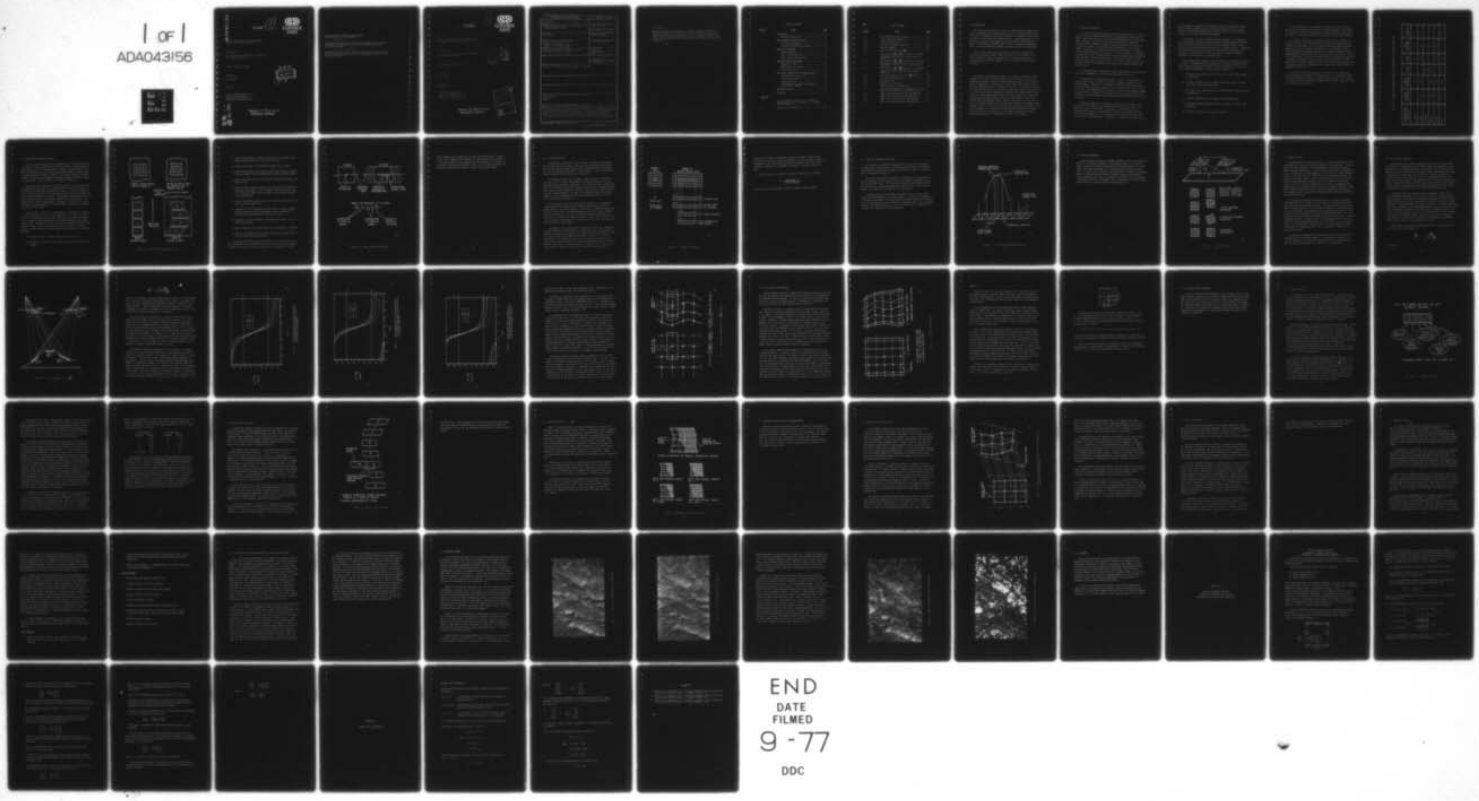
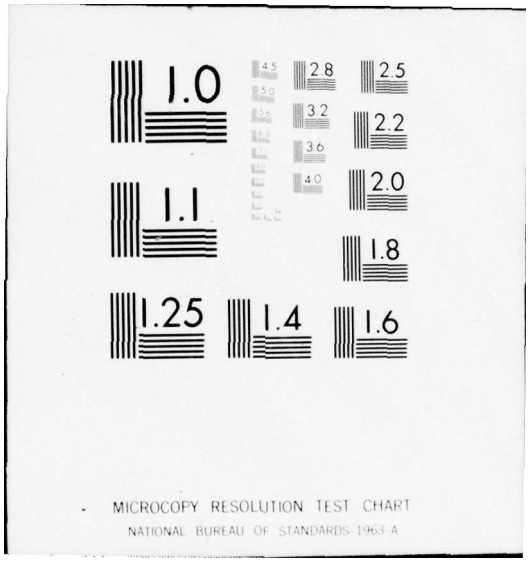
AD-A043 156 CONTROL DATA CORP MINNEAPOLIS MINN DIGITAL IMAGE SYS--ETC F/G 8/2
DIGITAL CARTOGRAPHIC STUDY AND BENCHMARK.(U)
JUL 77 D J PANTON DAAG53-75-C-0195

UNCLASSIFIED ETL-0093 NL

1 OF 1
AD A043156



END
DATE FILMED
9-77
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ETL-0093

CONTROL DATA

Digital Cartographic Study and Benchmark
Fourth Interim Technical Report

Prepared for:

U. S. Army Engineer Topographic Laboratories
Fort Belvoir, Virginia

Contract DAAG53-75-C-0195

Prepared by:

D. J. Panton

July 1977

Digital Image Systems Division
Control Data Corporation
2800 East Old Shakopee Road
Minneapolis, Minnesota 55420

Destroy this report when no longer needed.
Do not return it to the originator.

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

The citation in this report of trade names of commercially available products does not constitute official endorsement or approval of the use of such products.

ETL-0093

CONTROL DATA

Digital Cartographic Study and Benchmark
Fourth Interim Technical Report

Prepared for:

U. S. Army Engineer Topographic Laboratories
Fort Belvoir, Virginia

Contract DAAG53-75-C-0195

Prepared by:

D. J. Panton

July 1977

72 p.

Digital Image Systems Division
Control Data Corporation
2800 East Old Shakopee Road
Minneapolis, Minnesota 55420

Approved For Public Release
Distribution Unlimited

408 732

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ETL-0093	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): DIGITAL CARTOGRAPHIC STUDY AND BENCHMARK, FOURTH INTERIM TECHNICAL REPORT	5. TYPE OF REPORT & PERIOD COVERED: Contract Report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): D.J. Panton	8. CONTRACT OR GRANT NUMBER(s): DAAG53-75-C-0195
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Control Data Corporation, 2800 East Old Shakopee Road, Minneapolis, Minnesota 55420	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Virginia 22060	12. REPORT DATE: July 1977
	13. NUMBER OF PAGES: 70
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Stereo Matching
Algorithm
Pixel

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
This report is the fourth in a series of Interim Technical Reports that cover the development and implementation of a stereo matching algorithm that can be used in automatic terrain data collection. In particular, the results of Phase D of the contract are contained herein. The primary purpose of this Phase was to generalize the algorithm that was developed under Phases A, B, and C to handle more uncontrolled cases of central perspective photography and to

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. continued

lay the groundwork for handling non-central perspective photography. Previous developments and algorithm logic modifications have been reported in a rather piecemeal fashion over the first three Phases. This report combines all these developments into a consistent description of the matching algorithm as it appears to date, including the modifications of Phase D.

TABLE OF CONTENTS

LIST OF FIGURES

1.0 INTRODUCTION

This report is the fourth in a series of Interim Technical Reports that cover the development and implementation of a stereo matching algorithm that can be used in automatic terrain data collection. In particular, the results of Phase D of the contract are contained herein. The primary purpose of this Phase was to generalize the algorithm that was developed under Phases A, B, and C to handle more uncontrolled cases of central perspective photography and to lay the groundwork for handling non-central perspective photography.

Previous developments and algorithm logic modifications have been reported in a rather piecemeal fashion over the first three Phases. This report combines all these developments into a consistent description of the matching algorithm as it appears to date, including the modifications of Phase D.

Throughout the continued development, the matching algorithm has become somewhat of a terrain and sensor analysis tool rather than a strict stereo compilation technique. Under the tuning parameter concept, it is possible to apply the algorithm to a wide range of sensor, image, and terrain conditions. Using the algorithm's built-in reliability analysis, it is possible to assess the difficulty and quality of automatic matching under these varied conditions. It has been found that no single digital technique can approximate the capability of a human stereo compiler when faced with varied sensor records and diverse image and terrain events contained in these records. This is the justification for algorithm tuning. Moreover, the stereo conditions that are optimal for human compilation are not necessarily optimal for automatic matching and vice versa. This is the rationale behind reliability monitoring. The overall objective is to exercise the algorithm under representative experimental conditions so that decisions can be made and parameters acquired regarding what actually is optimum for automatic stereo mapping.

1.1 Underlying Assumptions

The basic philosophy behind the block matching system design is that the ideal situation for matching the images of a stereo pair is to match each pixel of one image individually with its corresponding position on the other image. However, as long as the correlation coefficient is used as the similarity metric between the images, matching one pixel with one pixel is not possible because of the low statistical significance of such a small sample. Therefore, it is necessary to measure the similarity of a group of pixels surrounding the pixel in question. The size and shape of this group, or image patch, must be chosen carefully and may vary from image to image and also from area to area in the same image. The underlying objective is to choose as small a patch as possible such that the local image noise and lack of feature content do not dominate the value of the correlation coefficient for that patch.

