Conceptual Issues of Softcopy Photogrammetric Workstations

Toni Schenk and *Charles K. Toth*
Department of Geodetic Science and Surveying, The Ohio State University, Columbus, OH 43210-1247

ABSTRACT: Digital photogrammetry involves ultimately all levels of the computer vision and machine perception paradigm. Despite all the progress in vision and digital photogrammetry, we believe that there is still a considerable lack of understanding of theories and methods which would allow the development of a machine capable of automatically producing maps, for example. We argue that major progress toward autonomous softcopy workstations depends more on advances on the conceptual level rather than on the refinement of system components such as hardware and algorithms. Past experience in computer science allow one to predict, and recent surveys confirm this, that the development of hardware is progressing much more rapidly than software. Consequently, the design of softcopy workstations should not be totally constrained by today's limitations of system components. In this paper we address problems of automating photogrammetric processes on digital photogrammetric workstations. After presenting our conceptual system, we describe the automatic orientation and the surface reconstruction module, followed by examples.

INTRODUCTION

DIGITAL PHOTOGRAMMETRY is rapidly emerging as a new subfield of photogrammetry. As always when new technologies and methods develop, there is no unified terminology, let alone an accepted definition of *Softcopy Photogrammetric Workstations*, the central theme of the workshop in Boulder. While we intuitively associate such workstations with digital image processing stations of some sort, there is a wide diversity of opinion about the functionality and the degree of automation softcopy photogrammetric workstation should exhibit.

An example may illustrate the case. Suppose a digitized stereopair is displayed on two monitors and viewed stereoscopically, employing one of the three-dimensional viewing techniques. The operator uses a cursor, like the floating mark of an analytical plotter, to identify points and features whose image coordinates are then recorded. Model or object coordinates can easily be computed, again in a manner similar to analytical plotters. If the only process consists of displaying the digital imagery and straightforward methods from analytical photogrammetry are then applied, do we call this digital photogrammetry?

Inevitably linked with the emergence of new disciplines and methods is a strong temptation to apply them without a clear concept. Reports about the capabilities and benefits of digital photogrammetry, perhaps not read and interpreted in the appropriate context, may push the expectations beyond what can be delivered in the foreseeable future. Photogrammetric processes are well understood and products such as maps are generated with great success. Why not do the obvious and automate the processes, perhaps by endowing a softcopy photogrammetric workstation with an expert system?

When a mapping project is given to an operator, the instructions are relatively short, even ambiguous or incomplete, because the human operator has enough knowledge and common sense to solve the problem. On the other hand, for a computer to perform the same task requires a great deal more instruction, for lack of common sense or specific knowledge. If the task were to digitize all buildings, then this is the only instruction the human operator will need. Why does the same task pose an unsurmountable problem to the computer? When viewing the stereomodel, the operator constantly interprets the three-dimensional (3D) model, basically by reconstructing the surface followed by scene interpretation and object recognition. This is the very essence of computer vision and machine perception: reconstruct and interpret a 3D scene from its 2D projections. We all perform this task with great ease but there are only vague ideas on how the human visual system makes the problem appear so simple. As long as our knowledge remains rudimentary there is no hope that a machine will understand the instruction "Digitize all building." Hence, further progress in automating photogrammetry depends on a precise understanding of how human operators solve problems. Artificial intelligence and cognitive science help to analyze these processes. Once understood methods can be devised and appropriate algorithms can be designed. We argue that more efforts should be undertaken to make the knowledge of mapmaking explicit and to elucidate the human operator role in photogrammetric processes.

This approach is in contrast to the popular method of trial and error: different methods and algorithms are tested until a satisfactory solution is found—only until it fails in the next application. Gaps in the theory and lack of concepts cannot possibly be compensated for by adding more memory to a softcopy photogrammetric workstation. This is not to downplay the importance of refining hardware and software. The evolution of photogrammetric workstations and a clear understanding (theory) of photogrammetric processes will have to interact to further advance the exciting field of digital photogrammetry.

This paper has a strong computer vision flavor. We hope this will provoke thought and stimulate discussions. We address more conceptual issues than implementation details. In the next section we provide some background information and describe our approach towards digital photogrammetry. As an illustration, we present in the following sections two examples of modules which are in various development stages. The basic approach of the automatic orientation module is summarized followed by results. Next, we present the concept of the surface reconstruction module, including some illustrative examples.

BACKGROUND

Figure 1 depicts the major building blocks of our conceptual system to generate photogrammetric products, such as digital maps, automatically. It involves the entire computer vision paradigm. Computer vision is the construction of explicit, meaningful descriptions of physical objects from images (Ballard and Brown, 1982). Abstract descriptions are essential for object recognition. To reason automatically from raw images to a fully interpreted scene is a gigantic task and usually grossly underestimated because humans solve it without conscious effort, in real time. A key element is to break the task into several modules and to design useful intermediate representations.

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING,
Vol. 58, No. 1, January 1992, pp. 101–110.

0099-1112/92/5801–101$03.00/0

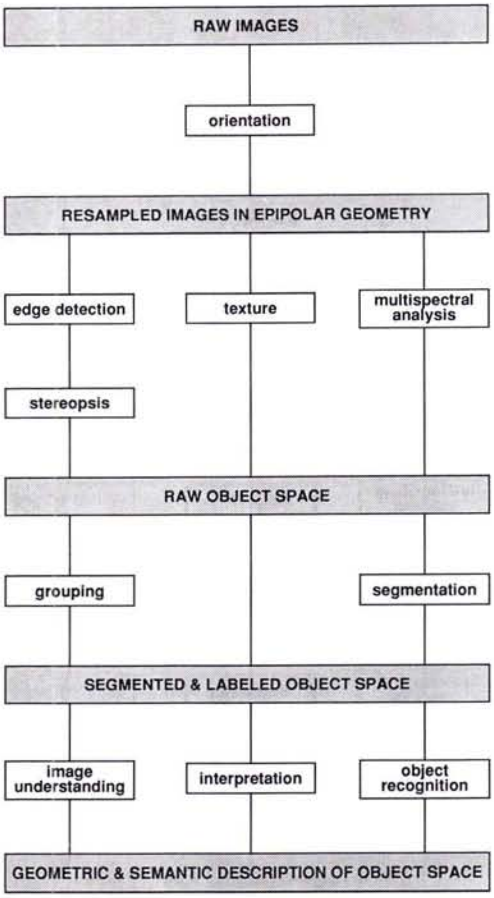


FIG. 1. Building blocks of conceptual digital photogrammetric system.

