AERIAL IMAGE MATCHING BASED ON ZERO-CROSSINGS

Jia Zong **Jin-Cheng Li** **Toni Schenk**

Department of Geodetic Science and Surveying
The Ohio State University, Columbus, Ohio 43210-1247
USA

Commission III

ABSTRACT

One of the basic tasks in digital photogrammetry is to find conjugate points in a stereo pair and to reconstruct the 3-D object space (DEM). Edges play an important role in that they may indicate breaklines in the surface. We use the LoG operator to extract edges (zero-crossings). In this paper the problem of matching zero-crossings is addressed. Zero-crossings computed from one image are matched with area-based method. A hierarchical matching approach is adopted by the use of both, interpolated disparity maps at each level of the image pyramid, and knowledge from image analysis at very high level of image pyramid. The method is particularly suited for matching aerial images for the purpose of restructing surfaces of urban areas.

KEY WORDS: Zero-crossing, Correspondent point, Figural Continuity, Disparity Interpolation, Image Analysis.

1. INTRODUCTION

One of the major research areas in digital photogrammetry is image matching for reconstructing the three-dimensional surface of the object space. This process involves a fundamental problem of stereo vision: to find corresponding points in an stereo-pair. Once correspoinding points are determined their three-dimensional positions can be easily computed, and the surface is obtained from matched points by interpolation.

Two methods are commonly used in image matching: area-based image matching and feature-based image matching. Aera-based matching is predominantly used for the object space (DEM). Here, the corresponding points are found by comparing the gray levels of correponding areas (image patches) in a image stereo-pair. This approach is favored in photogrammetry because of its high accuracy potential. However, there are several critical factors that need special consideration in area-based matching. For example,

- good approximations for the corresponding image patches are required
- matching in flat area or of sharp relief changes is extremely hard and it produces bad results. Both cases usually occur in urban aerial images
- recovering the surface, especially in urban areas, from randomly distributed matched points is difficult
- the reliablity control of the matching is low
- computations are intensive

Some of these problems are avoided in feature-based matching. Here, properties (features) derived from the gray levels are matched, rather than gray levels themselves. This method usually proceeds in two steps, the first being a local similarity matching such as comparing the parameters of detected features, and the second being a global matching such as checking continuity constraints. Features detected monocularly may differ and may include spurious data due to differences in reflectance which are not caused by the surface shape. This problem is quite acute in large-scale aerial images of urban areas. Another point to bear in mind is that matched features (e.g. edges) do not necessarily consist of conjugate points. In general, feature-based matching is more robust and less computationally intensive. But most important, matched features are more meaningful than randomly matched points if it comes to automatically analyzing image.

The motivation for this research is to combine the merits of both area-based and feature-based matching methods. First, edges or zero-crossings (ZC) are detected as features. The edges are more likely to represent prominent features of the surface, such as breaklines. Instead of matching edges as entities as described in [Schenk *et. al.* 1991], here we match every point of an edge by correlation. A match is accepted if it satisfies epipolar geometry and figural continuity constraints. This strategy proved to be quite successful [Li *et. al.* 1990]. In order to cope with urban areas where correlation must be applied with caution, we have modified the strategy by including a surface analysis step in the hierarchical matching scheme. At each level of the image pyramid an interpolated disparity constraint map is generated which provides the necessary approximations for the next level of matching. Knowledge gained from previous levels is used to guide matching in the subsquent level of image pyramid. With this new strategy the success rate of matching aerial images of complex urban scenes is greatly improved.

2. FEATURE EXTRACTION

Detecting zero-crossings as features for matching was first proposed by Marr and Poggio [Marr and Poggio, 1979] on the basis of a computational theory on the human stereo vision. Mathematically, zero-crossings are obtained by applying the convolution operator $\nabla^2 G$ over the image $f(x, y)$ as

$$G(x,y) = \frac{1}{2\pi\sigma^2} exp(-\frac{r^2}{2\pi\sigma^2})$$

$$\nabla^2 G(x,y) = (\frac{r^2 - 2\sigma^2}{2\pi\sigma^2}) exp(\frac{-r^2}{2\sigma^2})$$