The correlation and consequent match point determination of an individual patch is not independent of neighboring patches and match points.

The correlation of image areas introduces a certain averaging effect on the actual point that is matched; the effect increasing as the size of the areas increases. Therefore, it is necessary to shape the image data within a correlation patch so that it conforms to all the match points in the vicinity. In this way, the averaging effect is confined to small linear segments between the match points. In addition, the correlation patch is shaped such that its projection in three-dimensional space approximates the terrain shape as closely as possible.

In matching images, heavy reliance on values of the correlation coefficient alone is a rather unreliable approach in terms of the accuracy of any given match point. Driven solely by correlation maximum searching over extended areas, patches and match points can wander considerably from their true positions depending on the geometric distribution of image noise and high frequency signal components. Therefore, it is necessary to apply a great

deal of geometric constraint to correlation patches and match points based on known geometric parameters and the continuity and slope limits of natural terrain. Stereo frame imagery offers an unique opportunity in this respect because of its inherent, well-defined geometry.

The concept of a coarse correlation search followed by a fine search is less than desirable in most cases because of the aforementioned unreliability of the correlation coefficient alone. The better approach seems to be to perform a fine correlation search followed by a refinement process. In general, well behaved image areas and terrain will match up well on the first fine correlation, and the results will be the same on subsequent refinements. Refinement is necessary only in difficult areas.

When matching stereo imagery, there are essentially six coordinate systems to consider. Depending on the circumstances of the stereo exposures, some of these systems may coincide with others; but for the general case, they must all be handled independently. These coordinate systems are:

- The digital scan coordinate system on the left image (denoted image A), two-dimensional.
- The photo coordinate system on image A defined by fiducial or other calibration marks, three-dimensional.
- The digital scan coordinate system on the right image (denoted image B), two-dimensional.
- The photo coordinate system on image B, three-dimensional.
- The epipolar coordinate system that relates image B to image A, two-dimensional.
- The model coordinate system, three-dimensional.

In stereo matching, it is desirable to define an evenly spaced grid in one of these coordinate systems to drive the matching process. The location of this grid is a factor in determining the number, complexity, and speed of the image processing functions that are collectively called stereo matching.

For example, pixel gray scale correlation must occur in a digital scan coordinate system so an evenly spaced grid in this system facilitates the resampling of image data. It has been found that this resampling and patch shaping are the most time-consuming operations in the matching process. However, prediction mechanisms that make use of epipolar geometry and other photogrammetric transformations deal with photo coordinate systems. Thus, an evenly spaced grid in one of these systems minimizes the prediction inaccuracy. In addition, match points from the stereo pair are intersected to produce terrain data in a model coordinate system. An evenly spaced grid defined here eliminates the need for postprocessing the terrain data. These concepts are summarized in Table 1-1.

In the stereo matching system that is described in this report, the evenly spaced grid is defined in the image A digital scan coordinate system. This has been done to minimize the resampling time and inaccuracy. However, the resulting terrain data must be postprocessed to produce a regularly spaced set of profiles in model space.

TABLE 1-1. PROCESSING REQUIREMENTS BASED ON THE PLACEMENT OF THE EVENLY SPACED GRID

Coordinate System Containing Evenly Spaced Grid	A Image Resampling and Shaping	B Image Resampling and Shaping	A Image Search	B Image Search	Post-Processing Terrain Data	Prediction Iteration
A Digital		Yes		Yes	Yes	
A Photo	Yes	Yes		Yes	Yes	
B Digital	Yes		Yes		Yes	
B Photo	Yes	Yes	Yes		Yes	
Epipolar	Yes	Yes		Yes	Yes	
Model	Yes	Yes	Yes	Yes		Yes

2.0 BLOCK MATCHING CONCEPTUALIZATION

The basic idea behind the current implementation of the matching method is to define an evenly spaced grid of points on image A, and then for each of these points to find its conjugate point on image B. The conjugate point is found by correlating a group of pixels surrounding the point on image A with a sequence of groups of pixels on image B. The groups of pixels are termed blocks or patches, and the sequence of patches on image B defines a correlation search area. These concepts are illustrated in Figure 2-1.

Digital scan lines are generally oriented normal to the flight direction and direction of major parallax. Processing occurs from left to right, or in the direction of increasing x coordinate on image A. A column is defined as a line of patches which lie on the same digital scan line or whose centers all have the same x coordinate on image A. The sequence of processing is to correlate all the patches of a column before moving on to the next column. This scheme allows a rather straightforward management of image A data. That is, the image A buffer window need only be wide enough to contain one correlation patch width of scan lines.

As can be seen in Figure 2-1, the conjugate line on image B of a single scan line on image A can be rather non-linear due to the effects of terrain relief. This conjugate line cuts across many digital scan lines on image B; thus, the image B buffer window must be considerably wider than the image A buffer window. In addition, the size, shape, and orientation of the conjugate patches on image B can be quite different from the nominally rectangular patches on image A, also due to terrain relief and the geometry of the stereo exposures.

The conceptual steps in finding a match point are as follows:

- Determine the next point to be matched along the evenly spaced grid of image A.

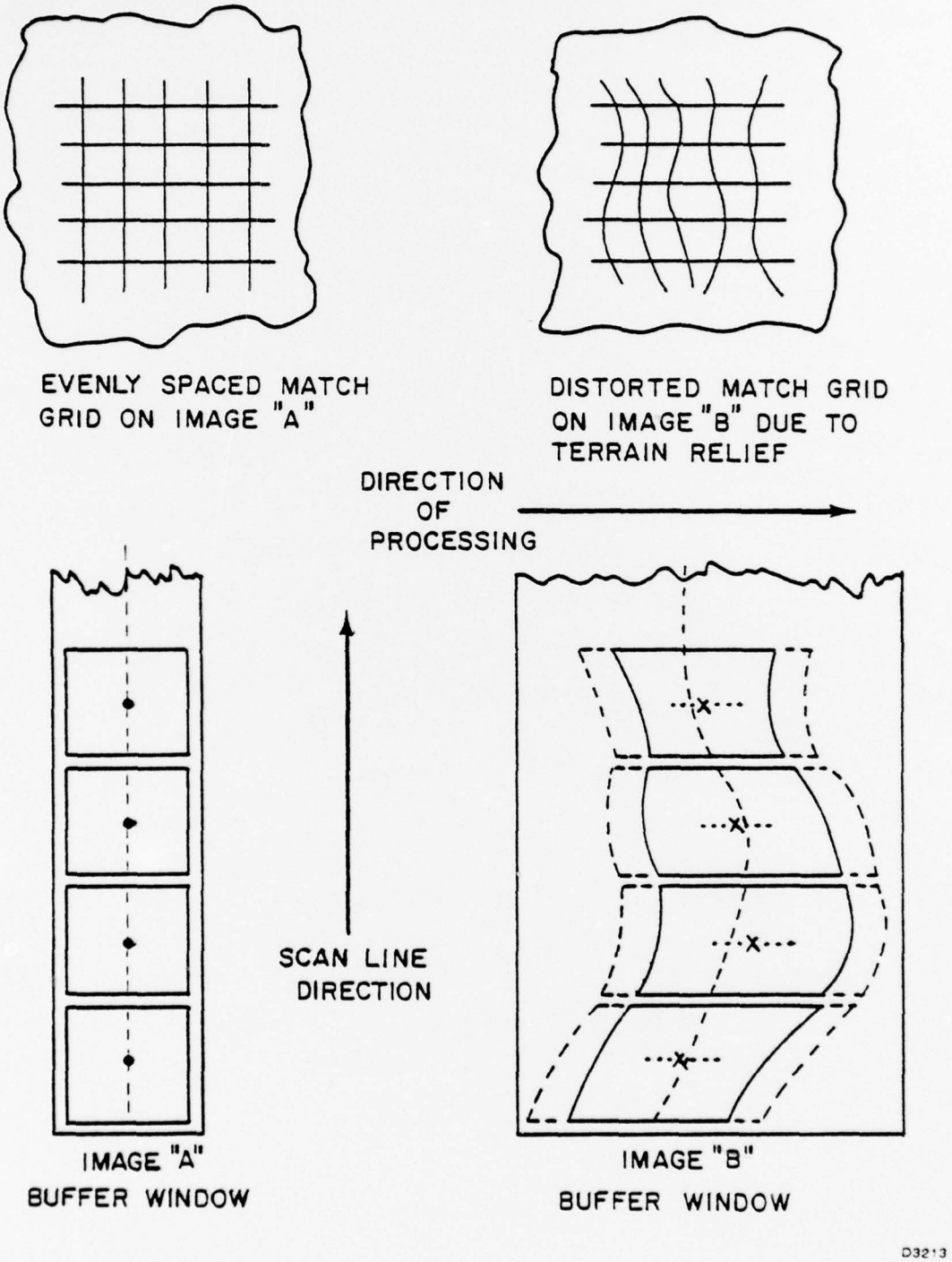


Figure 2-1. Block Matching Conceptualization

- Using epipolar geometry, compute the equation of the epipolar line passing through the point to be matched on image A.

- Compute the equation of the conjugate epipolar line on image B.

- Predict the location of the conjugate match point position on image B along the epipolar line using neighboring, previously matched points.

- Define correlation sites on each side of the predicted location on the epipolar line.

- Shape the image B patch and search area using previous and predicted match point information so that the image B patch most closely lies on the terrain and most closely conforms to the information content of the image A patch.

- Compute a correlation coefficient for the predicted match point location and for each search site.