Because original imagery in our applications is not registered in epipolar geometry we begin with the orientation module. The result of our automatic orientation procedure are resampled stereopairs such that corresponding points are found in the same row of the digital image. Moreover, many (hundreds) of points are precisely known in object space, giving rise to a first surface approximation.

The raw object space is an important intermediate representation. It corresponds to the 2.5D primal sketch proposed by Marr (1982), or to the intrinsic image proposed by Barrow and Tenenbaum (1982). The raw object space contains basic information about the surface: breaklines, depth, and surface normals. Unlike the images, the representation of the raw object space is not iconic; it is already a step toward the geometric and semantic description of the desired end result. That is, the degree of explicitness is increased. The reader should contrast this with a DEM and orthophotos, both of which are traditional object space representations. However, most of the relevant information is still implicit: an orthophoto does not show explicitly edges or regions and the DEM is not segmented into smooth surface patches.

The raw object space is the interface between early vision and late vision processes[1]. Early vision deals with pixels—it processes signals—thus, it is data-driven. Late vision is model-driven; it deals with properties of groups of pixels. The quest is to transform the raw object space into new representations that would allow matching the descriptions with a model world. Clearly, late vision is application dependent for it is not conceivable to model all possible objects and their relations and store them in a model base. Additionally, perceptual processes require common knowledge.

Figure 1 should not be confused with the architecture of a real system, say a softcopy photogrammetric workstation, where much more care for efficient communication between modules and implementation of hardware and software components is required. The concept serves more as a master plan to coordinate our research activities.

The conceptual system may also serve as a base to compare the degree of automation a real system offers. What in the long chain of tasks required to generate a map is performed automatically?

AUTOMATIC ORIENTATION

The orientation of stereopairs is a well-known process in photogrammetry. Numerous analog solutions have been developed mainly during the first part of this century. The analytical solution can easily be described: measure a number of well distributed corresponding points, use a suitable mathematical model, and determine the parameters by a least-squares adjustment.

What is the problem when it comes to solving this relatively simple task automatically? The challenge is obviously to find enough well-distributed corresponding points of sufficient accuracy. A general solution should not be restricted by severe assumption or by introducing too many constraints.

In this section we summarize the major steps and present results. For more detailed descriptions the reader is referred to Schenk *et al.* (1990a, 1991), Stefanidis *et al.* (1991), and Zong *et al.* (1991). The orientation module incorporates early vision processes and also illustrates nicely the coarse-to-fine strategy for deriving approximations. Because the stereopairs are not registered in epipolar geometry, most of the solutions from computer vision for solving the correspondence problem are not applicable[2].

We have combined the advantages of feature-based matching (robust, good approximations) with those of area-based gray-level correlation (high point precision).

STRATEGY AND IMPLEMENTATION

Figure 2 shows the major steps to be performed by the orientation module. It is governed by the following strategy:

- Coarse-to-fine approach[3] conquers the problem with weak assumptions about the camera parameters. Matching on a coarse level is less ambiguous and constrains the search space on the next levels.
- Feature-based matching (edge matching) assures that matches are only sought in areas of sufficient gray-level variations. Also, it greatly reduces the search space.
- Precise point matching renders the required sub-pixel accuracy for corresponding points. It may be viewed as a refinement of positions found by edge matching.
- Consistency checks throughout the scale space take into account previous results and check new observations by a least-squares adjustment with blunder detection. Points identified as blunders are no longer traced through the scale space.

The coarse-to-fine strategy is implemented as an image pyramid. Most of our digitized aerial photographs have a resolution of 4096 by 4096 pixels. The image pyramid is obtained by

[1]Also termed low-level and high-level vision processes.

[2]Computer vision almost exclusively deals in controlled imaging environments where the camera axes are parallel and at right angles to the camera base, thus rendering images in epipolar geometry.

[3]Also termed hierarchical, multiscale, or scale-space approach.

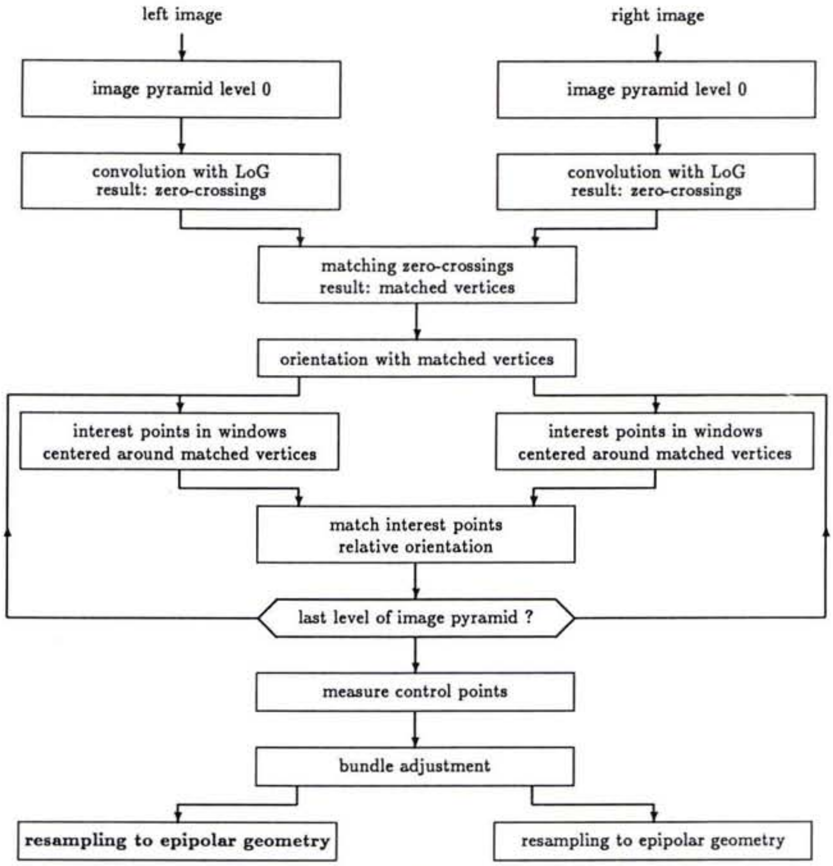


FIG. 2. Overview of automatic orientation module.

smoothing the original image with a Gaussian filter of varying σ. The top of the pyramid is usually represented by 512 by 512 pixels in our application. Corresponding points found on one level are projected to the next level where their location is refined because more detailed images are available.

Features to be matched are zero crossings that are obtained from convolving the image with the LoG operator[4]. We select the width of the central lobe of the LoG operator, ω, such that enough zero crossings are obtained for unambiguous matching. Zero crossings are statistically separated from one another by ω. With $\omega = 15$ pixels, we may therefore expect approximately 30 zero crossings. Assuming that one third of zero crossings can be matched, we will obtain well over 100 corresponding points.