$$f'(x,y) = \nabla^2 G(x,y) * f(x,y)$$

where $G(x,y)$ is a Gaussian filter, $\nabla^2 G$ is the Laplacian of a Gaussian (called LoG), and $f(x,y)$ is the image gray level function. Convolution us denoted by $*$, and $r = (x^2 + y^2)^{1/2}$ implies that the operator is rotationally symmetric. The advantage of the LoG operator is that it combines smoothing and differentiating into one operator. Moreover, it is localized in space and frequency domains. The filtered image $f'(x,y)$ is divided into positive and negative regions with average frequency of $\sqrt{2}/\sigma$. The boundaries of these regions are the zero-crossings. Zero-crossings occur wherever the gray levels change sharply. The degree of change can be described by the first-derivative of the gray level function, or the gradient of the gray levels. Zero-crossings are separated by an average distance which is equal to the window size of LoG operator, the diameter of positive central region of LoG curve $\omega = 2\sqrt{2}\sigma$. The larger the window size, the larger the dislocalization of detected zero-crossings from the real boundaries.

Edges in aerial images represent object boundaries or markings (e.g. shadows). Many object boundaries correspond to surface breaklines. The LoG operator is applied to both left and right image to obtain the zero-crossings. Several parameters are chosen to control feature detection. The window size ω of LoG operator is selected according to the quality and the scale of the images to ensure surface feature detection. In order to supress noise or less important features, a threshold value t is chosen according to the distinctness of the zero-crossing. The result of applying LoG operator are two binary images. Zero-crossings as feature entities are obtained in the left images as following:

- The location of zero-crossings is obtained by an edge following algorithm. The connected zero-crossing points form the zero-crossing curve as feature entity.
- Then each zero-crosssing curve is segmented using local curvature maxima as end points of each segment.

As a result, edges are detected as individual zero-crossing curves connected by several possible segments.

3. CORRELATION MATCHING

The flow chart of the matching scheme is shown in Fig. 1. Like most area-based matching algorithms, epipolar geometry is employed to constrain the searching to one dimension [Cho *et. al.* *1992*]. At each level of the image pyramid, the image patches are first enhanced since area-based matching methods require good image quality. Next, zero-crossings are determined in both images. For each zero-crossing point in the left image the corresponding point on the right image is found along the epipolar (scan) line by area correlation. right image zero-crossings only help to define the searching window for the correlation matching. The matching is performed in two steps: initial point to point correlation, and figural continuity checking acceptance criterion. During the initial matching, points with maximum correlation values larger than the preset threshold value are selected as matched points. The key point here is to find a good approximation of the search window in the right image. This is accomplished by using the disparity constraint map at each level of the image pyramid. Once matching is completed, an interpolated disparity image is generated, providing the approximations needed for the next level matching . In the highest level of the image pyramid, knowledge gained from surface analysis is also fed back to the matching process through the use of the disparity map. After the initial matching, all matched points must satisfy the figural continuity constraint for final acceptance as conjugate points.

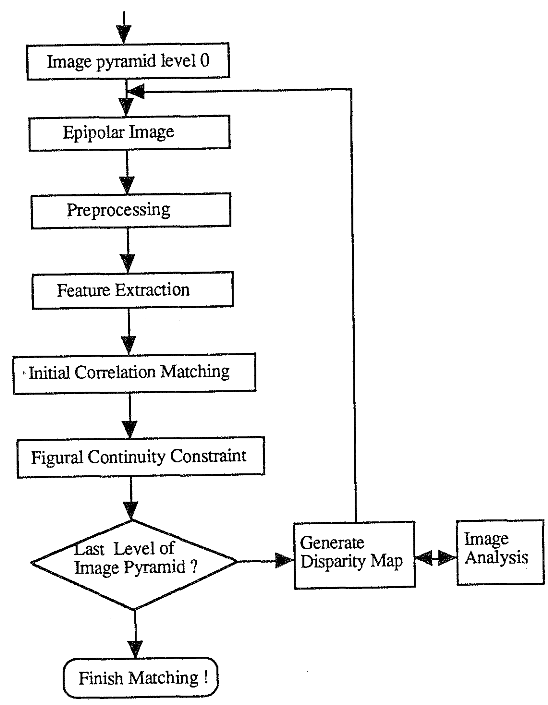

Fig. 1. Flowchart of matching scheme

3.1. Hierarchical Disparity Constraint

The search window in correlation matching is defined by two parameters: location and size. Obviously, the window size depends on the goodness of the approximations. We determine the window size (search range) dynamically based on the disparity map. If the search window is close to a zero-crossing contour detected in the right image, it is adjusted accordingly, because this zero-crossing is likely to be the conjugate point.