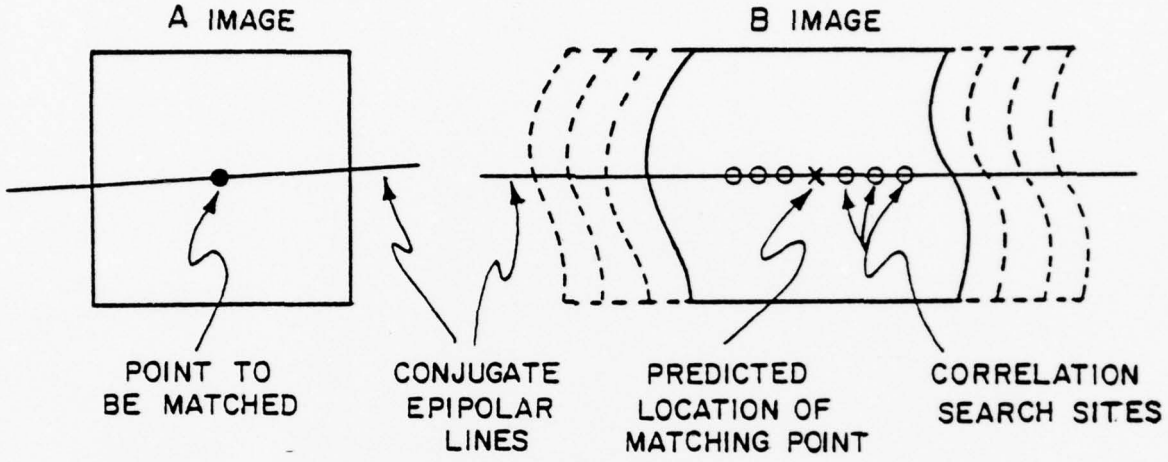
- Determine the site of maximum correlation and fit a smooth quadratic function through it and its two neighbor sites to determine the correlation function maximum to a fraction of a pixel.

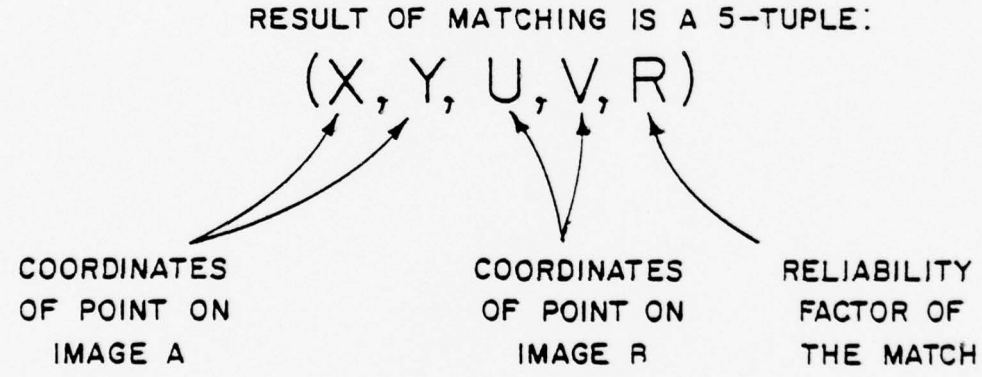
- Compute the reliability factor of the match point based on a set of reliability criteria.

- Apply a correction to the match point if it is excessively unreliable.

- Update the correlation history data and prediction mechanism based on this new match point and reliability.

The primary output of the matching algorithm is a file of five-tuples, one corresponding to each match point found. This five-tuple is illustrated in Figure 2-2 along with the basic terminology of the algorithm. x and y





D3205

Figure 2-2. Stereo Image Block Matching

are the digital scan coordinates of an evenly spaced grid point on image A. u and v are the digital scan coordinates of the conjugate point on image B: v is actually computed in the epipolar line determination, u is found by the correlation search along the epipolar line. R is the reliability factor of the match, and will be discussed in a later section of this report.

2.1 Correlation Strategy

In determining a match point, it is necessary to place the correlation patch at each site of the image B search area, and accumulate over the patch area the necessary gray-scale sums and cross products needed for the computation of the correlation coefficient. The net result is a value of the correlation coefficient for each site of the search area.

The assumption here is that the image B pixel data that is contained in the search area has already been resampled and shaped according to previous matching history and current predictions. The size of the image A correlation patch and the number of correlation sites on either side of the predicted match point along the search segment are variable, being initially defined as part of the input tuning parameters. The pixel data management scheme using buffer windows is rather straightforward and has been described previously in the First Interim Technical Report (1).

The procedure to avoid in the correlation strategy is the independent accumulation of pixel data for each individual patch placement along the search segment. This scheme results in a great deal of redundant computations, and each image B pixel is accessed repeatedly, once for each patch placement that it is contained in. So, under the current correlation strategy, the basic idea is to access each pixel only once for a given search and to accumulate its gray-scale value when it is available in all the sums and cross products for which it has influence.

The mechanics of the strategy are illustrated in Figure 2-3. Pixels are accessed from the buffers row by row. As each pixel of the image A patch is accessed, its value is accumulated in the patch sum and sum of squares. Likewise, as each pixel of the image B search area is accessed, its value is accumulated in its respective column sum and column sum of squares. Also, cross products are generated for each patch placement along the search segment. When the accumulation of sums and products is complete, the variance of the image A patch is computed; and the array of column sums and the array of

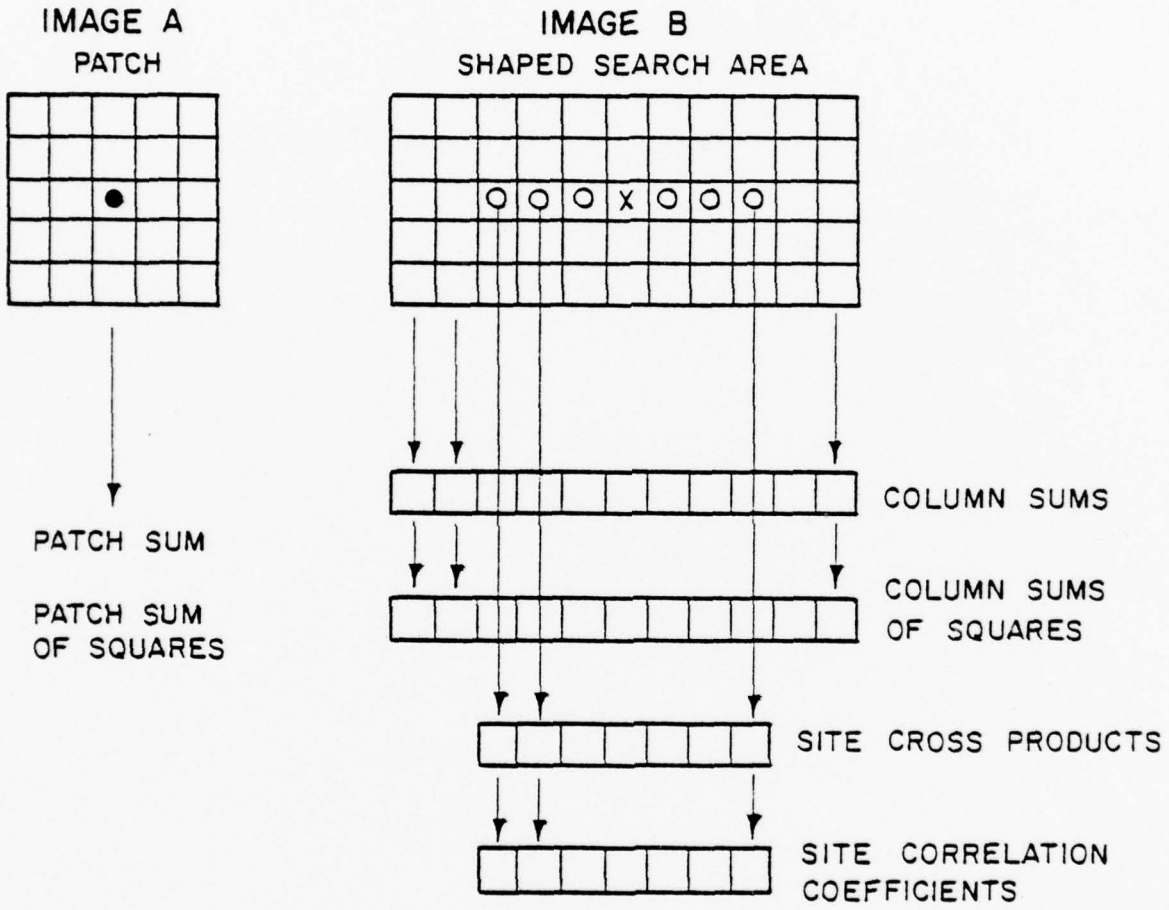


Figure 2-3. Correlation Algorithm

column sums of squares are traversed to generate an image B variance and a covariance for each site of the search segment. The array traversal involves the addition of the next column sum and the subtraction of the last column sum from a running total to simulate the movement of the patch from site to site.

The final result is an array of correlation coefficients of the form,

$$R = \frac{COV\ (A,B)}{\sqrt{VAR\ (A)\ VAR\ (B)}} \quad ,$$

one for each site or placement of the patch along the search segment.

2.2 Correlation Maximum Determination

The correlation coefficients computed by the above strategy correspond to shaped image B pixel centers. To determine a match point, it is necessary to interpolate the correlation maximum to a fraction of a pixel. This procedure is illustrated in Figure 2-4.

The array of correlation coefficients is searched for the maximum value. Then, by using one value on either side of this maximum, a parabolic curve is fit over the three values. Thus, the point along the search segment at which the derivative of the parabolic correlation function is zero is the point of maximum correlation and is defined to be the match point.

