Dr. Marsha Jo Hannah*
Institute for Advanced Computation
NASA/Ames Research Center
Moffett Field, CA 94035

Error Detection and Correction in Digital Terrain Models†

Algorithms to detect and correct errors focus on the use of constraints on both the allowable slope and the allowable change in slope in local areas around each point.

Introduction

Digital terrain models are playing an increasingly important role in the production of topographic maps, in urban planning, in highway routing, and in many other map-related tasks. These terrain models can be produced in several ways; our research has concentrated on elevation data resulting from digital correlation of sub-areas from stereo imagery (Hannah, 1974; Crombie, 1976; Panton, 1978).

A significant problem in correlation-derived digital terrain models is the introduction of errors into the elevation data when the stereo correspondence algorithm produces mismatches. These mismatches can result from a variety of conditions, including sensor noise, low contrast in portions of the images, relief-induced distortions between the images, and the presence of ambiguities due to identical objects or highly periodic textures on the terrain. It is not yet feasible for a correlation algorithm to handle all of these difficult situations without error. For this reason, post-processing techniques have been sought to detect and correct errors which occurred in the correlation process. Work to date has centered on the fitting of polynomials to the data (Jancaitis and Junkins, 1973), filtering in both the spatial and frequency domains (Johnson, 1978), and other global techniques.

Abstract: *Digital terrain models produced by computer correlation of stereo images are likely to contain occasional gross errors in terrain elevation. These errors typically result from having mismatched sub-areas of the two images, a problem which can occur for a variety of image- and terrain-related reasons. Such elevation errors produce undesirable effects when the models are further processed, and should be detected and corrected as early in the processing as possible.*

Algorithms have been developed to detect and correct errors in digital terrain models. These algorithms focus on the use of constraints on both the allowable slope and the allowable change in slope in local areas around each point. Relaxation-like techniques are employed in the iteration of the detection and correction phases to obtain best results.

Global techniques have the drawback that they give identical treatment to all areas of a digital terrain model. Terrain is rarely uniform in roughness, so uniform application of a global technique can produce over-smoothing in rough areas while failing to correct errors in relatively flat areas. Local techniques, on the other hand, have the potential for coping with different terrain types within a model. Also, local techniques can easily incorporate other terrain model information, such as land-use classifications.

† Revised from a poster presentation at the 1979 ASP-ACSM Fall Technical Meeting, Sioux Falls, South Dakota, 17-21 September 1979.

* The author is currently employed at the Lockheed Palo Alto Research Laboratory, Dept. 52-53, Bldg. 204, 3251 Hanover Street, Palo Alto, CA 94304.

Photogrammetric Engineering and Remote Sensing,
Vol. 47, No. 1, January 1981, pp. 63-69.

In September 1978, the Institute for Advanced Computation (IAC) was asked by the U.S. Army Engineer Topographic Laboratories (ETL) to explore local methods for the detection and correction of errors in digital terrain elevation data. The algorithms developed use constraints on the allowable slope and the allowable change in slope around each point in the data set. These measures are applied iteratively to achieve the desired results.

In this paper we describe the basic constraining techniques which we have employed in this research. We also describe the error detection processes we have developed, cover our error correction algorithm, and present some of the results we have obtained.

Terrain Constraints

Our algorithms are designed for grid-based digital terrain models. We assume that a terrain model represents a continuous surface which, for the most part, varies smoothly in elevation, i.e., that any data points causing sharp discontinuities in the elevations or sudden changes in the surface slopes can be suspected of being in error. These we detect by applying constraints to the slopes and to the changes in slope at each point.

Except for boundary points, each grid point [I,J] has eight neighbors, which can be identified by their direction vectors [DI,DJ] relative to [I,J].