A crucial step in any matching system is the approximation of the matching location (center of search window). At the top of the image pyramid we have two options. One is using an average disparity value for all matching positions. This average approximation is computed from the matched points generated during automatic orientation [Schenk *et al.* 1992, and Zong *et al.* 1991]. The second option is to convert the matched points from the automatic orientation to disparity values. These disparity values are then interpolated to generate a disparity map to determine the location of the search window. After the matching process is completed in one level of the image pyramid, the disparity map is updated to ensure that better approximations in the next level are available. This is quite important, particularly in urban areas where the disparity values may abruptly change. It should be noted that the disparity map always corresponds to the resolution of current level in the image pyramid.

3.2. Surface Analysis in High Level Matching

The use of the disparity map provides not only good approximations for correlation matching but also a closer surface approximation after each level of matching. In higher levels of the image pyramid such a surface can even provide a lot of three-dimensional object information. This information can then be used for real object surface analysis, as discussed in [Wang *et. al.* 1992]. On the other hand, the information gained from the 3-dimensional analysis can be fed back to guide the matching.

One of the most difficult matching cases are urban aerial images in which there exist man-made features with extreme height, such as tall buildings or chimneys. The deformations and disparities of such features in stereo images can be very large causing the matching to be either incomplete or unsuccessful. The solution here is to analyze the disparity map. As one application of 3D feature analysis discussed in [Wng *et. al.* 1992], a contour map can be generated after segmenting a disparity map. The following rules are implemented to superimpose knowledge to the existing disparity map and used to guide the program to find potential high features.

- A cluster of close-centered contours indicates a potential hump
- If the inner disparity values are much larger than the outer ones, a potential high hump is indicated
- For a potential high hump, information is fed back to guide a future matching
- The boundary of a potential high hump is the second closed outer contour since the first one may indicate boundary of the environment
- The potential disparity values inside the selected boundary are the average disparity values of the matched points inside the boundary
- The obtained disparity values of the potential hump are appended to the matched data and a new disparity map is interpolated

3.3. Figural Continuity Constraint

The figural continuity criterion implies that the disparity values along zero-crossings must be continuous. We implemented the figural continuity constraint by performing a Hough transformation of all the matched points belonging to one segment of a zero-crossing contour. Continuous disparity values show up as clusters in the Hough space. If fewer than 15 points fall into the cluster a flag is set to indicate that there is no corresponding zero-crossing segment. Finally, the location of the corresponding segment in the right image is determined by the Hough transformation and the correlation threshold.

4. EXPERIMENTS

The matching algorithm was tested with several pairs of aerial photographs. In this paper, we present the results from stereo-images (193, 195) taken over the campus of the Ohio State University. This model represents a very typical urban area of all the different models tested, it was the most difficult one. The photo scale here is approximately $1:4000$. The diapositives were scanned to a resolution of 30μ pixel size by Intergraph Corporation using the PhotoScan. However, we only used a resolution of 60μ which yielded a 4096×4096 pixel image. The ground coverage of a pixel is approximately 25×25cm.

Fig. 2 and 3 show the original aerial images at the coarsest resolution of 512×512. Zero-crossings were first detected with the LoG operator ($\omega = 5$), and then matched with a single average disparity approximation. The range of the search window was set to 10 pixels in order to avoid wrong matching. The matched zero-crossings are shown in Fig. 4 and 5. A disparity map was interpolated by using Modular function on Intergraph workstation. The result is shown in Fig. 12. which outlines the surface of the whole overlapping area of the model. Some humps are clearly visible.

The interpolated disparity map was then converted into an image of 512×512 resolution and the disparity values were treated as graylevels. Fig. 6 and 7 show matched zero-crossings of a 512×512 image patch selected from stereo images of $1K \times 1K$ resolution. Fig. 13 shows the interpolated 3D disparity map. The humps are now more prominent. Fig. 8 and 9 depict matched zero-crossings of a 512×512 image patch from $2K \times 2K$ resolution images. The interpolated disparity values are shown in Fig. 14. The surface is fairly well approximated at this level.