A problem exists when the maximum correlation value at a pixel center occurs at either extremity of the search segment. In this case, a parabolic fit cannot be performed; so the pixel center of maximum correlation is defined to be the match point, and this situation is recorded as part of the reliability factor. The underlying idea here is that this match point is suspect and is a candidate for correction or further processing.

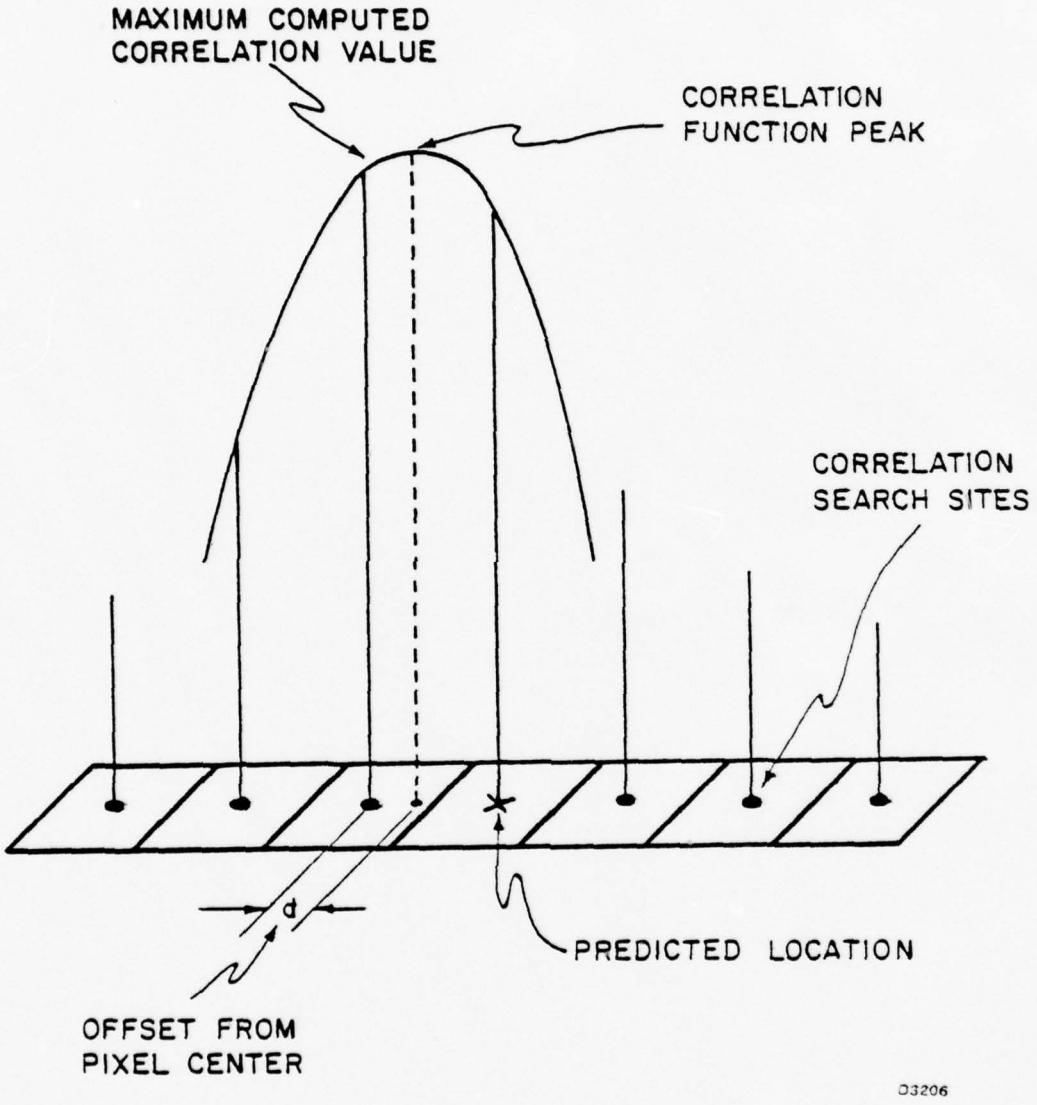
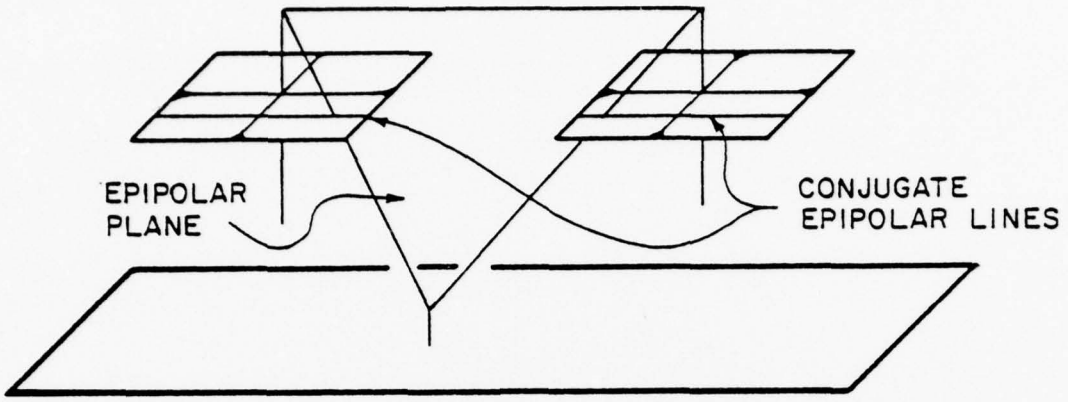


Figure 2-4. Correlation Maximum Determination

3.0 PREDICTION MECHANISM

Epipolar geometry provides the primary constraint on the stereo matching process. This idealized geometry is illustrated in Figure 3-1, showing a number of cases for different orientations of stereo photography. The basic idea is that match points must be on conjugate epipolar lines except when there is excessive film distortion or air refraction. Therefore, this geometry offers a convenient means for making correlation searches one-dimensional along an epipolar line. That is, in forming a match point, the v coordinate is derived strictly from the geometry, thus eliminating the need for a correlation search in a direction normal to the epipolar line.



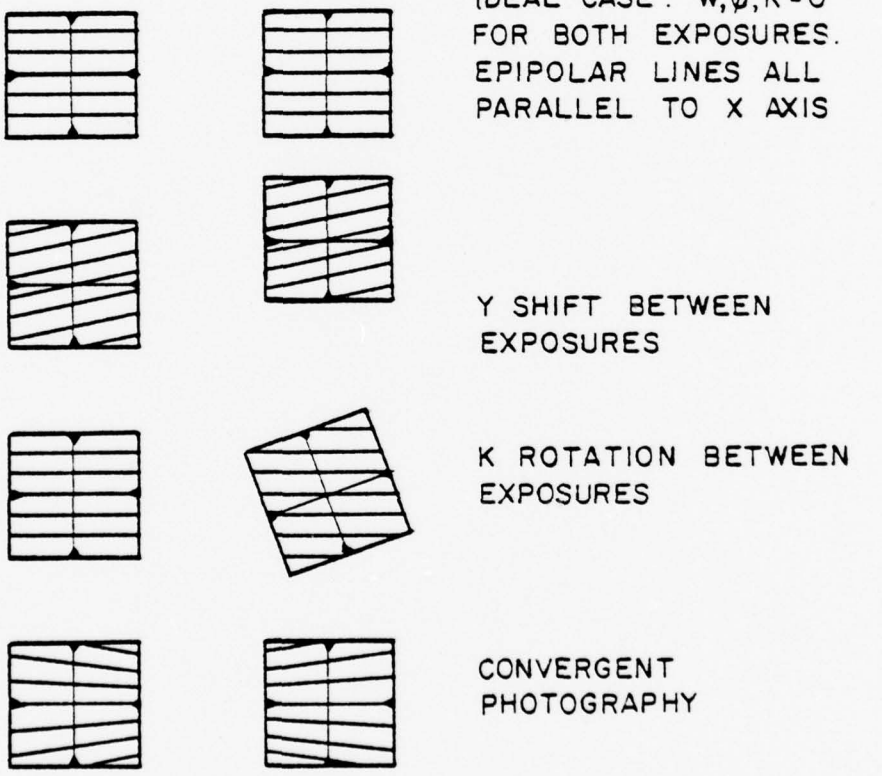


Figure 3-1. Epipolar Geometry

3.1 Epipolar Geometry

The implementation of epipolar geometry in a prediction mechanism requires a knowledge of the relative orientation elements of the stereo pair. These are provided as input parameters. The photogrammetric computations required to produce epipolar predictions on a point by point basis are given in Appendix B. It must be noted that all these computations take place in the photo coordinate systems of the stereo pair, the systems that are functionally related by the relative orientation elements. This is inconsistent with the fact that digital correlation and match point formation occur in the digital scan coordinate systems of the pair. What the algorithm must do, then, to accommodate this inconsistency is to constantly switch back and forth between coordinate systems; making epipolar predictions in photo coordinates, correlating and shaping in scan coordinates, and feeding back match point information in photo coordinates for the next prediction.

This coordinate switching necessitates the use of interior orientation transformations that accurately relate the digital scan coordinate systems to their respective photo coordinate systems. The word accurately here cannot be taken too lightly because it has been shown repeatedly that a great percentage of faulty correlations, and consequently a great percentage of errors in final terrain data, are attributable to inaccurate interior orientation. It is not sufficient to merely digitize the imagery and then process it without a precise mensuration exercise. The point here is that the digitization of an image results in an image transformation that must be modeled. That is, the digital image has its own geometry, incorporating all of the photogrammetric deformations present on the original film as well as distortions created by the digitization process.