Direction (K)	DI	DJ
1	0	1
2	1	1
3	1	0
4	1	−1
5	0	−1
6	−1	−1
7	−1	0
8	−1	1

The slope between point [I,J] and its K-th neighbor is defined as

$$\mathrm{SLOPE}[I,J,K] = (H[I + DI[K],J + DJ[K]] - H[I,J])/\mathrm{DIST}[I,J,K] \qquad (1)$$

where H[I,J] is the elevation datum associated with grid point [I,J], and DIST[I,J,K] is the Euclidean baseplane distance between point [I,J] and its K-th neighbor. Three sets of tests are performed on these slopes: a set of slope constraining tests, a set of local neighbor slope consistency tests, and a set of distant neighbor slope consistency tests.

The slope constraining tests check each of the eight slopes immediately surrounding a point to see that they are not unreasonable, i.e., that the absolute value of each slope does not exceed the slope constraint (also called the SLope THRESHold = SLTHRESH).

$$|\ \mathrm{SLOPE}[I,J,K]\ | < \mathrm{SLTHRESH} \qquad K = 1,2,\ldots,8 \qquad (2)$$

The local neighbor slope consistency tests check the four pairs of slopes crossing a point to see that each pair is consistent, that is, that the absolute value of the difference in slopes in each pair is not greater than a specified amount (Difference in SLope THRESHold = DSLTHRESH). If we define the local change in slope (Difference in SLOPe, Local = DSLOPL) as

$$\mathrm{DSLOPL}[I,J,K] = \mathrm{SLOPE}[I,J,K] - \mathrm{SLOPE}[I - DI[K],J - DJ[K],K] \qquad (3a)$$

then this test requires that

$$|\ \mathrm{DSLOPL}[I,J,K]\ | < \mathrm{DSLTHRESH} \qquad K = 1,2,3,4 \qquad (3b)$$

The distant neighbor slope consistency tests check that the pairs of slopes approaching a point across each of the eight neighbors are consistent. Here we define the distant neighbor change in slope (Difference in SLOPe, Distant = DSLOPD) as

$$\mathrm{DSLOPD}[I,J,K] = \mathrm{SLOPE}[I,J,K] - \mathrm{SLOPE}[I + DI[K],J + DJ[K],K] \qquad (4a)$$

and require

$$|\ \mathrm{DSLOPD}[I,J,K]\ | < \mathrm{DSLTHRESH} \qquad K = 1,2,\ldots,8 \qquad (4b)$$

For points interior to the grid of elevation data, all of these constraints are applied. For points on or adjacent to the edge of the elevation grid, as many as possible of these tests are performed.

Error Detection

Before we can reasonably rectify the data, we must first establish which data points are to be trusted, so that we can give preference to valid data points when forming the adjustments. We begin by calculating an indicator of the correctness of each data point.

Our correctness indicator has two components. The first uses the change-in-slope constraints to produce a slope consistency evaluation of the data. The second uses the slope constraints, forming an elevation consistency evaluation. These two indicators are combined into an overall evaluation of the correctness of each terrain data point.

SIMPLE INDICATORS OF CORRECTNESS

The slope consistency evaluation is based on the application of constraints (or thresholds) to the differences in slopes involving both the local and distant neighbors of a point, as shown in Equations 3 and 4. Inherent in the slope differencing is information about whether the point should have its elevation increased, decreased, or preserved as is. We use this information by converting each of the DSLOPL and DSLOPD quantities to a +1, −1, or 0 value, denoting the need for an elevation increase, and elevation decrease, or no change in elevation, respectively. Thus we form

$$\mathsf{TL}[\mathsf{I,J,K}] = \begin{matrix} 1.0 \text{ if } \mathsf{DSLOPL}[\mathsf{I,J,K}] > \mathsf{DSLTHRESH} \\ -1.0 \text{ if } \mathsf{DSLOPL}[\mathsf{I,J,K}] < -\mathsf{DSLTHRESH} \\ 0.0 \text{ otherwise} \end{matrix} \tag{5a}$$

and

$$\mathsf{TD}[\mathsf{I,J,K}] = \begin{matrix} 1.0 \text{ if } \mathsf{DSLOPD}[\mathsf{I,J,K}] > \mathsf{DSLTHRESH} \\ -1.0 \text{ if } \mathsf{DSLOPD}[\mathsf{I,J,K}] < -\mathsf{DSLTHRESH} \\ 0.0 \text{ otherwise} \end{matrix} \tag{5b}$$