The procedure is repeated at the finest resolution, again with an image patch size of 512×512 pixels. In this example the disparity values range from 0 to 118. The segmented disparity image resulting from matching is shown in Fig. 16 where the hump is clearly indicated. Fig. 10 and 11 show the matching results superimposed to the resampled images, while Fig. 15 and 17 show the interpolation of the final matching results in the disparity map and the three-dimensional object space, respectively.

5. CONCLUSION

The presented matching scheme combines the merits of both area-based and feature-based matching methods and proved succussful in the aerial image matching. The use of a hierarchical approach and surface approximation makes this approach particularly suited for urban area image matching. It is found that the precise detection of prominent features is helpful for recovering the object surface. The reliability of the correlation matching is improved by the employing the figural continuity constraint. Finally, this matching scheme shows a great potential for object surface analysis and reconstruction.

6. REFERENCES

Cho, W., M. Madani and T. Schenk, 1992. Resampling Digital Image to Epipolar Geometry. *Proceedings of ISPRS*

Li, J.C., and T. Schenk, 1990. Aerial Image Matching Based on Zero-crossings and Least-squares Correlation, *Technical Notes in Photogrammetry No. 9, Department of Geodetic Science and Surveying, The Ohio State University*

Marr, D. and T. Poggio, 1979. A Computational Theory of Human Stereo Vision, *Proc. Royal Society*, B 204 301 – 328

Schenk, T., J.C. Li and C. Toth, 1990. Hierarchical Approach to Reconstruct Surfaces by Using Iteratively Rectified Imagery , *Proceedings of the ISPRS*, Commission V, 28(5/1) : 464 – 470

Schenk, T. and C. Toth, 1992. Conceptual Issues of Softcopy Photogrammetric Workstations. *Photogrammetirc Engineering and Remote sensing*, Vol. 58, No. 1, Jan. 1992, pp. 101-110

Wang, Z. and T. Schenk, 1992. 3D Urban Aera Surface Anlysis *Proceedings of ISPRS*

Zong, J., J.C. Li and T. Schenk, 1991. Application of Forster Interest Operator in Automatic Orientation System, *Proceedings of the ASPRS-ACSM Annual Convention*, 5:440-448

Fig. 2. Original left epipolar image 193 in 512 × 512 res.

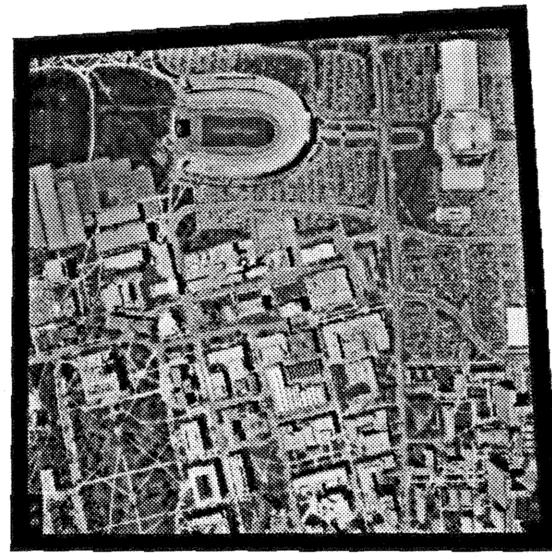

Fig. 3. Original right epipolar image 193 in 512 × 512 res.

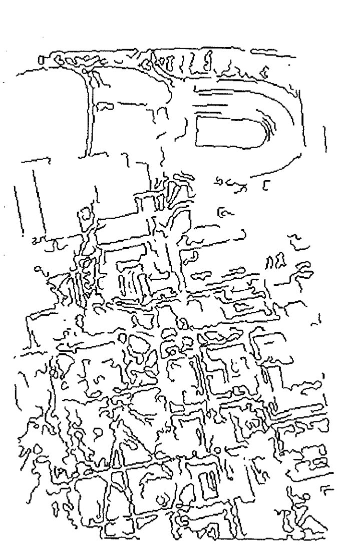

Fig. 4. Matched ZCs of L-image (512 × 512 res.)