A comprehensive procedure for constructing accurate interior orientation transformations is outlined in Appendix A. This procedure, together with relative orientation, is considered a preprocess to the matching algorithm. The transformation coefficients are supplied as input parameters.

3.2 Rate of Change Functions

As explained earlier, epipolar geometry provides a means for predicting the match point v coordinates. But in order to apply this prediction, a prediction for u, the parallax coordinate, is required. This u prediction must primarily take into account the local terrain relief displacement on the images in the area to be correlated. The basic idea is to make this prediction accurate enough so that correlation only has to be performed for one or two pixel sites on either side of the predicted location. The ideal running situation in automatic matching occurs when the tuning parameters are set correctly for the imagery, the image quality and feature content are high, and correlation maxima are found within one pixel of their predicted locations.

A convenient means for characterizing the differential terrain relief displacement between images of a stereo pair is the use of velocity or rate of change functions. This concept is illustrated in Figure 3-2. In the present matching algorithm, the evenly spaced grid to be matched is defined on image A. Therefore, the distance between grid lines of constant x is constant across the image and is termed Δx. The corresponding distances, Δu, on image B are variable and depend on the exposure station positions and orientation and on the slope of the imaged terrain. As can be seen in the figure, the rate of change of image feature placement, $\Delta u/\Delta x$, is one in all cases for flat terrain. Then, $\Delta u/\Delta x$ is less than one or greater than one depending on whether the terrain slopes toward or away from exposure station A.

Analytically, the relationship between the velocity function $\Delta u/\Delta x$ and the actual terrain slope $\Delta h/\Delta X$ is:

$$\frac{\Delta u}{\Delta x} = 1 - \frac{B \frac{\Delta h}{\Delta X}}{H + X, \frac{\Delta h}{\Delta X}}$$

or inversely,

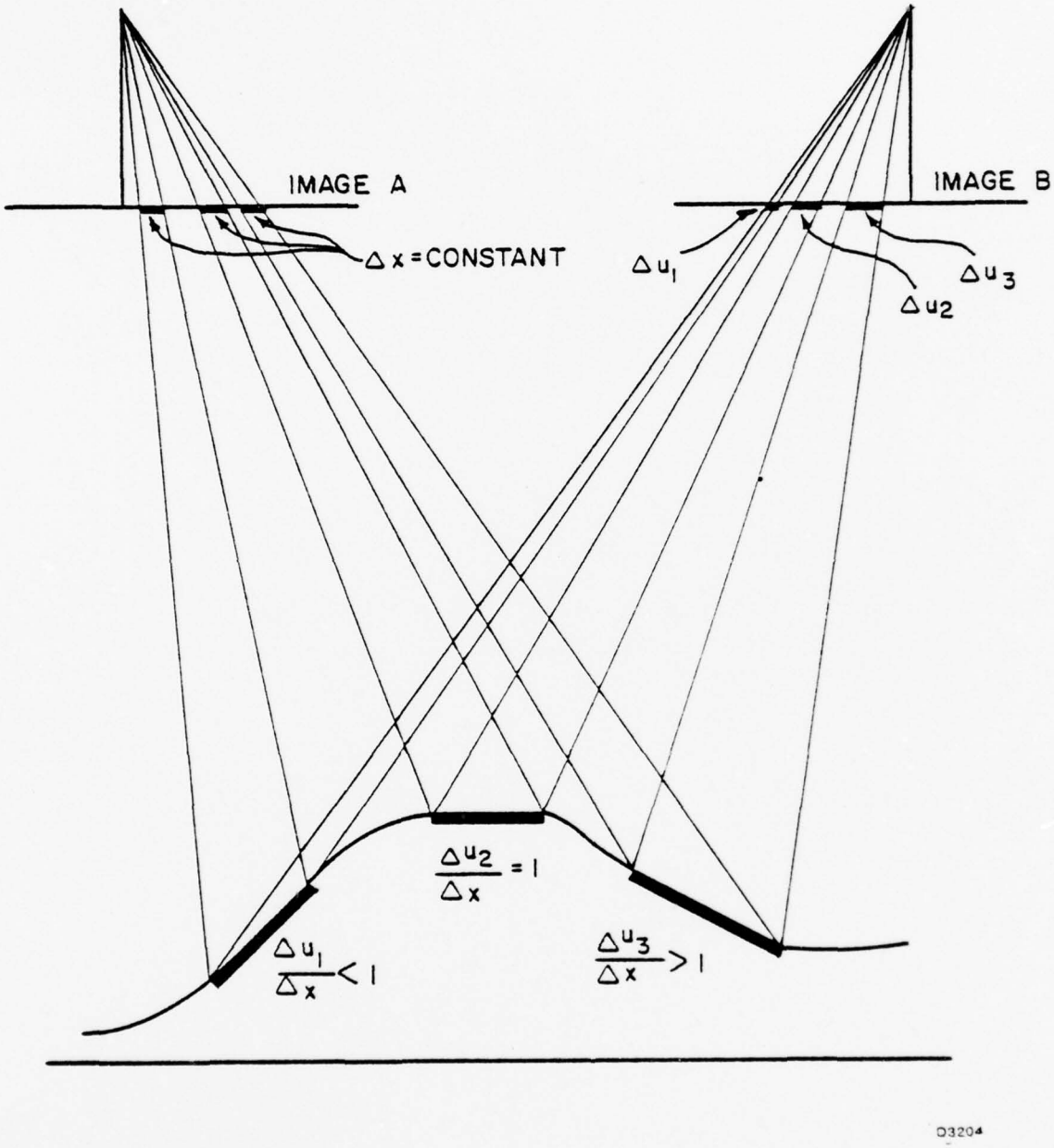


Figure 3-2. Rate of Change Function: $\frac{\Delta u}{\Delta x}$

$$\frac{\Delta h}{\Delta X} = \frac{H\left(1 - \frac{\Delta u}{\Delta x}\right)}{B - X_1\left(1 - \frac{\Delta u}{\Delta x}\right)}$$

where B is the baseline distance between exposure stations, H is the altitude of the exposure stations above some reference datum, and X_1 is the model distance of a point of interest from the nadir of exposure A relative to the baseline distance. These relationships are for the ideal case of vertical photography; the effects of platform tilt are not considered. They have been included in this form for simplicity of illustration.

One can see from the equations that a constant terrain slope does not produce the same $\Delta u/\Delta x$ value in all locations of the model. This is the reason for the X_1 factor. In addition, the relative magnitude of $\Delta u/\Delta x$ varies with the base-height ratio of the exposure stations. Figures 3-3, 3-4, and 3-5 illustrate these variations for base-height ratios of .3, .6, and .9. The three figures are plots of $\Delta u/\Delta x$ against terrain slope for three different positions in the model. For example: If H is taken to be 1.0 and B is 0.3, then X_1 is .075 for Figure 3-3; representing a position one-fourth of the distance from the nadir of exposure station A to the nadir of exposure station B. Likewise, X_1 is .15 for Figure 3-4, and X_1 is .225 for Figure 3-5. The analogy is similar for base-height ratios of .6 and .9.

The important assumption underlying the use of $\Delta u/\Delta x$ as a prediction function is that natural terrain rarely has a slope exceeding $\pm$ 45 degrees. Within this range, as indicated in the figures, the $\Delta u/\Delta x$ function is rather well behaved. Outside of this range, the function becomes asymptotic and discontinuous. For this reason, the current matching algorithm design using $\Delta u/\Delta x$ as a prediction function is not effective when matching large scale images containing angular structures or man-made structures such as buildings, towers, storage tanks, etc. There situations require additional constraints on the $\Delta u/\Delta x$ function or an entirely new prediction scheme. One alternative is to shift the evenly spaced matching grid from the image A coordinate system to the model coordinate system. In this way, image feature velocities can be