The desired indicator of the correctness of each point is a number between 0.0 (signifying that this point can not be trusted) and 1.0 (meaning that this point appears to be correct). The TL and TD of Equation 5 each produce a number between −1.0 and 1.0, so that the sum of these indicators is a number between −12.0 and 12.0. Taking the absolute value of this sum, divided by 12.0, gives a number in the proper 0.0 to 1.0 range, but with its sense reversed: 1.0 means that all of the tests indicated a bad point, while 0.0 means either that there were no objections or that the positive and negative objections cancelled out. Subtracting this quantity from 1.0 produces the desired range and polarity of values for our indicator of correctness, based on the difference in slopes:

$$\mathsf{RD}[\mathsf{I,J}] = 1 - \frac{1}{12}\left| \sum_{\mathsf{K}=1}^{4} \mathsf{TL}[\mathsf{I,J,K}] + \sum_{\mathsf{K}=1}^{8} \mathsf{TD}[\mathsf{I,J,K}] \right| \tag{6}$$

A similar indicator can be constructed from the slope constraints if we define

$$\begin{aligned} \mathsf{TS}[\mathsf{I,J,K}] &= 1.0 \text{ if } |\ \mathsf{SLOPE}[\mathsf{I,J,K}]\ | > \mathsf{SLTHRESH} \\ &= 0.0 \text{ otherwise} \end{aligned} \tag{7}$$

then form the analog of Equation 6

$$\mathsf{RS}[\mathsf{I,J}] = 1 - \frac{1}{8} \sum_{\mathsf{K}=1}^{8} \mathsf{TS}[\mathsf{I,J,K}] \tag{8}$$

WEIGHTED ITERATION OF CORRECTNESS INDICATORS

In forming the above correctness indicators, we are performing simple averaging of the contributions made by each confidence measure. This is the equivalent of a weighted averaging with all of the weights being equal to 1.0. We know that the confidence measures are not all of the same validity, since the data points which formed them vary in validity. Therefore we should use a weighted averaging in which the confidence measures each have different weights. For the slope constraints, this would be

$$\mathrm{RS2}[\mathrm{I,J}] = 1 - \frac{\sum\limits_{\mathrm{K}=1}^{8} \mathrm{WS}[\mathrm{I,J,K}] * \mathrm{TS}[\mathrm{I,J,K}]}{\sum\limits_{\mathrm{K}=1}^{8} \mathrm{WS}[\mathrm{I,J,K}]} \tag{9}$$

The weight for a slope confidence value should be related to the validity of the data points which produced the slope. Equation 9 is calculating the correctness of one of these data points, so the information needed is the validity of the other point. Thus we use

$$\mathrm{WS}[\mathrm{I,J,K}] = \mathrm{RS}[\mathrm{I} + \mathrm{DI}[\mathrm{K}],\mathrm{J} + \mathrm{DJ}[\mathrm{K}]] \tag{10}$$

In reality, Equations 8 and 9 are both instances of a more general iterative form,

$$\mathrm{RSI}[\mathrm{I,J,N}] = 1 - \frac{\sum\limits_{\mathrm{K}=1}^{8} \mathrm{RSI}[\mathrm{I} + \mathrm{DI}[\mathrm{K}],\mathrm{J} + \mathrm{DJ}[\mathrm{K}],\mathrm{N} - 1] * \mathrm{TS}[\mathrm{I,J,K}]}{\sum\limits_{\mathrm{K}=1}^{8} \mathrm{RSI}[\mathrm{I} + \mathrm{DI}[\mathrm{K}],\mathrm{J} + \mathrm{DJ}[\mathrm{K}],\mathrm{N} - 1]} \tag{11}$$

reminiscent of relaxation techniques (Rosenfeld, *et al.*, 1976).

