Stereo Correlation: Window Shaping and DEM Corrections

F. Raye Norvelle
Research Institute, U. S. Army Engineer Topographic Laboratories, Fort Belvoir, VA 22060

ABSTRACT: The accuracy of digital elevation models (DEMs), extracted automatically from terrain images using digital correlation methods, can be enhanced by performing simple, linear, window shaping operations within the correlation method. Window shaping reduces the size and shape differences between small corresponding areas on stereoimages and contributes to a more accurate determination of stereoimage correspondence and, consequently, DEM values.

The slow and tedious process of DEM editing can be expedited by taking advantage of the geometric relationship between orthophotographs and the DEMs that are used to produce them. For example, a pair of digital orthophotos, supposedly identical, can be generated from each of the original stereoimages that were used to create the DEM in the first place. The mismatches in the pair of orthophotos can be measured automatically with a digital correlation method, converted to equivalent elevation errors, and used to correct the original DEM.

INTRODUCTION

MOST PHOTOGRAMMETRIC SOFTCOPY WORKSTATIONS provide a digital correlation method to automatically extract elevation data (DEMs) from stereopairs of digital terrain images. In the correlation process, a small patch of gray shade data taken from one digital image is compared in some manner for correspondence with a patch of gray shade data from its stereomate. The goal is to determine a "match" point between the two images. Ideally, the two patches of imaged terrain should have the same size and shape. This condition is generally not met because sloping terrain surfaces, when photographed from different camera exposure stations, will be recorded differently on the stereophotos. The resulting dissimilarities in corresponding image areas will limit the accuracy of the digital correlation method in determining the coordinates of match points and subsequent DEM values.

The U.S. Army Engineer Topographic Laboratories (USAETL), Fort Belvoir, Virginia, has conducted research in digital correlation techniques since the early 1970s. One important conclusion from the research is that some effort must be made to minimize the geometric differences between the stereoimages before, or as part of, the correlation computation to achieve the maximum correlation accuracy. This effort is referred to herein as "window shaping."

Some correlation methods do not perform a window shaping operation, possibly because its value is not clearly appreciated. One purpose of this paper, therefore, is to demonstrate how window shaping can significantly enhance correlation accuracy.

The DEM values obtained by an automatic digital correlation process are seldom error-free. Interactive editing of the derived data set can be a tedious and time-consuming exercise. The second purpose of this paper is to show how DEM corrections can be expedited by taking advantage of the inherent geometric relationship between pairs of orthophotos[1], created from stereoimages, and the DEMs used in the orthophoto transformation process. This DEM correction technique is termed the "Iterative Orthophoto Refinements (IOR)" method at USAETL. It is similar in concept to that reported by Schenk (1989).

This paper presents experimental results obtained using three basic methods. The following section describes these methods starting with "match," USAETL's correlation method. This is followed by a description of point prediction and window shaping methods, and, finally, by a description of the IOR technique. The next section reports the results of tests in two parts. The first part concerns tests made with and without window shaping operations. The second gives the results of the IOR method when used, *in lieu* of manual or interactive editing methods, to refine a DEM. The results are discussed and conclusions presented in the final two sections of this paper.

[1]All references to "orthophoto" products imply the digital image equivalent.

DESCRIPTION OF METHODS

MATCH ROUTINE

Early digital correlation research by USAETL led to the development of a correlation program termed the "Digital Interactive Mapping Program(DIMP)." It was written in FORTRAN for a CDC mainframe computer. For many years, DIMP served as USAETL's research tool for conducting correlation experiments and for generating elevation data in support of other research projects. In the late 1980s, the DIMP software was reprogrammed, with some enhancements, in the C-language for a Silicon Graphics Iris 4D/85GT engineering workstation and renamed MATCH. A complete description of MATCH, its parameter settings and options, is beyond the scope of this paper. The reader is referred to Norvelle (1981) which describes DIMP in detail and is generally applicable to MATCH.

MATCH is an area-based digital correlation method that operates in image space and is referenced to an evenly spaced grid of points on the left image of a stereopair. That is, for equally spaced image points (grid) on the left image, MATCH attempts to determine the coordinates of the corresponding points on the right image. A small, reshaped patch of gray shade data (window), centered about the point in question on the left image, is extracted and compared (correlated) with a larger patch (search area) taken from the right image. The search area is centered about a point on the right image where maximum correspondence between the window and search area is predicted to occur. The size of the search area is a function of the window size and the limits ("pull-in" range) over which a search for the match point is to be performed.