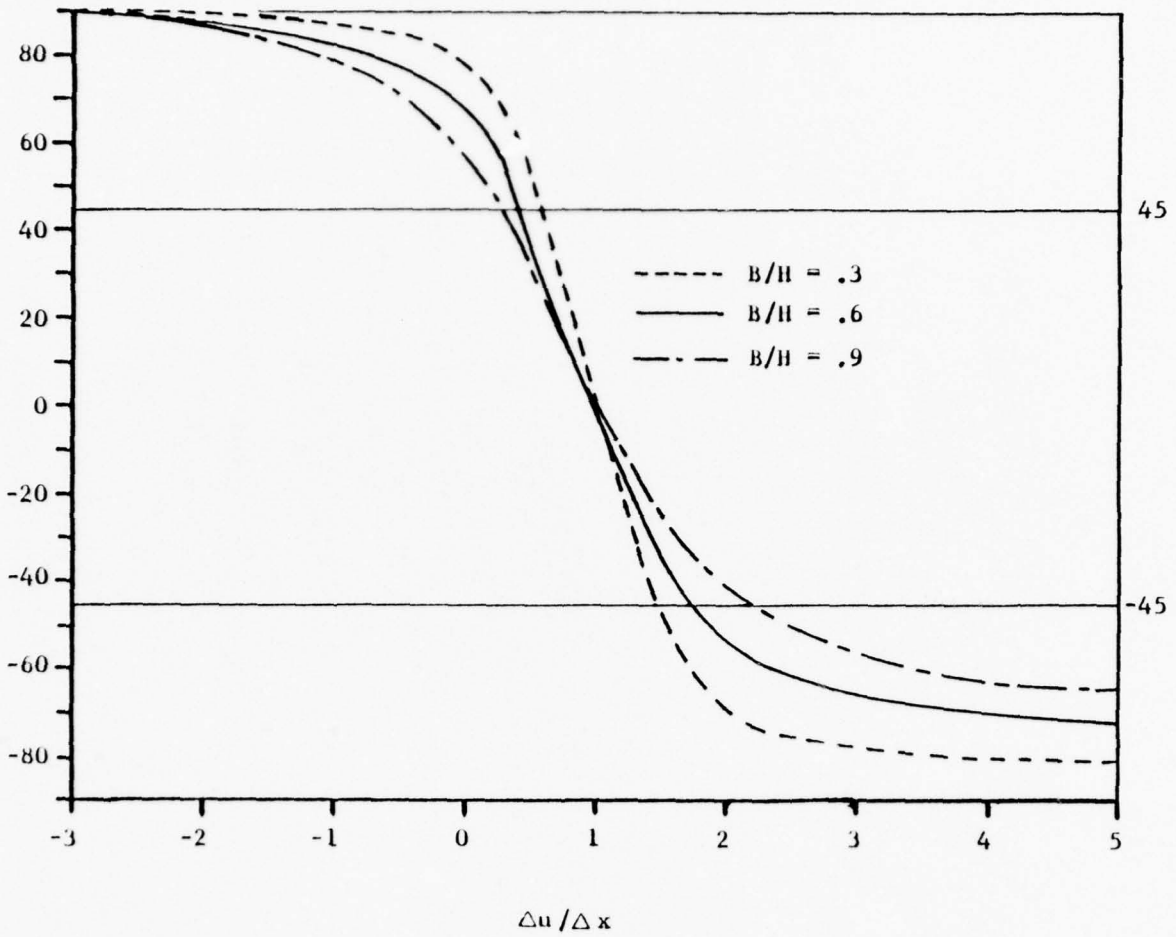


Figure 3-3 Relationship of $\Delta u/\Delta x$ to $\Delta h/\Delta X$ for Points in The Model Between Exposure Station A and The Center of The Model

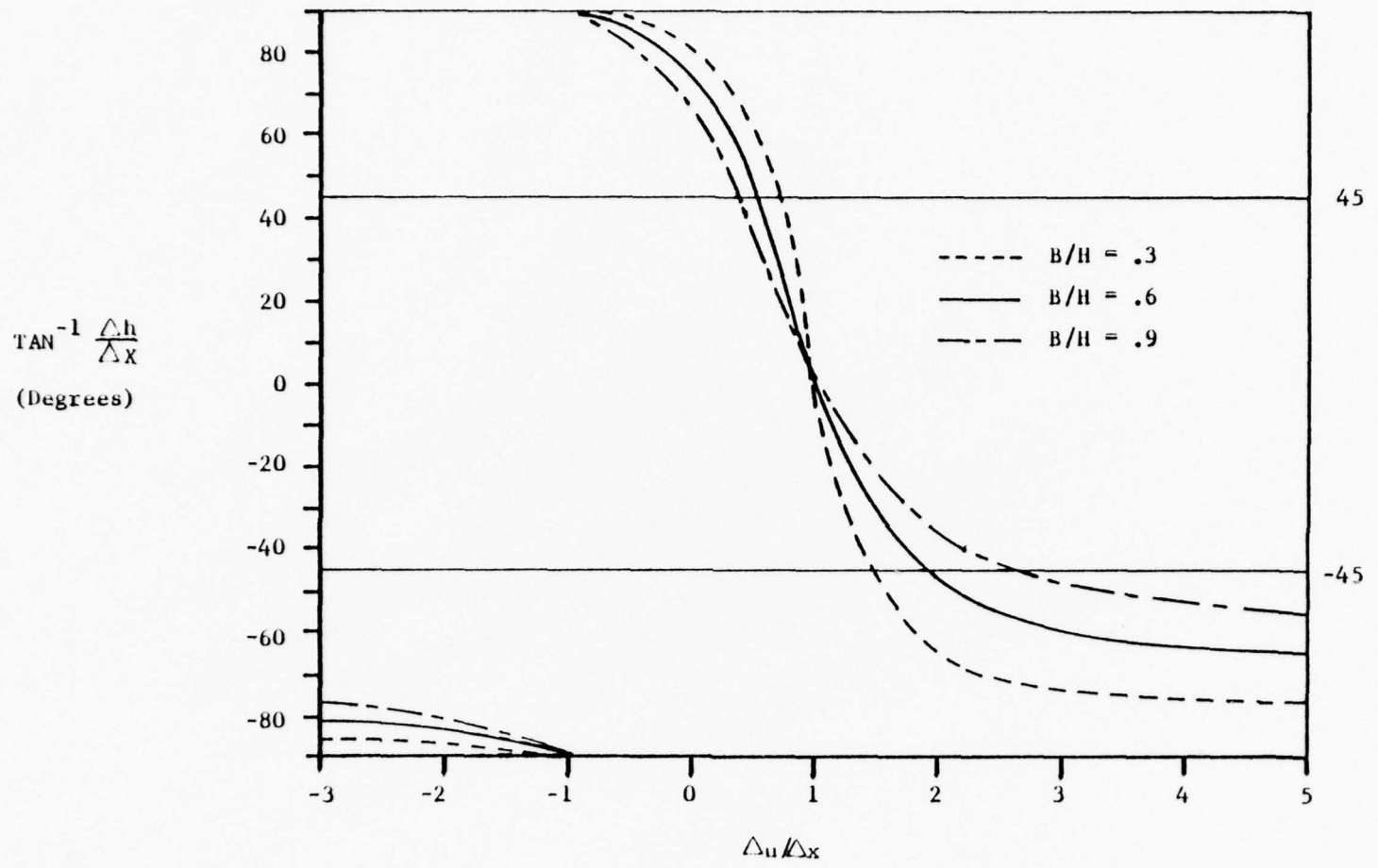


Figure 3-4 Relationship of $\Delta u/\Delta x$ to $\Delta h/\Delta X$ for Points in The Model Midway Between Exposure Station A and Exposure Station B