The RSI[I,J,0] can be initialized to 1.0 (or any other constant) to produce the effect of Equation 8 on the first iteration. However, if other, *a priori* knowledge exists about the relative validity of the data, this information can (and should) be used as the initial weights.

The final slope correctness indicator RS[I,J] is defined to be RSI[I,J,N] after sufficient iterations that significant changes are no longer being made. We have used the criterion that the correctness indicators have convered when 99 percent of them are changing less than 0.05. On the data sets we have processed, this criterion has usually been met by the third iteration.

A similar iterative correctness indicator can be developed for the slope consistency tests. Here we have

$$\mathrm{RDI}[\mathrm{I,J,N}] = 1 - \frac{\sum\limits_{\mathrm{K}=1}^{4} \mathrm{WL}[\mathrm{I,J,K,N}] * \mathrm{TL}[\mathrm{I,J,K}] + \sum\limits_{\mathrm{K}=1}^{8} \mathrm{WD}[\mathrm{I,J,K,N}] * \mathrm{TD}[\mathrm{I,J,K}]}{\sum\limits_{\mathrm{K}=1}^{4} \mathrm{WL}[\mathrm{I,J,K,N}] + \sum\limits_{\mathrm{K}=1}^{8} \mathrm{WD}[\mathrm{I,J,K,N}]} \tag{12}$$

The weights for the change-in-slope measures depend on the validity of the three data points which produced the two slopes. Because we are calculating the correctness of one of these points, we need examine only the validities of the other two. In both the local and distant neighbor weights, we have used the smaller (MINIMUM) of the correctness indicators of the neighboring points:

$$\begin{aligned}\mathrm{WL}[\mathrm{I,J,K,N}] &= \mathrm{MINIMUM}(\mathrm{RDI}[\mathrm{I} + \mathrm{DI}[\mathrm{K}],\mathrm{J} + \mathrm{DJ}[\mathrm{K}],\mathrm{N} - 1], \mathrm{RDI}[\mathrm{I} - \mathrm{DI}[\mathrm{K}],\mathrm{J} - \mathrm{DJ}[\mathrm{K}],\mathrm{N} - 1]) \\ \mathrm{WD}[\mathrm{I,J,K,N}] &= \mathrm{MINIMUM}(\mathrm{RDI}[\mathrm{I} + \mathrm{DI}[\mathrm{K}],\mathrm{J} + \mathrm{DJ}[\mathrm{K}],\mathrm{N} - 1], \mathrm{RDI}[\mathrm{I} + 2 * \mathrm{DI}[\mathrm{K}],\mathrm{J} + 2 * \mathrm{DJ}[\mathrm{K}],\mathrm{N} - 1])\end{aligned} \tag{13}$$

These weight terms are produced on the assumption that the correctness of a slope difference measure can be no better than the least valid elevation which went into it. As before, the final difference-in-slope correctness indicator RD[I,J] is defined to be RDI[I,J,N] after it has converged.

For the correction process, we need to form one overall correctness indicator at each point from the two indicators we have developed. We achieve this by iterating the RSI and RDI until each has converged, then use the square root of the product of the two indicators as our final correctness indicator:

$$\mathrm{R}[\mathrm{I,J}] = \sqrt{\mathrm{RSI}[\mathrm{I,J,M}] * \mathrm{RDI}[\mathrm{I,J,N}]} \tag{14}$$

where M and N are the iterations at which the RSI and RDI are judged to have converged, respectively.

The change-in-slope threshold DSLTHRESH and the slope threshold SLTHRESH can be simple constants or locally computed quantities. Since one of our terrain data sets included land-use classification data, we have implemented these thresholds through two tables, indexed by the classification associated with each elevation point. For data sets with no accompanying classifications, all grid points are assigned the same classification, and our tables have only one entry each.