The window is placed, in template fashion, over equal sized subsets of the search area at all possible unique positions. At each position, the degree of correlation between the gray shades in the window and those of the search area is computed using the normalized cross correlation coefficient of Equation 1.

$$R = \sigma_{xy}/(\sigma_x * \sigma_y) \tag{1}$$

0099-1112/92/5801–111$03.00/0

In Equation 1, σ_x and σ_y are the standard deviations of the gray shade data in the window (x) and search area subsets (y). The σ_{xy} is the covariance between the two patches of gray shade data. The R values obtained for the subsets in the search area, and the pixel positions where they occur, are termed the "correlation function." The correlation function is analyzed to determine the position of the maximum correlation value. This maximum value, and R values to the left and right of the maximum, are used to define the equation of a parabola that passes through them. The equation is evaluated to determine (1) its peak value (RMAX), (2) the sharpness of the peak (CX), and (3) where the peak occurs (DX) relative to the coordinates of the predicted match point. These three values are combined in Equation 2 to define a measure of confidence for each match point.

$$\text{TEST} = (\text{RMAX} * \text{CX})/(1 + (\text{DX}/4)^2) \tag{2}$$

The confidence value, TEST, is compared to a minimum threshold value, specified by the user, to judge if the derived match point is acceptable or not. The degree of confidence in a match point dictates how it is treated by other operations in MATCH. For example, if the point is judged to be unacceptable (TEST is too low), its coordinates will be replaced by straight line interpolation across points that are acceptable.

Anaglyphic techniques are used both to display the stereoimages in three dimensions (3D) and to display the results of the correlation process in 3D over the stereoimages. That is, after each row of correlations is completed, red and cyan lines are drawn column-wise on the monitor between corresponding match points on each image. Red lines connect match points on the left image and cyan lines connect those on the right image. When viewed through the anaglyphic glasses, the lines appear as 3D profiles overlaying the 3D stereomodel area. If the match points are correct, the profiles will appear to rest on the terrain surface. They will "float" above or below the surface if they are incorrect. This technique helps the user evaluate correlation accuracy without recourse to ground truth. In cases where the correlation process appears to have difficulty, the operator can halt the program and make manual measurements of match points or change various operating parameters to get the process back on track.

As previously stated, point prediction and window shaping operations are used in MATCH. Detailed descriptions of these are given in the section that follows.

Point Prediction and Window Shaping

Figure 1 shows the relationship between the positions of evenly spaced points on the left image and the corresponding match points on the right image. On the left image, the points marked by crosses are spaced in the column and row directions by a constant "NCRE" (for incremental) number of pixels. The positions of the match points on the right image are not equally spaced because they reflect the parallax between images caused primarily by terrain relief.

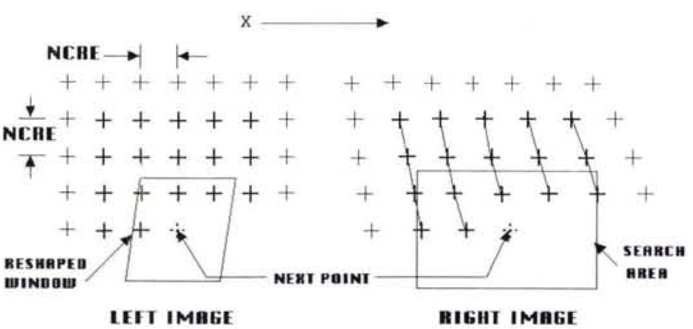


Fig. 1. Match points on the left and right images used for point prediction and window shaping.

For point prediction purposes, the coordinates of crosses on the left image are functionally related in the X (column) direction to those on the right image according to Equation 3.

$$X' = a + bX + cY \tag{3}$$

The X and Y values are the coordinates of a match point on the left image. The X' value is the coordinate of the corresponding match point on the right image. The Y' coordinate of the next point in the search area is taken to be the same as that of the previous point in the same row, because very little change in the y direction is encountered from one match point to the next.