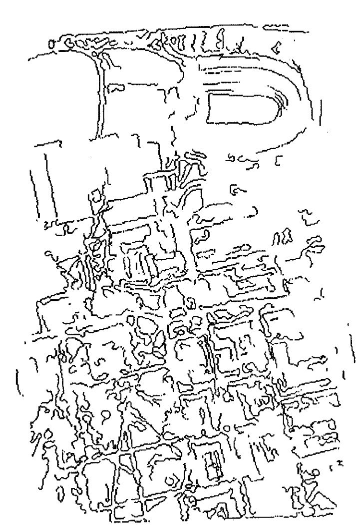

Fig. 5. Matched ZCs of R-image (512 × 512 res.)

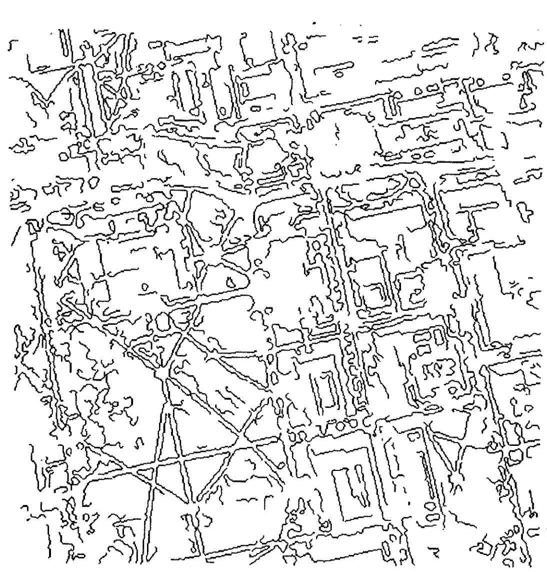

Fig. 6. Matched ZCs of L-image ($1K \times 1K$ res.)

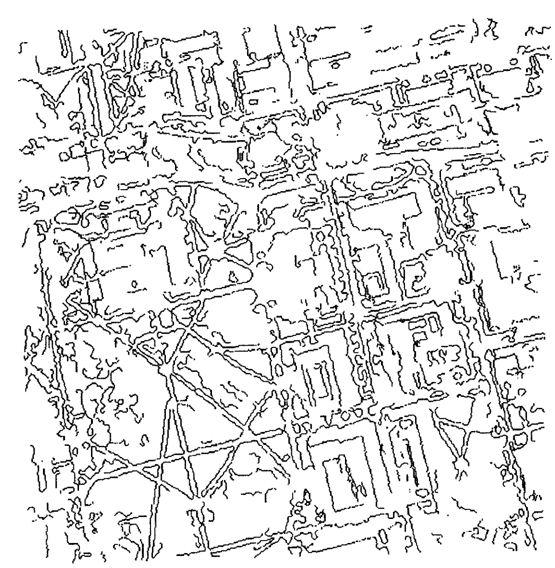

Fig. 7. Matched ZCs of R-image ($1K \times 1K$ res.)

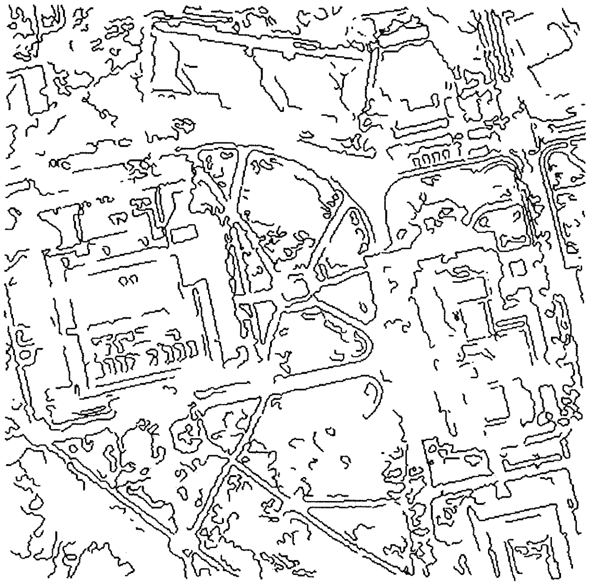

Fig. 8. Matched ZCs of L-image ($2K \times 2K$ res.)