Zero crossing contours look similar in both images, but sometimes they are connected differently or are only present in one image. Also, the beginning and the end of contours may be different in both images. Thus, a general solution must be found for detecting corresponding contours. We match entire zero-crossing contours by considering similar shape, similar convolution gradients, and certain constraints in terms of their locations in both images[5].

Matching edges takes place in the $\psi - s$ domain. The $\psi - s$ representation of a line is a function $\alpha = \psi(s)$ where the length s is the parameter of the tangent ψ. The length s corresponds to the number of pixels in the chain-code representation of the contours. The chief advantage of the $\psi - s$ representation is that it is invariant with respect to rotation and position in the original x,y domain.

To that end, we segment the $\psi - s$ representation. This corresponds to a piecewise circular approximation in the x,y domain. Edges of similar shape are characterized by similar vertices. After all vertices with similar angles are found, they are grouped together and analyzed for line consistency. Vertices of a contour in one image correspond to vertices of one or several lines in the other image. However, usually it is a one-to-one correspondence.

With all the matched vertices a first adjustment is performed, mainly to detect blunders. In our experience the root mean square error of the parallaxes is ± 1 pixel. That is, matching zero crossings on the top level of the image pyramid renders approximations to an accuracy of ± 0.5 mm[6]. The vertices which passed the blunder detection test are projected to the next level where they serve as centers of windows within which interest points are determined and matched.

RESULTS

We have performed experiments with several digitized aerial stereopairs. The results are summarized in Table 1. Except model "Campus" all diapositives were digitized with an EIKONIX camera to a resolution of 4096 by 4096 pixels. The camera was

[4]The Laplacian of a Gaussian is a combination of a Gaussian smoothing filter with the Laplacian ∇^2 of the smoothed image.

[5]As an example of a weak assumption, we expect both images of the stereopair in normal position. That is, they are not rotated by 180° to one another. In that case, it is not possible that a contour in the upper left corner of one image has its corresponding contour in the lower right corner in the other image.

[6]One pixel at the 512 by 512 resolution is approximately 0.5 mm.

TABLE 1. RESULTS OF AUTOMATIC ORIENTATION PERFORMED ON FOUR DIFFERENT STEREO MODELS. ALL MODELS HAVE BEEN DIGITIZED WITH A RESOLUTION OF 4096 BY 4096 PIXELS.

Model	Corresponding Points	Blunders	σ [µm] x	y	Control Points	Photo Scale
Campus	112	12	0.5	3.5	6	1 : 4,000
Hunter-Ligget	112	15	0.2	3.3	4	1 : 80,000
Munich	88	11	1.4	3.9	6	1 : 3,800
Daun-Mehren	118	10	1.5	3.7	6	1 : 12,000

carefully calibrated to an accuracy of about 6 micrometres. The camera calibration procedure and test results are described in Li and Schenk (1990a). The "Campus" model was digitized by courtesy of Intergraph Corporation with the PS1 PhotoScan at various resolutions.

To further illustrate our approach we include in Figure 3 model "Daun-Mehren," the zero crossings obtained on the top level of the image pyramid with an LoG of $\omega = 15$ pixels as well as the matched zero crossings.

The results presented in Table 1 were achieved with a rigorous bundle adjustment (without additional parameters). Note the accuracies achieved. The standard deviation σ in y is, at 4 micrometres, consistently around 1/15 of the pixel size at the finest resolution. Not surprisingly, the standard deviation in x is almost zero because errors "disappear" as incorrect elevations[7].

Identifying and measuring control points is the only interactive step in the orientation module. We measure the control points on Intergraph's Interpro 3055 workstation by reading the cursor position. The accuracy is approximately 1/3 of a pixel. Current research addresses the difficult problem of locating control points automatically (see Al-Tahir *et al.*, 1990; Toth and Schenk, 1990).

From our experiments we may conclude that a resolution of 4096 by 4096, or approximately 60 micrometres, is sufficient to automatically orient digital stereopairs to an accuracy comparable to manual orientation on analytical plotters. This may come as a surprise, but closer examination reveals that many pixels contribute to the determination of corresponding points. Also, the redundancy with a hundred points in the adjustment is remarkable[8] and contributes not only to a better accuracy but, perhaps more important, to an increased reliability.

SURFACE RECONSTRUCTION

In the context of our application, surfaces play an important role for several reasons. For one, the topographic surface of the Earth is itself, in the form of a DEM, a frequent end result of photogrammetric processes. Perhaps more important, surfaces are an intermediate, essential step toward image understanding and object recognition. Following the spirit of this paper, we elucidate the concept of our approach and present results. We will make appropriate references for readers who are interested in a more detailed treatment of the subject.

BACKGROUND

The 3D shape of surfaces is essential for image understanding and object recognition. The human visual system is remarkably adept at perceiving shape information and at organizing it into continuous, smooth surface patches. Research in psychophysics, psychology, and neurobiology has provided strong evidence that early vision modules, such as stereopsis, texture, motion, and color contribute towards understanding 3D shape (see Marr, 1982).

Reconstructing surfaces is an ill-posed problem. Individual depth cues are in themselves not robust; they just contribute evidence toward a hypothesis of the surface. Other modules may confirm or reject the hypothesis. The combination of different depth cues has gained much interest recently, see, e.g., Aloimonos and Shulman (1989), Grimson (1984), and Horn (1986). Poggio suggests integrating information from several modules within the framework of Markov Random Fields (Poggio, 1988).

Regularization theory deals with ill-posed problems. Poggio *et al.* (1985) and Terzopoulos (1986) report about the regularization of ill-posed early vision processes. Basically, the goal is to restrict the space of acceptable solutions by imposing physically plausible constraints. Constraints were introduced long before the theoretical framework of regularization theory became available. Virtually every work in surface reconstruction deals with various kinds of assumptions and constraints that are applied to assure a solution. Quite often the assumptions severely limit the range of applications, manifested by the fact that no general purpose, universally accepted automatic surface reconstruction method is available.