The coordinates of those points denoted by bold crosses are used in a least-squares solution to compute the coefficients (a, b, and c) of Equation 3. Once the coefficients have been determined, the equation is used to predict, by extrapolation, the X' coordinate for the next point to be matched on the right image.

In Figure 1, there is a considerable difference between the size and shape of the pattern of match points on the two images. The two patterns are related, however, by the "b" (scaling) and "c" (shear) coefficients of the point prediction Equation 3. In MATCH, the window is shaped to agree with the search area because it is smaller and has to be shaped only once per match point computation. This means that an inverse form of Equation 3 is used for shaping as follows:

$$X = (j - ci)/b \tag{4}$$

In Equation 4, i and j are the row and column indices of a rectangular array of pixels with a size equal to that of the desired reshaped window. For each i and j value, an X coordinate is obtained that dictates where on the left image a gray shade value should be extracted to form the reshaped window. In Figure 1, the X values will fall inside the parallelogram shown on the left image, because it is the inverse shape of the match point positions on the right image. Because X is a decimal pixel value, the gray shades are resampled from the left image using bilinear interpolation methods. No shaping is performed in the y direction because y-axis dissimilarities are very small on stereoimages typical associated with topographic mapping.

The point prediction and window shaping operations, as expressed by the linear Equations 3 and 4, are based on the assumption that local areas of the terrain surface are flat and have a uniform slope. There are certain pitfalls involved with this assumption. For example, if the next point to be matched falls precisely on a sharp ridge line (breakpoint), half of a properly shaped window will be on one side of the slope while the other half will be on the opposite slope. Equation 4 is not completely appropriate in such a case. Also, if the selected pixel spacing between match points is too large, the assumption of linearity will not be valid. That is, previously matched points will be too far apart to be "local" and cannot be used in a meaningful manner for point prediction and window shaping. The point to emphasize here is that, while simple, linear point prediction and window shaping operations can be used successfully, strategies must be available in the correlation method to detect nonlinear terrain conditions and compensate for their effects.

The "Iterative Orthophoto Refinements (IOR)" Method

Operating on an SGI 4D/85GT workstation, MATCH can typically produce 360,000 to 1 million match points in an hour, depending on window size, "pull-in" range, and other factors. Many hours may be required to carefully inspect and edit the data using conventional manual and semiautomatic techniques. There is, therefore, great benefit to be derived from new editing methods that can perform at speeds commensurate with that of the correlation process. In this regard, USAETL has experimented with a new DEM correction technique it terms the "Iterative Orthophoto Refinements(IOR)" method.

The IOR is an experimental technique for correcting DEM values based on the geometric relationship between pairs of orthophotos and the DEM used to produce them. Originally, the IOR technique was conceived as an editing tool. Experiments indicated, however, that the method was so successful that it can be used to automatically generate, rather than just edit, high resolution DEMs. Consider the steps used in Figure 2 to generate a DEM. Here an experiment is outlined where, as a first step, the stereoimages to be correlated are reduced in scale and resolution by a factor of four, for instance, using simple pixel averaging. Because of (1) their smaller size, (2) the reduction in x parallax by scaling, and (3) the elimination of high frequency noise by pixel averaging, the reduced images can be correlated very rapidly and successfully by MATCH. The resulting match point coordinates are used to generate an approximate DEM. The approximate DEM is then used to generate orthophotographs of the original, full-resolution stereoimages.

If the DEM is correct, and if the orientation parameters of the original photographs are accurate, the pair of digital orthophotos should be nearly identical. Because the DEM is approximate, there will be image mismatches (x parallax) between the orthophotos. The mismatches can be detected automatically by a digital correlation method, converted to equivalent elevation errors and used to refine the original DEM. This process is performed in an iterative fashion until there are no significant corrections to be made to the latest version of the DEM. Note that, on each iteration, new orthophotos are generated using the original stereoimages and not by correcting the previous versions of the orthophotos. This eliminates the image degradation problem that can occur when images are resampled repeatedly.