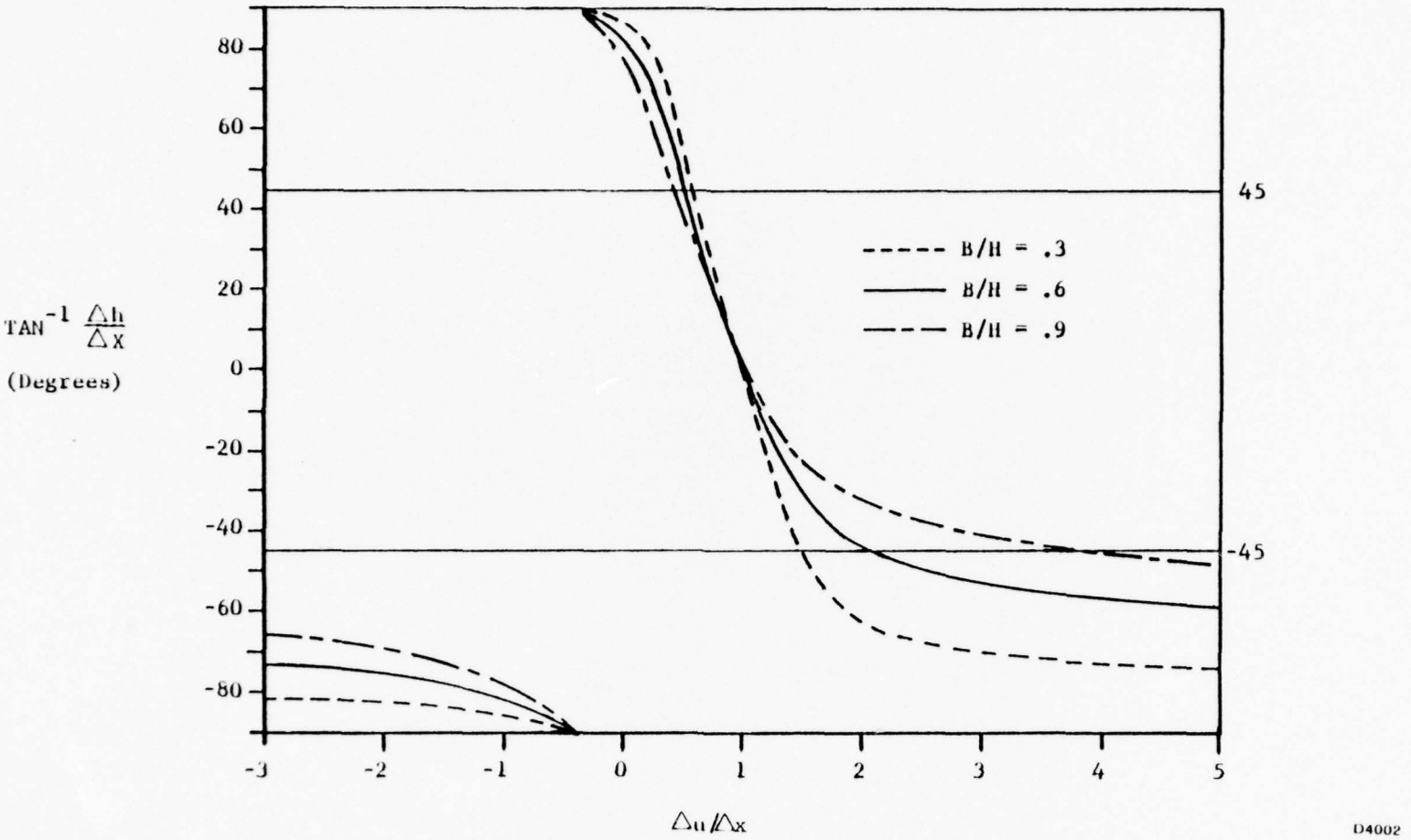


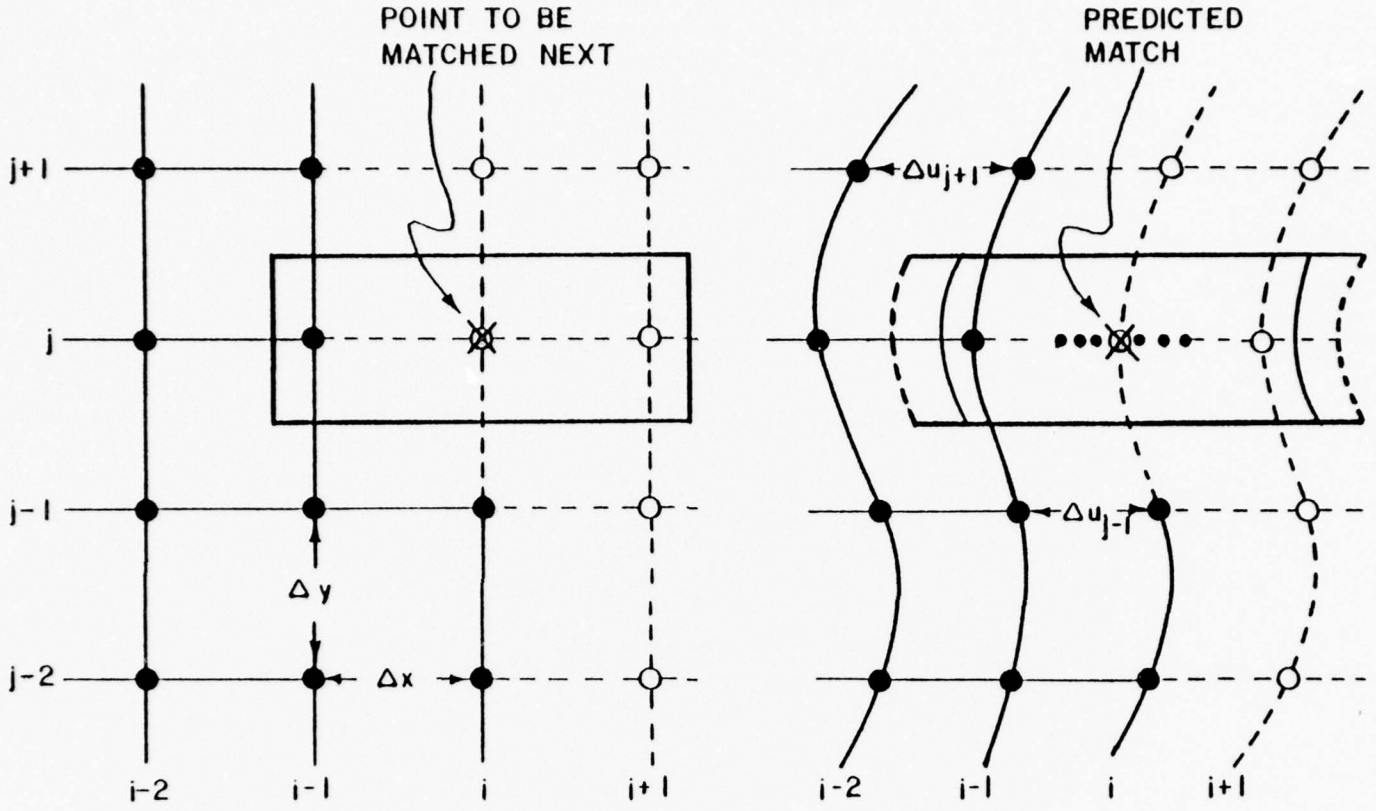
Figure 3-5 Relationship of $\Delta u/\Delta x$ to $\Delta h/\Delta X$ for Points in The Model Between The Center of The Model and Exposure Station B

measured with respect to their model placement; that is $\Delta x/\Delta X$ would be the prediction function for image A, and $\Delta u/\Delta X$ for image B.

An additional consideration that is brought out by the figures is that the range of $\Delta u/\Delta x$ values over the allowable terrain is much greater for a base-height ratio of .9 than for a base-height ratio of .3. Therefore, the matching algorithm cannot respond as rapidly in the .9 case to greatly varying terrain as in the .3 case. Also, the algorithm can be more unstable in tracking the terrain slope along an epipolar line.

The implementation of the $\Delta u/\Delta x$ function as a match point predictor is illustrated in Figure 3-6. In the current matching concept, blocks can be thought of as moving in the parallax direction in paths. These paths are labeled j-2, j-1, j, etc. in the figure. On image A, this movement is in jumps of Δx. On image B, the jump is Δu which may be smaller or larger than Δx depending on whether the terrain is rising or falling in that particular interval. The solid circles in the figure represent points already matched, and the $\Delta u/\Delta x$ values between these points are retained by the algorithm as match point history. The open circles represent points yet to be matched, and must be extrapolated from the local history. The collection of these $\Delta u/\Delta x$ values is a complete characterization of the geometry of the terrain surface; the higher order surface is being approximated by the small linear segments between the match points. The size of the match point grid interval, then, defines the coarseness or fineness of this modeling.

Predicting the next match point involves extending the local terrain surface one grid interval in the process direction; that is on path j in the figure, the extension is made from column i-1 to column i. This is carried out by the equation at the bottom of the figure. Block paths are not independent of one another. Therefore, the necessary cross-coupling between paths is achieved by the path weights W_1, W_2, W_3. These weights are set up as tuning parameters, and their relative values depend to a great extent on the size of the grid intervals, Δx and Δy. A general guideline is that the block path interval closest to the point to be predicted should have the most weight.



$$u_{i,j} = u_{i-1,j} + \left[W_1 \left(\frac{\Delta u}{\Delta x} \right)_j + W_2 \left(\frac{\Delta u}{\Delta x} \right)_{j-1} + W_3 \left(\frac{\Delta u}{\Delta x} \right)_{j+1} \right] \text{WHERE } W_1 + W_2 + W_3 = 1$$

D3207

Figure 3-6. Match Point Prediction by Use of $\frac{\Delta u}{\Delta x}$ Function.

3.3 Patch Shaping and Resampling

Previous matching studies have shown that correlation patches that are the same size and shape on both images of the stereo pair are by no means adequate for performing terrain mapping. The only case in which similar patches apply is the case of flat, non-varying terrain. A quantitative justification for this can be found in the First Interim Technical Report (1).

Referring to Figure 3-7, depending on the selected patch size on image A with respect to the selected matching grid intervals (Δx and Δy) there can be any number of "subpatches" defined by the grid within the total patch area. On image B, these subpatches can all be of different shape in highly varying terrain. Therefore, the patch shaping algorithm has been designed to use the $\Delta u/\Delta x$ values on the sides of the subpatches as shaping factors for each individual subpatch. For example, if the subpatch side (or Δx on image A) is ten pixels, and if the $\Delta u/\Delta x$ value for that interval is .6, then the corresponding subpatch side on image B is six pixels. For pixels lying within a subpatch, the shaping factors are linearly interpolated from the well-defined subpatch side values. The result is that the terrain surface within a patch is being modeled in terms of planar facets such that the patch conforms to the terrain as closely as possible, and the shaped image B patch resembles the image A patch in feature content as closely as possible.

This shaping algorithm design represents a generalization of what was reported previously. In previous designs, the entire patch area was shaped with the same shaping factor. In addition, the patch sides were not permitted to slant; and patch compression and expansion were allowed to occur only in the parallax direction in accordance with the central $\Delta u/\Delta x$ value. The motivation for the generalization came about when small-scale imagery was encountered. Here, the patch had to be large enough to contain enough pixel samples for the correlation coefficient to have significance; yet, this patch was too large in ground distance with respect to the matching grid to allow homogeneous shaping. When the generalized scheme was implemented there was, in fact, an increase in the values and reliability of the correlation coefficients on this small-scale