In Dhond and Aggarwal (1989) an extensive review about extracting 3D structures from stereo is provided. In Doorn *et al.* (1990) similar efforts undertaken in the photogrammetric community are summarized. Here the predominant method for solving the correspondence problem is area-based matching. Some of the shortcomings of this approach (for a critique, see Horn (1983)) have been overcome by introducing additional constraints, such as shape parameters of the correlation window which lessens the foreshortening problem as outlined in Förstner (1986). In addition to the straight-forward cross-correlation method to determine the maximum correlation factor, the technique of minimizing the gray-value differences between two or more image patches using least-squares techniques has been developed and successfully employed for precise point determination (see, e.g., Ackermann, 1984). Grün and Baltsavias (1988) combine least-squares matching with the collinearity condition to further constrain the correspondence problem. In Rosenholm (1987) this concept is taken a step further by proposing a global matching method which entails matching multiple windows simultaneously by constraining the surface patches to continuity of the surface normals. Recently, the concept of adding to the geometric model the radiometric model and adjusting both simultaneously has been proposed in Helava (1988), Wrobel (1988), and Ebner and Heipke (1988). Conceptually, this compares to combining stereo with shading, as reported in Grimson (1984).

While reliability is improved with these additional constraints, area-based matching is still plagued by false matches. This is an inherent pitfall of matching gray values (or their differences) as they are mere abstract representations of the object space and do not themselves represent physical object-space features. Compared to feature-based matching, it offers the advantage of simpler implementations and, provided good approximations are available[9], a higher accuracy.

SURFACE RECONSTRUCTION STRATEGY

We now turn to the description of our approach to reconstruct surfaces from aerial imagery. The schematic diagram of Figure 4 provides an overview. The strategy can be summarized as follows:

[7]For normal aerial photography, the epipolar lines are almost parallel to the x- axis. Bundle rays with an error in the epipolar line still intersect precisely in object space, but at a wrong location.

[8]In contrast to manual orientation where seldom more than a dozen corresponding points are measured.

[9]Experience shows that the correlation window should be as close as three pixels to its corresponding position.

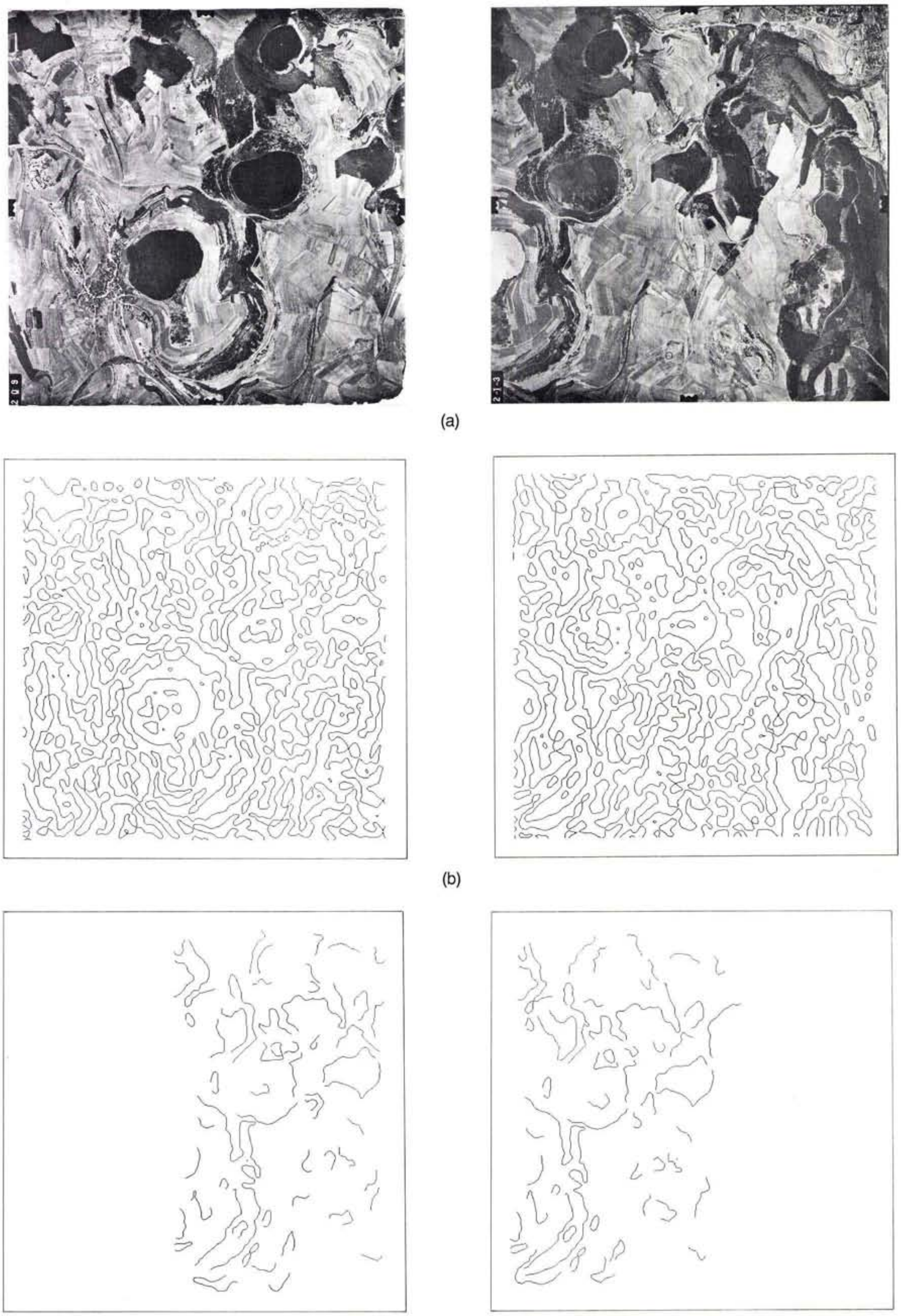


FIG. 3. (a) Model "Daun-Mehren," photoscale ≈1:12,000. (b) Zero crossings obtained with an LoG, $\omega = 15$ pixels. (c) Matched zero crossings using $\psi - s$ method.

- With the orientation parameters determined from the orientation module, the images are resampled to epipolar geometry and an image pyramid is generated. In the coarse-to-fine approach, matching is performed on every scale. Thus, a DEM is obtained from every scale space. In fact, these surfaces may be considered as the discrete scale space representation of the true surface.
- Matching is performed in object space by warping the original images with respect to the current DEM. On the finest level the warped images are in fact orthophotos. Matching with warped images greatly reduces the foreshortening problem.
- Matching on any level leads to a sparse set of points in object space. Information about the surface stemming from other sources than stereopsis, for example, from texture or multispectral analysis, is integrated. This set is analyzed for discontinuities in the surface such as breaklines or occlusions.

Matching in Object Space

In our method of matching in object space, we warp the images on every level of the image pyramid to the current surface to reduce the foreshortening problem. It is not possible to repeat the entire development here, but we will present salient points. The reader is referred to Schenk *et al.* (1990b) for a detailed description.

