system with self-adaptive capability should be pursued. An important requirement is fast sampling and digitising of images in epipolar format, which requires improvements in the hardware.
Consideration should also be given to alternative strategies and algorithms for image preprocessing, to the ranges of the corresponding parameters values, and to the adaptive capabilities in all stages.
In further development, special attention should be given to the super-fine matching. The information on terrain relief from stage 3 can be analysed in conjunction with the corresponding intensity distribution in images. By applying suitable decision rules, the strategy and algorithms for fine matching, inspection and editing of the output could be selected and implemented automatically.

REFERENCES

[1] Forney, G.D., "The Viterbi Algorithms", IEEE Proceedings, Vol.61, Nr. 3, 1973.

[2] Pereira, L.M.G., "Automatic stereopsis: A robust Technique for Image Matching", M.Sc. Thesis, ITC, The Netherlands, 1988.

[3] Makarovic, B., "Digital Image Registration by Means of Compressed Data", ASP/ACSM Fall Technical Meeting, Little Rock, USA, October 1977.

[4] Makarovic, B., "Automatic Off-line Generation of DTM", ISP Congress, Com.II, Hamburg, 1980.

[5] Makarovic, B., "Image Correlation Algorithms", ISP Congress, Com.II, Hamburg 1980.

[6] Makarovic, B., "Automatic Production of DTM Using Digital Off-line Technique", ISPRS Congress, Com.II, Rio de Janeiro, 1984.