Error Correction

Our error correction algorithm is also based on change-in-slope analysis. Under our assumptions, a point is most compatible with its neighbors if it causes minimal changes in the slopes surrounding and across that point. Therefore, to bring a point which has been judged to be in error into conformity with its neighbors, we look for the elevation which will minimize the local changes in slope. Because some of those neighbors may also be in error, this process is weighted by the correctness indicators described in the previous section.

Mathematically, this involves searching for the H[I,J] which minimizes

$$\frac{\sum_{K=1}^{4} \mathrm{WL}[\mathrm{I,J,K}] * \mid \mathrm{DSLOPL}[\mathrm{I,J,K}] \mid + \sum_{K=1}^{8} \mathrm{WD}[\mathrm{I,J,K}] * \mid \mathrm{DSLOPD}[\mathrm{I,J,K}] \mid}{\sum_{K=1}^{4} \mathrm{WL}[\mathrm{I,J,K}] + \sum_{K=1}^{8} \mathrm{WD}[\mathrm{I,J,K}]}$$

where we redefine

$$\begin{aligned} \mathrm{WL}[\mathrm{I,J,K}] &= \mathrm{MINIMUM}(\mathrm{R}[\mathrm{I} + \mathrm{DI}[\mathrm{K}],\mathrm{J} + \mathrm{DJ}[\mathrm{K}],\mathrm{N} - 1], \mathrm{R}[\mathrm{I} - \mathrm{DI}[\mathrm{K}],\mathrm{J} - \mathrm{DJ}[\mathrm{K}],\mathrm{N} - 1]) \\ \mathrm{WD}[\mathrm{I,J,K}] &= \mathrm{MINIMUM}(\mathrm{R}[\mathrm{I} + \mathrm{DI}[\mathrm{K}],\mathrm{J} + \mathrm{DJ}[\mathrm{K}],\mathrm{N} - 1], \mathrm{R}[\mathrm{I} + 2 * \mathrm{DI}[\mathrm{K}],\mathrm{J} + 2 * \mathrm{DJ}[\mathrm{K}],\mathrm{N} - 1]) \end{aligned} \tag{16}$$

with DSLOPL and DSLOPD as defined in Equations 3 and 4, but recalculated for each trial H[I,J] in the search for the solution.

If every point in the elevation matrix is replaced in this manner, a great deal of smoothing of the data will have taken place. This usually results in the loss of significant detail, especially in areas of ridge tops and canyon bottoms. To avoid this, a constraint has been imposed on the use of these corrections, based on the statistical relationship of the adjusted elevation with its neighbors.

This constraint examines each data point in the context of its neighbors by calculating the standard deviation, σ, of the elevations of the eight neighboring points, then comparing the suggested change in elevation with K * σ, where K is a user-supplied constant which controls the amount of smoothing. If the change is larger than K * σ, it is likely that the original data point is in error, so the corrected elevation replaces the original. If the change is smaller than K * σ, then the original data point is likely to represent terrain detail rather than error and are retained. Because the neighbors might well be in error, σ is also calculated using the correctness indicators as weights.

Results and Conclusions

We now present some of the results we have obtained on an elevation data set provided to us by ETL. Figure 1 is a photograph of one of the digital stereo images which was used to produce the digital terrain model. Figure 2 is a contour map of the model, which was generated by computer correlation of the digital images, using a slope-compensating correlator developed at ETL (Crombie, 1976; Panton, 1978).

The elevation data were provided as a set of *X,Y,Z* triples for each data point. These were formed by placing a grid of points in one image, finding their matching points in the second image, then determining the three-dimensional points which correspond to these matches. The *X,Y* portion of the data thus approximates a grid, but is perturbed slightly at each point by the relief. For ease of algorithm development, we have assumed that the data are on a true grid, ignoring the exact *X,Y* data; it could easily be reintroduced, however, since it effects only the distance between data points in the slope computation of Equation 1. The contour maps presented in this section were generated at IAC, using a very simple linear interpolation algorithm on the original and the altered grids of elevation data.