Motivation.

For area correlation, it is essential that the two correlation windows which render the highest correlation factor correspond to the same area on the surface. The relative difference of relief distortion in both images affects the matching process.

Some area-based methods address this problem[10] by introducing shape parameters. The shape difference between the image patches is approximated by an affine transformation. The transformation parameters are determined independently for every point to be matched. We argue that by using warped images the shape of the surface is globally considered and therefore is a stronger constraint as opposed to the locally determined and independent shape parameters.

It is worth pointing out that feature-based matching is also affected by terrain relief. Edges are computed with an operator of some spatial extent. Suppose the circular shape of the LoG operator is projected onto the surface. Because the projection from the other image does not cover the same surface patch, the convolution values and subsequently the positions of the zero crossings and their gradients will be different. The relative difference of the two corresponding zero crossings translates directly into disparity errors.

Implementation.

As depicted in Figure 5, warped images on level i of the discrete scale space are computed with the current DEM^i. The tesselation of the warped images is identical to DEM^i. The gray level of a point with coordinates I,J of the warped image is found by projecting the coordinates I,J,H of the DEM back to the two images. Figure 5 depicts the warped images that correspond to DEM^i and the original stereopair. Note that the warped images are always computed from the original images to avoid any error propagation.

[10]Also referred to as foreshortening.

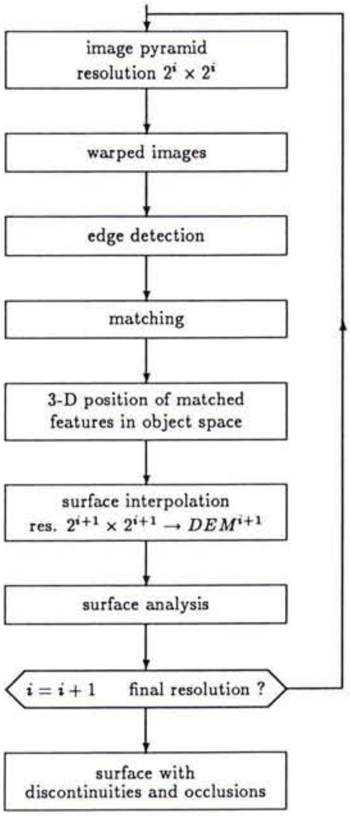


Fig. 4. Overview of OSU surface reconstruction module.

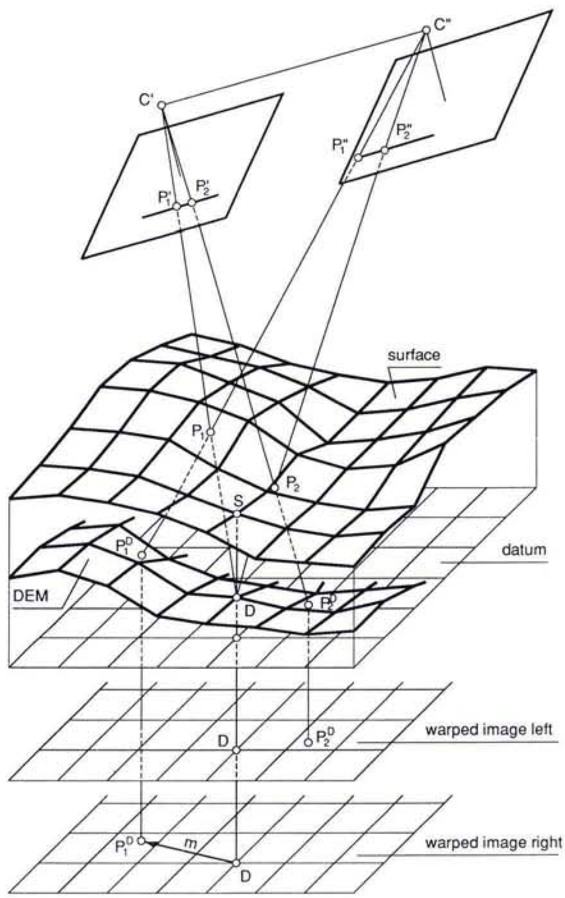


Fig. 5. Principle of object space matching. Shown is the true surface, its approximation by DEM, and the warped images left and right. Vector **m** matches gray levels corresponding to surface point P_1 whose location in the warped images is at D and $P_1{}^D$, respectively.

From matched warped images surface corrections can be computed. Let D be a point on the DEM. Because DEM^i is not yet the final surface, an error Δh exists between DEM point D and true surface point S. If the warped images are computed as outlined, the gray levels at position D do not correspond to the true surface point S, but to the two points P_1 and P_2 which are the intersection of the bundle rays from D to the projection centers C' and C''. If the two warped images are matched, then point D in the left image with its gray level $g(P_1')$ will correspond to point P_1^D of the right image. The matching vector $\mathbf{m}$ renders the position of point P_1^D. Now point P_1 on the true surface is obtained by intersecting the corresponding bundle rays P_1^D,C'' and D,C'. The two rays will not intersect precisely. The shortest distance between them gives a good indication of the quality of the match.

Experimental Results.

In Figure 6 we compare the results obtained with the conventional method of matching in image space versus matching in object space. Based on a synthetic surface and randomly distributed gray levels a stereopair was generated as well as its scale-space representation. We then matched the images by using area correlation[11]. In image space matching, shape parameters were used as described in Norvelle (1992). Basically, the same matching technique was used for the warped images except no shape parameters were introduced. In both cases, the matching results from one level were used to guide the matching on the next level.

As is evident from Figure 6, object space matching performs much better and is more robust than image space matching. For more details compare Doorn (1991).

MATCHING EDGES

Motivation

The most relevant information of 3D surfaces is contained in their breaklines. Breaklines are discontinuities in the surface function $z(x,y)$. They may occur at different scales. That is, there are "smooth" discontinuities, occuring over a large spatial extent. Breaklines are usually associated with "sharp" discontinuities. If all discontinuities are known, then the surface can be reconstructed[12] from them.

Discontinuities in the surface are related to discontinuities in the image which in turn are caused by gray-level changes in the image function $g(x,y)$. Because gray-level changes are caused by changes in illumination, discontinuities in images and surfaces are related. All surface breaklines will show up as edges in the images. The contrary does not hold: there may be edges (markings) that do not correspond to breaklines, for example, shadows. Thus, breaklines are under normal circumstances a subset of edges. The compelling conclusion is that by matching edges the breaklines of the surface are found.