Because each pixel of the digital orthophotos represents a known ground-sample distance in metres, mismatches, measured in pixels on the orthophotos, can be converted to equivalent horizontal displacements in a ground coordinate system. The amount of elevation error required to produce the displacements can be determined approximately by dividing the displacement (metres) by the base-height ratio of the original stereoimages that were used to generate the DEM and the orthophotos. This approximate approach is based on the assumption that the terrain is level. More rigorous methods must account for the actual slope of the terrain in the computations. In either case, the method is iterative because the true slope of the ground is not known until the correction process in completed.

TEST RESULTS

WINDOW SHAPING RESULTS

To demonstrate that window shaping enhances correlation accuracy, a stereopair of images was correlated with and without window shaping. "Point prediction" was used in both cases but, in the case without shaping, the "b" and "c" coefficients of Equation 4 were set to 1 and 0, respectively (no shaping performed). The test images are of the San Bernardino Mountains in California and have a scale of 1:40,000. They were digitized with a 50-micrometre spot size and, consequently, each pixel has a ground resolution of about 2 metres. A match point was determined for every 5th pixel on the left image which corresponds approximately to a 10-meter spacing on the ground. The tests were conducted on 1280- by 1024-pixel subsets of the original images and resulted in 32505 match points. The mountainous terrain is very rugged with some slopes approaching 45 degrees and with changes in slopes (ridge lines and valleys) of up to 70 degrees.

Preliminary tests indicated that a window 15 by 15 pixels in size produced the better correlation results on the test images. It was used in the window shaping tests along with a search area of 17 by 27 pixels. This combination allows refinements to the predicted match point of up to ± 1 pixel in the y-parallax direction and ± 6 pixels in the x direction. The large 6-pixel "pull-in" range was used to allow the correlation process to negotiate the numerous, nonlinear terrain areas characterized by sharp ridges and valleys. Only a 1-pixel correction was used in the y direction because, over a 5-pixel spacing between correlation points, y-parallax changes are very small. Other MATCH parameters were required for this test and they were the same for both cases. The point to emphasize is that the differences in the results are solely a function of whether window shaping was performed or not. The results of the tests are shown in Table 1.

The first two columns of numerical values are the number of match points that were judged to be either acceptable or unacceptable, based on the confidence values computed with Equation 2. RMAX is the average maximum correlation coefficient for all acceptable match points. It can range in value from zero to one units. The DX and DY values, in units of pixels, are the average of the corrections (absolute values) applied to the coordinates of the predicted match points.

The statistical results given in Table 1 are not adequate for drawing conclusion about the nature of the differences in the two sets of correlation data. The differences are clearly revealed, however, when the match points are displayed in anaglyphic fashion as 3D profiles overlaying the 3D stereomodel. Alternately, synthetic shaded relief images can be generated from the resulting DEM and used, as in Figure 3, to provide a visual representation of correlation results.

Figures 3a and 3b are gray shade relief images of elevation values (DEM) spaced horizontally at 10-meter intervals. The "sun" is in the upper-left or northwest corner at an altitude of 60 degrees from the vertical. The DEM used to create the left image, Figure 3a, was derived from the match points obtained with the window shaping feature. Figure 3b is from data acquired without window shaping. The "shadows," caused by terrain masking, give a good visual indication of the differences that exist between the DEM's represented in each figure.

- REDUCE FULL-SCALE IMAGES BY 4X.
- CORRELATE REDUCED IMAGES ("MATCH").
- COMPUTE APPROXIMATE DIGITAL ELEVATION MODEL (DEM).
- CREATE ORTHOPHOTOS OF FULL-SCALE INPUT IMAGES.
- CORRELATE ORTHOPHOTOS PAIRS TO DETERMINE MISMATCHES DUE TO ELEVATION ERROR ("MATCH").
- UPDATE DEM BASED ON MISMATCHES.