The terrain model represents an area near Phoenix, Arizona. The major feature of the area is the end of a very rough range of hills, occupying the upper left two-thirds of the model. To the right of this is a flatter area crossed by a network of arroyos; below it is an agricultural area, bounded by an irrigation canal and a highway (see Figure 1). The lower right corner contains an orchard also visible in Figure 1.

When we compared the model to the USGS topographic maps for this area, we found the hilly part of the model to be fairly consistent. There are a few places, such as in the ridge at the very top, where the model contours are suddenly jagged, and there are a few extraneous short closed contours. The most glaring error here is a large depression in the middle of the peak which is just above and to the right of center (coordinates [18,28], that is row 18, column 28). The area of arroyos to the right is mostly correct, erring only in a few short or jagged contours.

The agricultural area at the bottom of the model contains many errors, most of which appear as small closed contours, surrounding erroneous data such as the 30 metre depression at [43,39]. Also, the area occupied by an orchard on the aerial photo is covered in the terrain model by a large, steep 60-metre hill which we have nicknamed Error Mountain. This feature is highly improbable, given the very regular orchard, the irrigation canal, and the straight highway which are all in that vicinity. This, clearly is a model in need of correction, as ETL indicated.

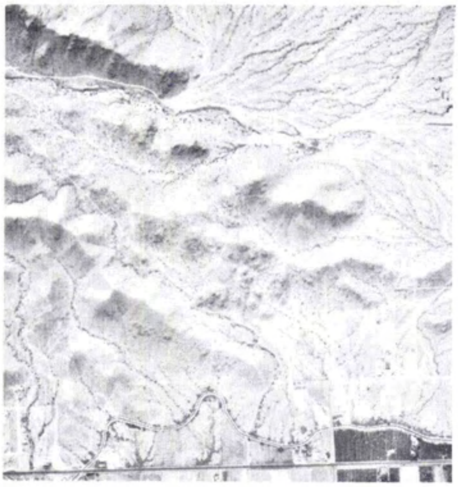

FIG. 1. Aerial photograph of the site of the digital terrain model. The stereo imagery from which the digital terrain model was produced was taken over Phoenix South Mountain Park, just outside Tempe, Arizona. This image, which is 2048 by 2048 picture elements in size, represents a two-inch square section digitized from the original photo by the U.S. Army Engineer Topographic Laboratories.

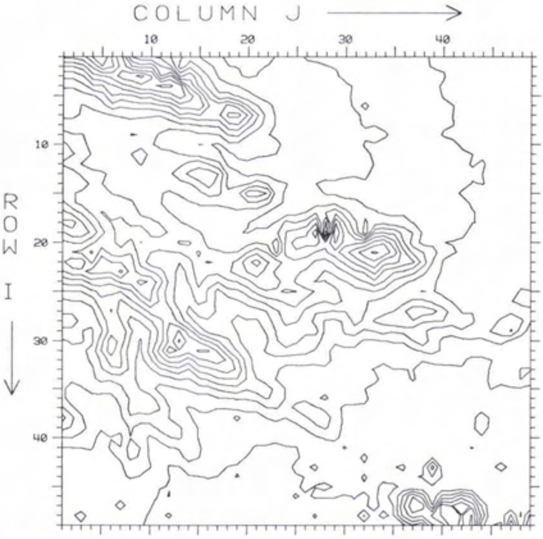

FIG. 2. Contour map representation of the raw elevation data. This is an interpolated contour map of the digital terrain model, as supplied by ETL. Note the large depressions at [18,28] and [43,39], and the systematic errors in the lower right corner. Contours are at 10-metre intervals, and range from 350 metres at [43,39] to 520 metres at [1,5]. Data points are at approximately 45-metre intervals on the ground.