Breaklines are a notorious problem in area-based matching because the correlation windows are badly distorted and the shape parameters cannot compensate for the discontinuity. Thus, area correlation performs badly precisely in those areas which contain the most interesting information about surfaces.

Another strong reason for matching edges is the relationship between surface discontinuities and object boundaries. It is quite likely that matched edges correspond to object boundaries. This is of great importance for object recognition.

[11]A version of ETL's program as described in Norvelle (1991).

[12]We can think of representing the surface in scale space, for example, by smoothing the surface function $z(x,y)$ by a Gaussian. It is then possible to reconstruct the surface from this representation. Moreover, this scale-space representation can be obtained from matching the appropriate image scale space.

Implementation.

In addition to the $\psi - s$ edge matching method used in the orientation module, we have developed other approaches. Again, zero crossings are determined on every level of the image pyramid. However, the method is not limited to zero crossings; it may well be that other edge operators would render edges which are even better suited for that purpose.

As explained in Li and Schenk (1990b) in detail, zero crossings in one image are matched by a modified least-squares matching method. From Figure 7, we see that the correlation window is centered on a zero-crossing position. The search window in the other image is placed according to the current DEM^i. Its dimension is derived from the expected depth range in that area. Because the zero crossing corresponds to a potential breakline, we model the matching window accordingly by introducing two separate planes whose common edge crosses the window. This amounts to shaping the matching windows by two separate affine transformations to the correlation window.

The matching results are checked for figural continuity. Because zero crossings indicate a discontinuity in the direction perpendicular to the contour, it is unlikely that a discontinuity occurs simultaneously along the contour. The principle of continuities along discontinuities (see Marr, 1982) provides a useful consistency check.

Experimental Results.

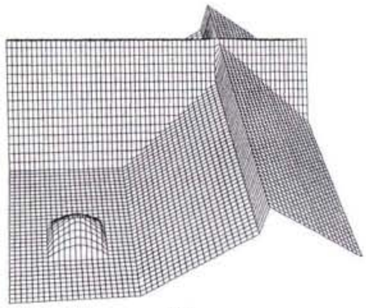
Figure 8 shows an example of matching edges. The two image patches, size 512 by 512 pixels, are from model "Munich" with the finest resolution of 4096 by 4096. With trees, shrubs, tall buildings, occlusions, and shadows, it represents a typical large-scale urban scene which poses many difficulties for matching. The major steps in matching zero crossings with a modified least-squares method are presented in Figure 8 and explained in the caption. The reader is encouraged to view the matched edges stereoscopically. They give a vivid impression about the complexity of the raw object space and an appreciation of the power of edge matching. It should also be stressed that many of the matched edges correspond to object boundaries.

CONCLUSION

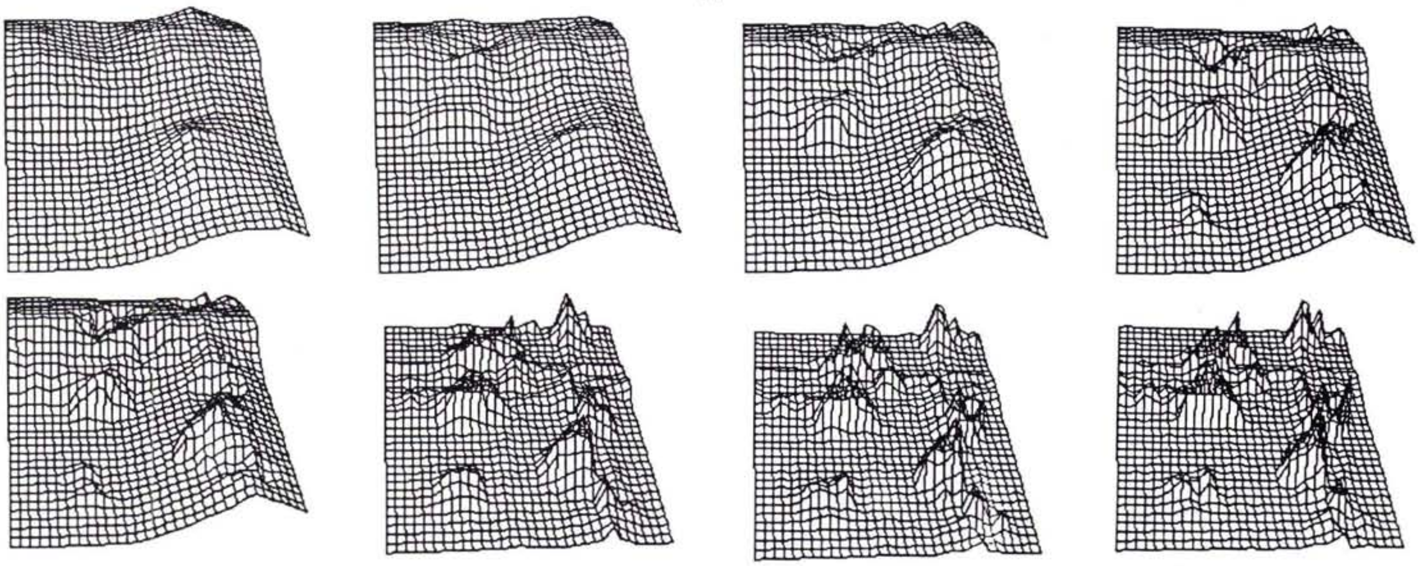
Research efforts in digital photogrammetry are beginning to bear fruit. Together with advances in hardware and system software, it is now possible to develop operational systems. Furthermore, we are witnessing the advent of softcopy photogrammetric workstations—the pendant to the analytical plotter, but now dealing with digital imagery.

As always when new technologies and methods emerge, expectations are raised that photogrammetric processes can be automated. This will be especially true for softcopy workstations: potential users may push expectations beyond what can be delivered in the foreseeable future. We argue that major progress toward autonomous softcopy workstations depends more on advances on the conceptual level rather than on the refinement of system components such as hardware, system software, and low-level algorithms. This entails adopting the computer vision paradigm for photogrammetric applications, understanding the methods and tools that are available, and developing skills which enable researchers to analyze and understand precisely how human operators perform their tasks to generate photogrammetric products.