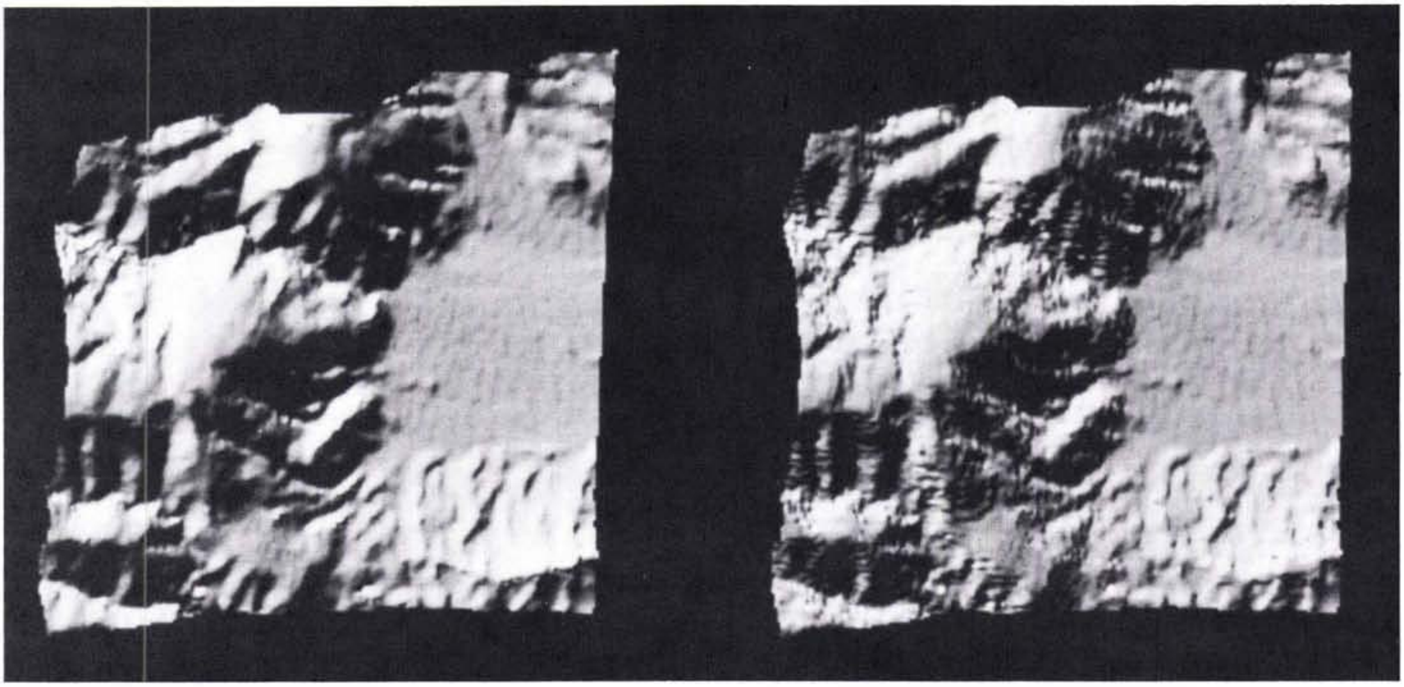
FIG. 2. Steps performed in the "Iterative Orthophoto Refinements (IOR)" method.

IOR RESULTS

The IOR method was applied to the same stereoimage data sets that were used in the window shaping tests. The match

TABLE 1. STATISTICAL RESULTS OF THE WINDOW SHAPING TESTS.

	Acceptable Pts		Average		
	Yes	No	RMAX	DX	DY
with shaping	32180	325	0.700	0.326	0.144
without shaping	30875	630	0.603	0.492	0.164



(a) (b)

FIG. 3. Shaded relief images of a DEM generated (a) with window shaping and (b) without window shaping.

points derived by correlation on reduced (by $4\times$) versions of the stereoimages were used to create an approximate DEM. These elevation values are presented in Figure 4a as a synthetic shaded relief image. Two iterations with the IOR method resulted in a corrected DEM that is represented in Figure 4b. For comparison purposes, elevations near the edges of the DEM area were not corrected. Consequently, the finished DEM appears to be an insert in Figure 4b, surrounded by uncorrected DEM values. The shaded relief images of Figure 4 were generated with the "sun" in the upper-left corner and at an altitude of 50 degrees from the vertical.

On the first iteration, a 15- by 15-pixel window and a ± 2-pixel "pull-in" range (in x) was used to detect mismatches between the pair of orthophotos generated with the approximate DEM. On the final (second) iteration, the new orthophotos were nearly identical and, consequently, could be correlated successfully with a 9- by 9-pixel window and a ± 1-pixel "pull-in" range.

Once the approximate DEM was obtained, approximately 30 minutes were required to perform the two IOR iterations that resulted in the corrected DEM. For comparison purposes, manual and interactive techniques were used to edit the equivalent DEM shown in Figure 3a. This more conventional editing process required five hours to complete.