Fig. 9. Matched ZCs of R-image ($2K \times 2K$ res.)

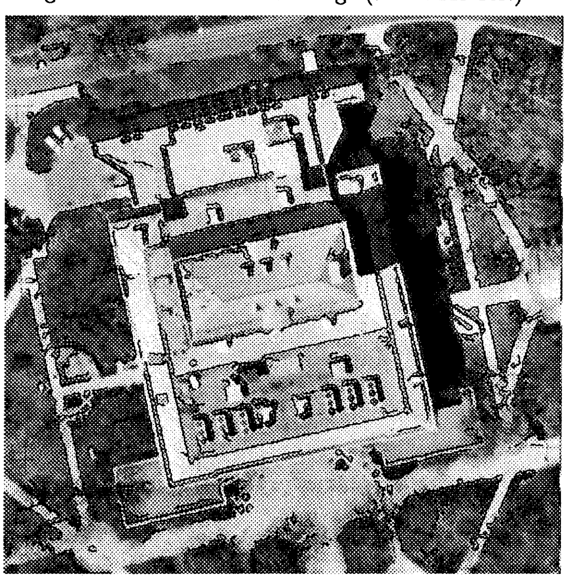

Fig. 10. Matched ZCs overlapping in L-image ($4K \times 4K$)

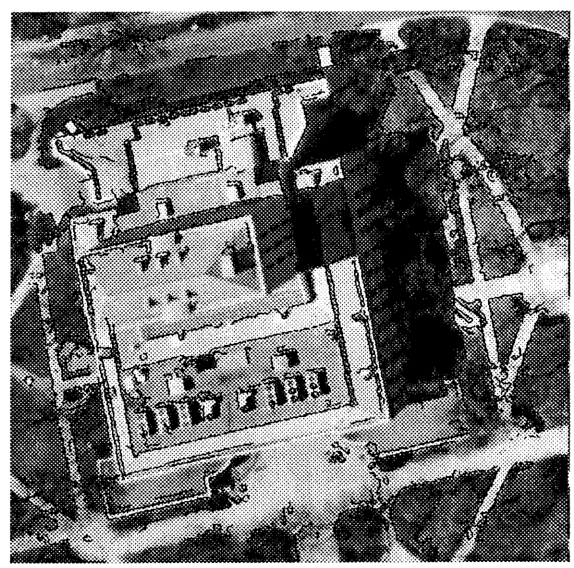

Fig. 11. Matched ZCs overlapping in R-image ($4K \times 4K$)

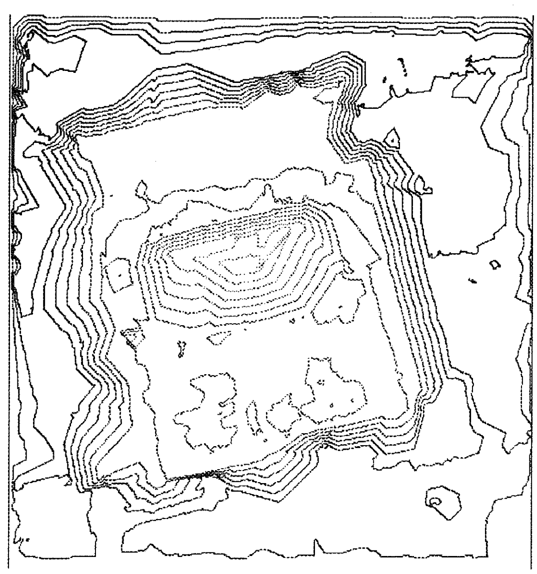

Fig. 12. Segmented disparity map from $4K$ matching

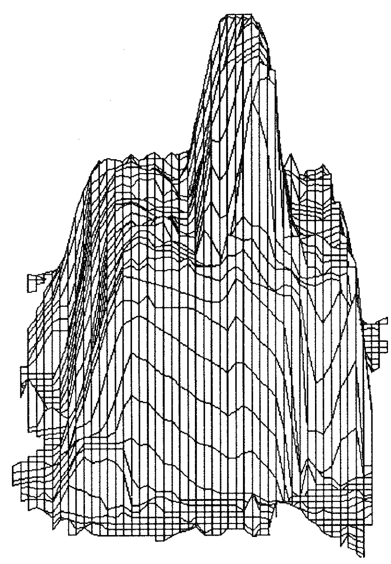

Fig. 13. DEM in object space from the final matching ($4K$)