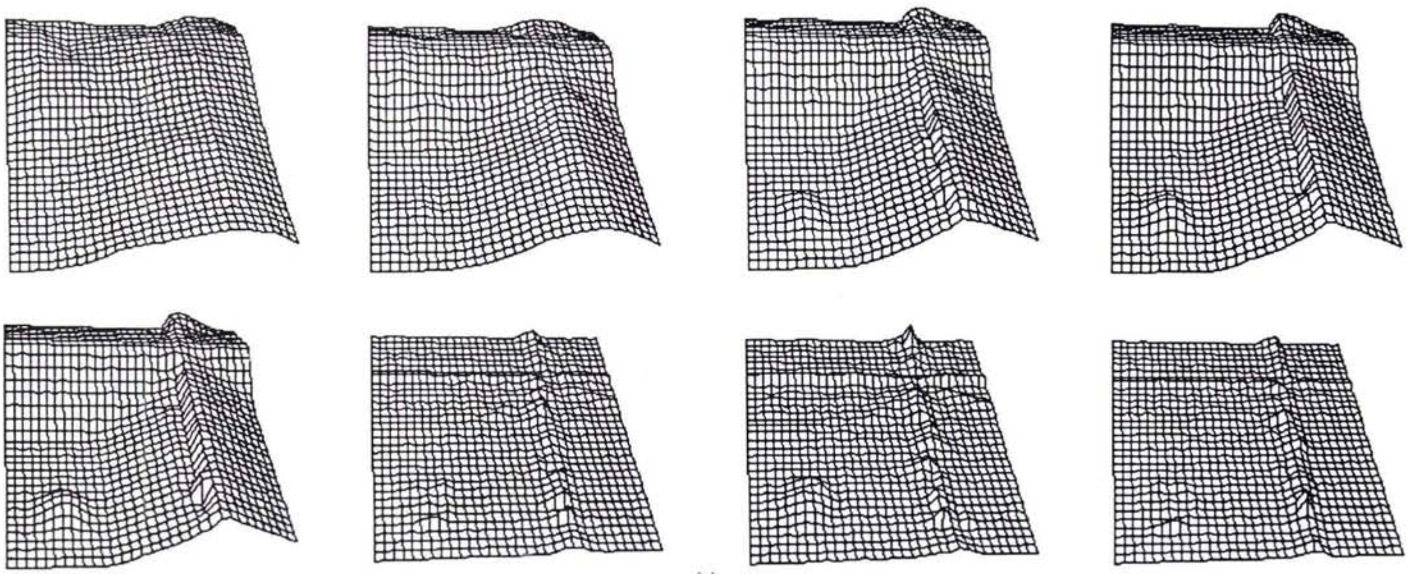
The problem of increasing the degree of automation is the problem of making implicit information, "buried" in the raw image, explicit. Photogrammetrists are perhaps too entrenched in the micrometre and sub-pixel world, deeply if not exclusively concerned with accurate point positioning. We have to shift the focus of our attention from data-driven, pixel-to-pixel operations to a more symbolic processing of abstract representations.



(a)



(b)



(c)

FIG. 6. Comparison between image space and object space matching. (a) Synthetic surface, based on which synthetic images are generated by adding randomly generated gray levels to the synthetic surface. (b) Results of matching in image space (conventional area correlation with shape parameters). The first surface in the sequence of six is the initial approximation, the second is the result of matching at a resolution of 8 by 8, the third at 16 by 16, and so on to the sixth surface that corresponds to the finest resolution of 128 by 128 pixels. The next row with three surfaces shows the differences between the surfaces obtained by matching on the 32 by 32, 64 by 64, and 128 by 128 level with the true surface. Note that matching in image space cannot catch the little hump on the surface and also has lots of difficulties along the breaklines. (c) Results of matching in object space. The sequence of representations is the same as in (b). On every level of the image pyramid the two synthetic images are warped to the surface obtained from the previous level. Matching was performed with the same area correlation algorithm, without shaping parameters, however. Note that the differences from the true surface (last row) are much smaller as compared to ordinary image space matching.

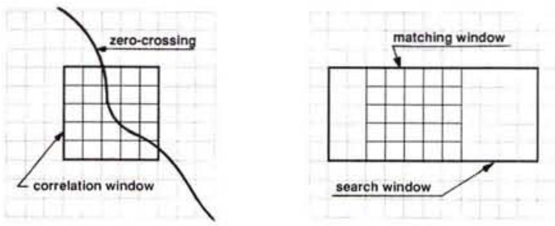


FIG. 7. Principle of matching zero crossing with modified least-squares method.

With the automatic orientation and the surface reconstruction problem, we have described a step in that direction, admittedly a small step considering the remaining problems.

ACKNOWLEDGMENTS

Funding for this research was provided in part by the NASA Center for the Commercial Development of Space Component of the Center for Mapping at The Ohio State University. We would like to thank Raye Norvelle of the U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Virginia, for the interactive matching algorithms used in this research.

REFERENCES

Ackermann, F., 1984. Digital Image Correlation: Performance and Potential Applications in Photogrammetry, Photogrammetric Record, 11(64):429–439.

Aloimonos, J, and D. Shulman, 1989. *Integration of Visual Modules*, Acamdemic Press, Inc.

Al-Tahir, R., Ch. Toth, and T. Schenk, 1990. Automatic Ground Control Point Recognition with Parallel Associative Memory. *Proceedings of the ACSM-ASPRS Fall Convention*, pp. B-10–B-18.

Ballard, D.H., and Ch. M. Brown, 1982. *Computer Vision*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Barrow, H.G., 1982. Computational Vision. *Proceedings IEEE 69*, No. 5, pp. 572–595.

Dond, U.R., and J.K. Aggarwal, 1989. Structure from Stereo—A Review, *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-19(6):1489–1510.

Doorn, B.D., 1991. *Multi-Scale Surface Reconstruction in the Object Space*, Dissertation, Department of Geodetic Science and Surveying, The Ohio State University.

Doorn, B.D., P. Agouris, R. Al-Tahir, A. Stefanidis, and O. Zilberstein, 1990. *Digital Stereo Matching: In Perspective*, Technical Notes in Photogrammetry No. 10, Department of Geodetic Science and Surveying, The Ohio State University.

Ebner, H., and C. Heipke, 1988. Integration of Digital Image Matching and Object Surface Reconstruction, *Proceedings of ISPRS Commission III*, Kyoto, Japan.

Förstner, W., 1986. Digital Image Matching Techniques for Standard Photogrammetric Applications, *Proceedings of the ACSM-ASPRS Annual Convention*, Vol. 4, pp. 210–219.

Grimson, W.E.L., 1984. Binocular shading and visual surface reconstruction, *CVGIP*, pp. 18–44.

Grün, A., and E.P. Baltsavias, 1988. Geometrically Constrained Multiphoto Matching, *Photogrammetric Engineering & Remote Sensing*, 55(5):581–586.

Helava, U.V., 1988. Object-Space Least-Squares Correlation, *Photogrammetric Engineering & Remote Sensing*, 54(6):711–714.