DISCUSSION OF RESULTS

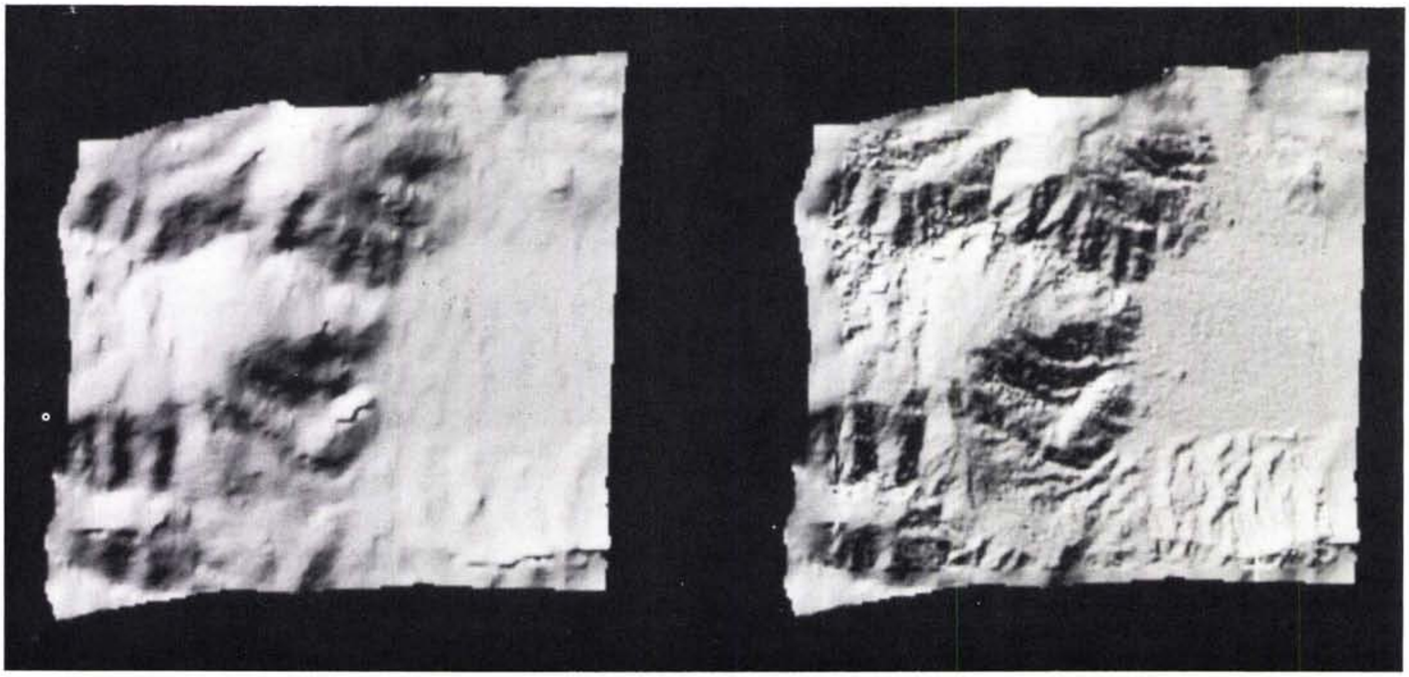
WINDOW SHAPING

In the MATCH routine, performance is rated by the number of unacceptable points obtained, the magnitude of the mean correlation coefficient (RMAX), and the size of the corrections (DX,DY) applied to the coordinates of the predicted match points. A low number of unacceptable points and a high RMAX value indicate high correlation quality. Low DX and DY terms indicate that the point prediction operation was performed correctly and, therefore, that the assumption of terrain linearity was valid. As shown in Table 1, the results obtained with window shaping are far superior to those without. This simply attests to the fact that the window shaping operation was highly successful in removing stereoimage dissimilarities and, thereby, providing the more accurate correlation results. The superior accuracy of the window shaping results is even more evident when the match point data are viewed as 3D profiles overlaying the 3D stereomodel. With window shaping, the profiles accurately conform to the stereomodel surface. In contrast, the profiles obtained without shaping have a distinctive "staircase" appearance. Figure 3 dramatically illustrates this effect.