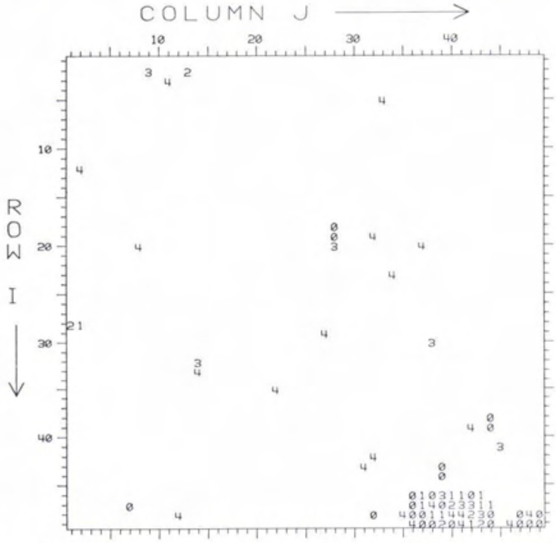

FIG. 3. Combined correctness indicators for the original data. These numbers are a one-digit printout of the correctness indicators multiplied by ten. Indicators greater than 0.5 have been suppressed to highlight erroneous points.

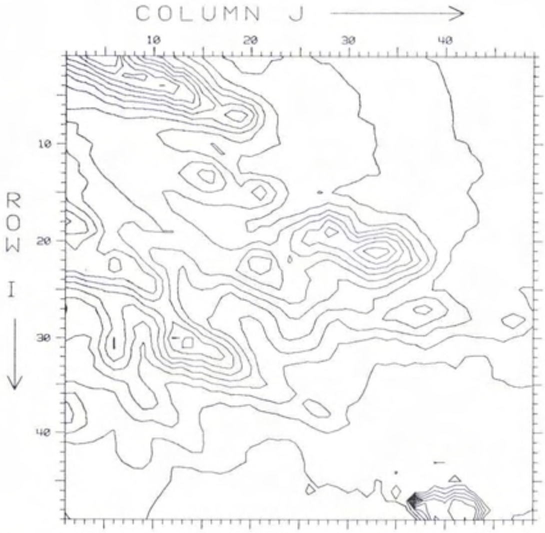

FIG. 4. Contours of the data after one cycle of correction without constraints. After one iteration of correctness analysis and correction, most of the one- and two-point errors have been removed.

The depressions and most of the small errors are one- and two-point mistakes, which ETL believes resulted from low image contrast in these areas. The algorithmic errors which created the large hill are more complicated; it appears that the correlator became confused by the periodic texture of the orchard, then systematically (and progressively) mismatched the trees.

Figure 3 is a printout of the correctness indicators produced by iterating the change-of-slope evaluator three times, and the slope evaluator once. Correctness indicators greater than 0.5 have been suppressed, to highlight the erroneous points. Note that all of the major errors have been indicated and very few

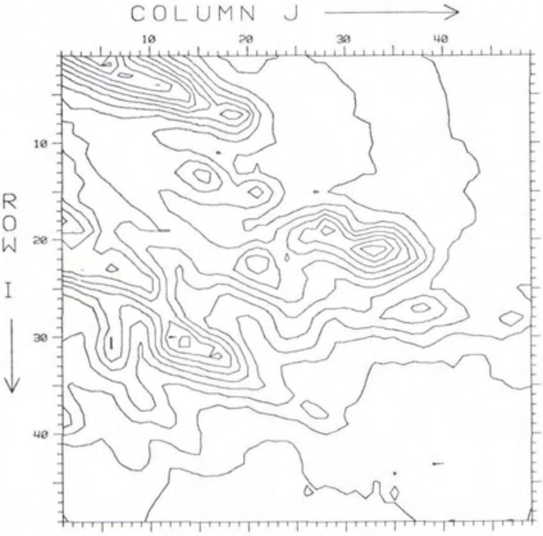

FIG. 5. *Contours of the data after seven cycles of correction* without constraints. After seven iterations of analysis and correction, the systematic errors at the lower right have been removed.