Horn, B.K.P., 1983. Non-correlation methods for stereo matching, *Photogrammetric Engineering & Remote Sensing*, 49(4):535–536.

———, 1986. *Robot Vision*, The MIT Press, Cambridge.

Li, J-Ch., 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.

———, 1990. An Accurate Camera Calibration for the Aerial Image Analysis, *Intern. Conf. on Pattern Recognition*, 1:207–209.

Marr, D., 1982. *Vision*, W.H. Freeman and Company, New York.

Norvelle, R., 1991. *Digital Photogrammetric Compilation Package*, U.S. Army Engineer Topographic Laboratories, Autonomous Technologies Divisions.

———, 1992. Stereo-Correlation: Window Shaping and DEM Corrections, *Photogrammetric Engineering & Remote Sensing*, 58(1):113-117.

Poggio, T., 1988. MIT Progress in Understanding Images, *Proceedings DARPA Image Understanding Workshop*.

Poggio, T., V. Torre, and C. Koch, 1985. Computational vision and regularization theory, *Nature*, (317):214–319.

Rosenholm, D., 1987. Multi-Point Matching Using Least-Squares Technique for Evaluation of Three-Dimensional Models, *Photogrammetric Engineering & Remote Sensing*, 53(6):621–626.

Schenk, T., J.C. Li, and Ch. Toth, 1990a. Zur automatischen Orientierung von digitalen Bildpaaren, *Zeitschrift für Photogrammetrie und Fernerkundung*, (4):1213–1218.

———, 1990b. Hierarchical Approach to Reconstruct Surfaces by Using Iteratively Rectified Images, *Proceedings of the International Society for Photogrammetry and Remote Sensing Symposium*, Commission V, 28(5/1):464–470.

———, 1991. Towards an Autonomous System for Orienting Digital Stereopairs, *Photogrammetric Engineering and Remote Sensing*, 57(8),1057–1064.

Stefanidis, A., P. Agouris, and T. Schenk, 1991. Aspects of Accuracy in Automatic Orientation, *Proceedings of the ASPRS-ACSM Annual Convention*, 5:334–343.

Terzopoulos, D., 1986. Regularization of inverse problems involving discontinuities, *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-8(1):413–425.

Toth, Ch., and T. Schenk, 1990. Pattern Recognition with Parallel Associative Memory, *Proceedings of the International Society of Photogrammetry and Remote Sensing Symposium*, 28(5/1):558–563.

Wrobel, B.P., 1988. Least-Squares Methods for Surface Reconstruction from Images, *Proceedings of the International Congress for Photogrammetry and Remote Sensing, Commission III*, Kyoto, Japan.

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

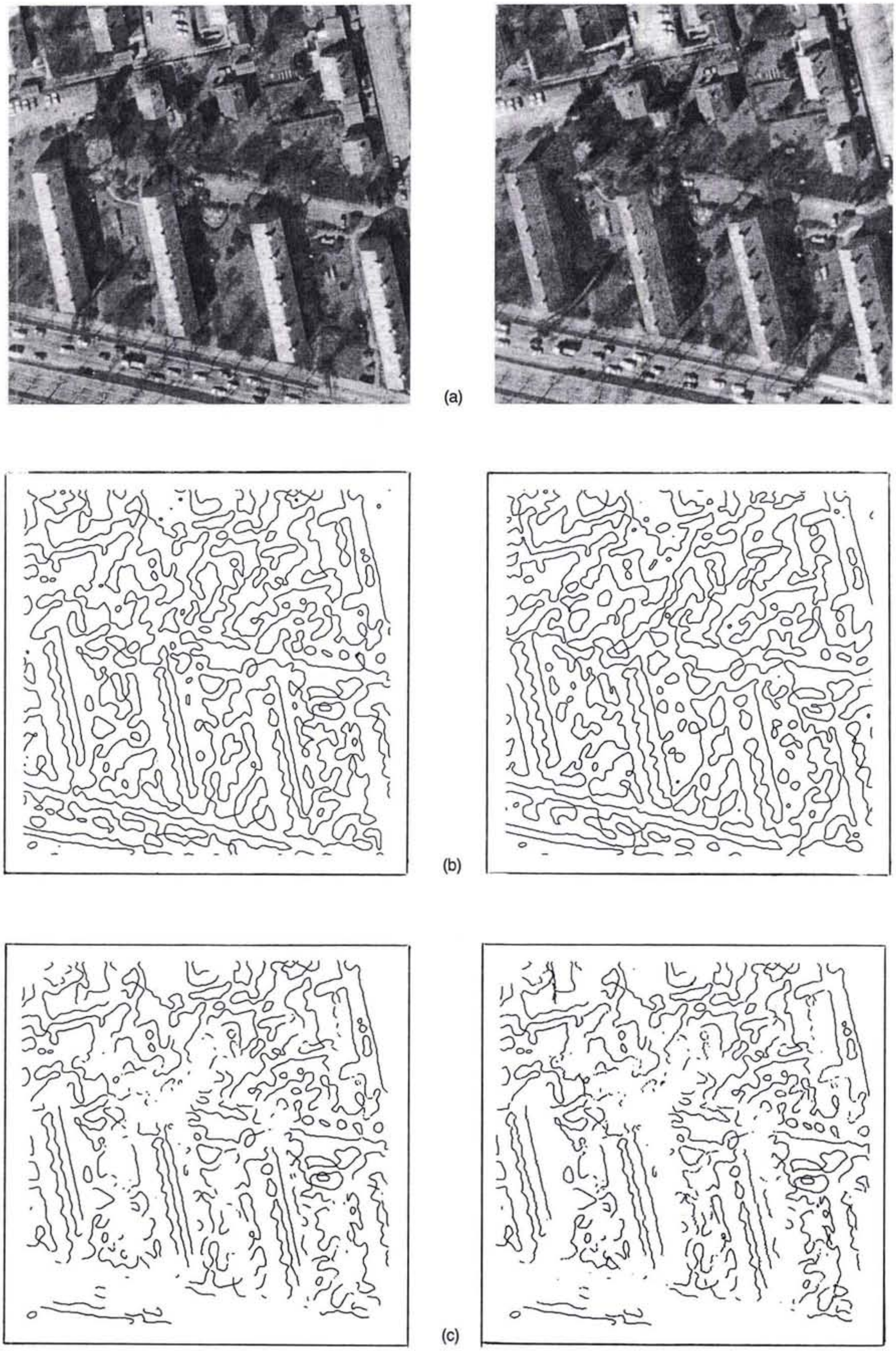


FIG. 8. Results of matching zero crossings with modified least-squares method. (a) Two corresponding image patches from model "Munich."