In Figure 3a, the DEM representation of match points obtained with shaping has the smooth and continuous appearance that is expected, and verifiable, from the stereoimages. The DEM values used for Figure 3b, however, approximate sloping surfaces in a manner similar to a step function. This "staircase" effect is created when the correlation process fails to detect small x-parallax changes from one match point to the next. Consequently, equivalent elevation values of adjacent match points define a level surface (the "step" of the "staircase"). Eventually, the difference between the predicted and the actual location of the next point to be matched (x parallax) becomes large and detectable, resulting in an abrupt elevation change (the "riser" of the "staircase"). The "staircase" effect further illustrates that size and shape dissimilarities between small imaged areas on a stereopair preclude precise match points results.

IOR METHOD

It is obvious, by inspection of Figures 4a and 4b, that the IOR technique was very successful in adding high resolution corrections to the original, approximate DEM values. The accuracy of the corrections was verified by 3D inspection of the final pair of orthophotographs generated from the finished DEM data. When viewed in 3D as a stereopair of images, the mismatches (x parallax) produce false, exaggerated "terrain" heights which are readily detected. No false terrain relief was noted, except in a



(a) (b)

FIG. 4. Shaded relief images of (a) the initial, reduced resolution DEM and (b) after correction by the IOR method.

few occluded areas. There can be circumstances, however, where mismatches between the pair of orthophotos do not appear in 3D as "relief" displacements. For example, monotone areas (such as water bodies) will appear to be the same on each orthophoto, even though there may be errors in the corresponding DEM data. In such cases, the user must select another means, such as manual editing, to correct the DEM independent of the orthophotos.

The success of the IOR technique can be attributed to several factors. Mainly, the transformation of stereoimages to orthophotos removes the major geometric dissimilarities that exist between the original stereoimages. Consequently, correlation methods are more successful on the orthophotos than on the original images. Furthermore, because the pair of orthophotos are nearly identical, a small window size and "pull-in" range can be used successfully in the correlation process. This provides a fast compilation rate but, more importantly, the small window includes only terrain shapes local to the desired match point. Consequently, more subtle and relevant mismatches can be detected, resulting in high resolution DEM corrections.

The high resolution DEM corrections obtained by the IOR method can be illustrated by comparing Figure 4b with Figure 3a, keeping in mind that the gray shade relief images were produced using a different "sun" altitude. The data represented in Figure 3a was acquired in the conventional manner. That is, correlation was performed on the full-resolution stereoimages using window shaping and a 15- by 15-pixel window. In the IOR experiment, a window size of 9 by 9 pixels was used on the final iteration to detect mismatches. The small window in the IOR experiment was instrumental in producing high resolution DEM corrections. The larger window used in the conventional approach produced smoother results. Because larger windows cover more, and probably varying, terrain areas, they tend to produce an averaging or smoothing effect over that area.

The noisy appearance of the elevation data in the valleys of Figure 4b is actually a true representation of tree stands. This level of detail was not detected with the conventional approach.

Thirty minutes were required to correct the approximate DEM using the IOR method. In contrast, manual and semiautomatic editing of the equivalent amount of match point data required five hours. This 10 to 1 improvement in speed comes about because the IOR method is essentially an automatic process and is not, therefore, slowed by manual and interactive operations.

CONCLUSIONS

- Correlation accuracy is significantly enhanced when window shaping operations are used to minimize the stereoimages dissimilarities caused by sloping terrain surfaces.
- The "Iterative Orthophoto Refinements (IOR)" technique is a highly effective, accurate, and fast method for generating or editing digital elevation models (DEMs).

ACKNOWLEDGMENTS

The stereoimages, and associated orientation data, of the San Bernardino Mountains, California, were furnished by the U.S. Geological Survey, Western Mapping Center, Menlo Park, California.

REFERENCES

Norvelle, F. Raye, 1981. *Interactive Digital Correlation Techniques for Automatic Compilation of Elevation Data*, ETL-0272, U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Va.

Schenk, A. F., 1989. *Determination of DEM Using Iteratively Rectified Images*, Photogrammetry Technical Report No. 3, Department of Geodetic Science and Surveying, The Ohio State University, Columbus, Ohio