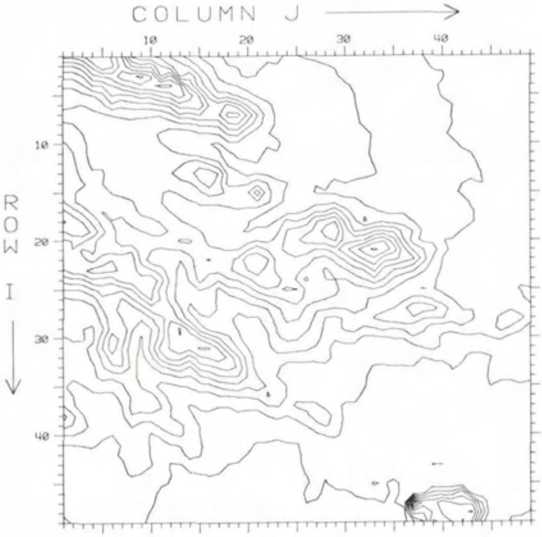

FIG. 6. Contours of the data after one cycle of correction with constraints. Using the local σ in elevation to screen out small changes to the model preserves much of the detail, yet does not greatly effect the error correction process.

plausible points have been marked as being suspect. These indicators, and all of the results which follow, were produced using land-type classification data to select the thresholds for each point.

Figure 4 shows the contours of the model after one cycle of correction in which all points were permitted to change. Note that the depressions and most of the small contours are gone.

Figure 5 shows the model after seven cycles, each consisting of the determination of the correctness indicators, followed by a full correction. Note that the hill created by the errors in the orchard has been completely removed.

When compared to the original model, the contours in Figures 4 and 5 appear very much smoothed. This is an undesirable effect, since it results in the loss of topographic detail in the areas of small stream beds and in the lowering of many of the prominent ridgelines. Figure 6 shows the results after one cycle of correction, using the neighborhood elevation σ criteria for determining when to apply adjustments. The same removal of errors as seen in Figure 4 is present, but the loss of topographic detail is considerably decreased.

These techniques have been applied to sections of other digital terrain models for which no classification information was available, with similar results. Based on these findings, we believe that these techniques have a significant potential in the detection and correction of errors in digital terrain models.

ACKNOWLEDGMENTS

This work was supported by the U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Virginia under ETL task number 18R3205HT08.

REFERENCES

Crombie, Michael A., 1976. *Stereo Analysis of a Specific Digital Model Sampled from Aerial Imagery,* U.S. Army Engineer Topographic Laboratories Report ETL-0072 (U.S. Government Accession Number AD-A033 567).

Hannah, Marsha Jo, 1974. *Computer Matching of Areas in Stereo Images,* Ph.D. Thesis, AIM-239, Computer Science Department, Stanford University.

Jancaitis, J. R., and J. L. Junkins, 1973. Modelling Irregular Surfaces, *Photogrammetric Engineering and Remote Sensing,* Vol. 39, pp. 413-420.

Johnson, Charles N., 1978. *Processing of Terrain Elevation Data for Improved Contouring,* Institute for Advanced Computation Technical Memo 5646, Sunnyvale, California.

Panton, Dale J., 1978. A Flexible Approach to Digital Stereo Mapping, *Photogrammetric Engineering and Remote Sensing,* Vol. 44, No. 12, pp. 1499-1512.

Rosenfeld, Azriel, Robert A. Hummel, and Steven W. Zucker, 1976. Scene Labeling by Relaxation Operations, *IEEE Transactions on Systems, Man, and Cybernetics,* Vol. SMC-6, No. 6.

(Received 12 July 1979; revised and accepted 6 August 1980)