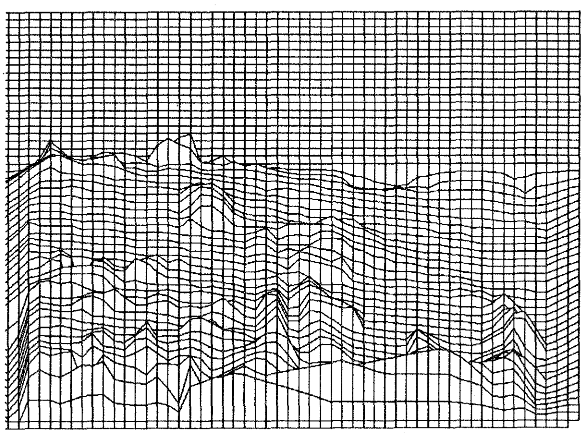
Fig. 14. Interpolated disparity map from 512 res. matching

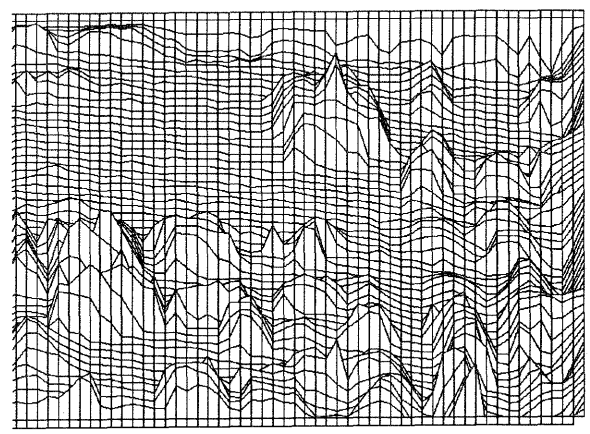
Fig. 15. Interpolated disparity map from 1k res. matching

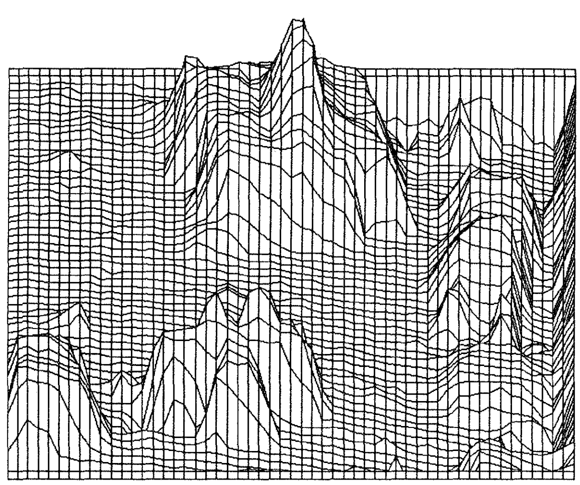
Fig. 16. Interpolated disparity map from 2k res. matching

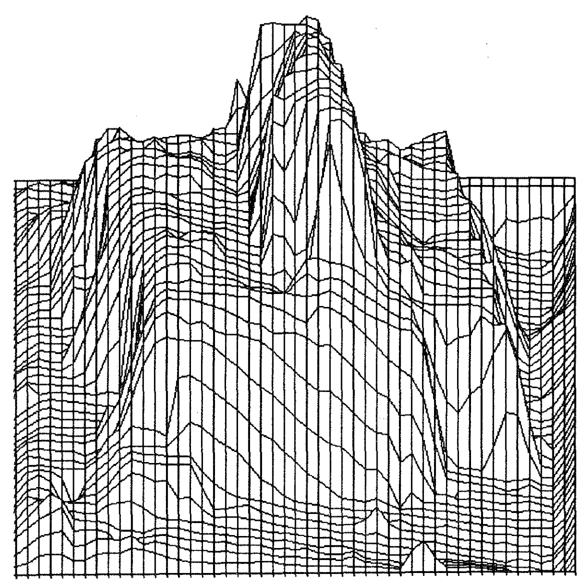
Fig. 17. Interpolated disparity map from 4k res. matching