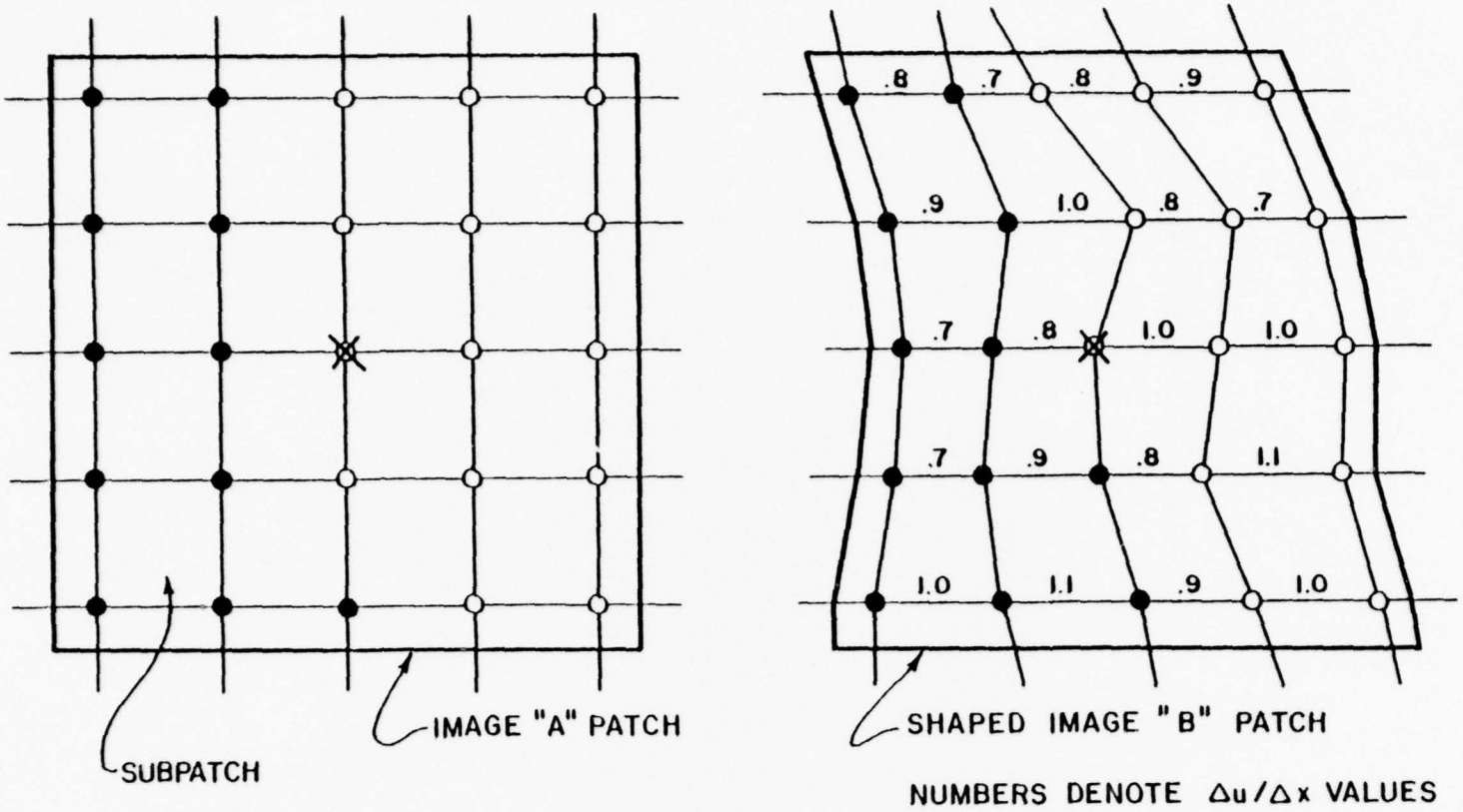


Figure 3-7. Correlation Patch Shaping

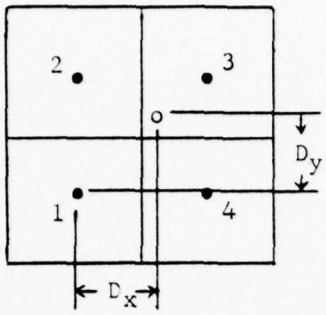
imagery.

Figure 3-7 shows that for any given patch, only a certain percentage of the shaping factors are actually known. These are the result of previous match point determinations. The rest, those lying between the open circles in the figure, must be extrapolated from the known factors as in the prediction mechanism above.

In the implementation of the shaping scheme, digital image resampling plays an integral part. The shaping algorithm must extract from the image B buffer data the correct number of gray scale samples for correlation with the image A patch. These samples do not necessarily lie at pixel centers on the image B digital raster. Using the example above, to correlate a ten pixel wide subpatch that covers only six pixels on image B, ten gray scale values must be resampled from the six.

This resampling and shaping takes place on a row by row basis going up the patch. In the initial implementation of the generalized shaping scheme, the method involved linear interpolation of the gray scale in only one direction, along a row. The epipolar center of the patch was determined and the nearest raster row was designated as the center row. Other rows of the patch fell at integral pixel increments from this center row. Also, it was assumed that the image B patch rows were not rotated with respect to the image A patch rows, and that all patch rows coincided with epipolar lines. Using this method, an averaging effect was observed not only in terms of gray scale (because resampling was nearest neighbor between rows), but also in terms of epipolar line slope with respect to the raster and the image B patch side slant. For large values of this slope and slant, the process can break down.

Therefore, the resampling procedure was also generalized. Resampling and shaping still occur row by row, but now a full bilinear resampling scheme is used so that the row to be resampled can lie anywhere on the image B digital raster. The bilinear formulation is as follows:



Consider the four pixels shown above whose centers are labeled 1, 2, 3, 4. The resampling task is to assign a gray scale value to the point whose distance components from pixel 1 are D_x and D_y, where D_x, $D_y < 1$. If G_i ($i = 1, 4$) are the gray scale values assigned to the pixel centers, then the desired value G is give by:

$$G = G_1 (1 - D_x) (1 - D_y) + G_2 (1 - D_x) D_y + G_3 D_x D_y + G_y D_x (1 - D_y).$$

In moving this four-pixel resampling window to correspond to a patch row, the transition from subpatch to subpatch is made using finite difference techniques.

After the generalized shaping and resampling concepts were implemented, the reliability of the matching over approximately 1,400 match points increased from 80% to 83%. This will be discussed in the next section.

4.0 ALGORITHM CONTROL MECHANISMS

The use of the correlation coefficient as the similarity metric in automatic matching systems gives rise to several pitfalls that must be dealt with. Some of these pitfalls are false correlation maxima, low correlation values with false peaks in featureless areas due to the low signal to noise ratio, and correlation dropout due to gross noise or dissimilar imagery. Therefore, it is necessary for the algorithm to monitor itself and apply match point corrections where appropriate, to overcome the possible instability of the correlation coefficient.

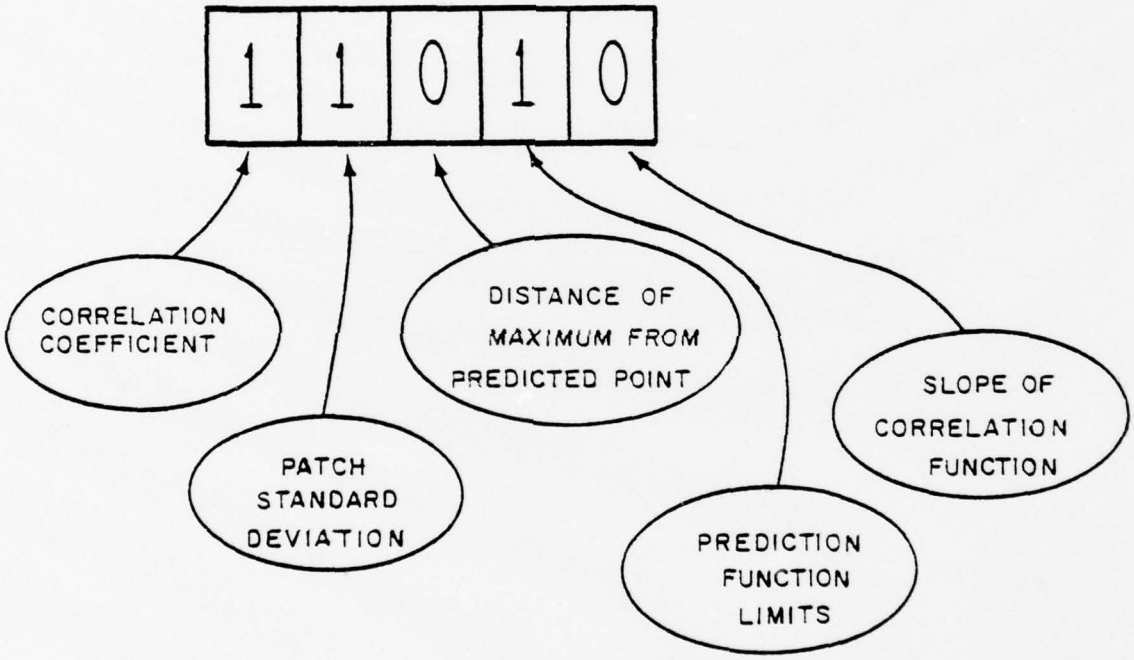
4.1 Reliability Factor

In an effort to allow for match point corrections, a reliability factor has been designed into the algorithm so that both the algorithm itself and the analyst running the algorithm can have an indication of how the matching is progressing. Referring to Figure 4-1, a number of reliability criteria have been set up such that the N digit reliability factor contains one digit for each criterion. In the current algorithm design, there are five criteria; but this number can be increased as more or different processs characterization functions come into use. There is a hierarchy among the criteria such that the most important ones occupy the most significant digits of the reliability factor. A reliability factor is generated for each point matched.

The first criterion is the value of the correlation coefficient. If this value falls below a certain threshold value that was initially input to the process as a tuning parameter, then the digit for this criterion in the reliability factor becomes a one. Otherwise, it is zero. Likewise, the standard deviation of the gray scale samples over the patch are evaluated. In a previous implementation the image A standard deviation was measured for this criterion. Now, the algorithm also generates an image B patch standard deviation and also the difference between these standard deviations. An excessive difference here suggests image dissimilarity which can be caused by feature-sized noise, such as film scratches and dust specks, or by occluded terrain in the case of large base-height ratios.

The third criterion is the flag that was mentioned in Section 2.2 indicating that the correlation maximum was encountered at the extremity of the search segment. The next criterion indicates whether the $\frac{\Delta u}{\Delta x}$ value for the current match point is outside the range for allowable terrain. Again, the allowable threshold is set as a tuning parameter. The last criterion deals with a judgement as to the quality of the correlation peak. Sometimes this is a rather vague indicator because for large, reliable patches the slope of the correlation peak is generally small; yet, for small patches where false peaks are prevalent, the slope can be rather large.

AN N DIGIT NUMBER, ONE DIGIT FOR EACH RELIABILITY CRITERION :



A RELIABLE MATCH POINT HAS A FACTOR OF 0

D3214

Figure 4-1. Reliability Factor

A match point, then, that is reliable with respect to all of these criteria has a reliability factor of zero. However, the reliability factor is more a general indicator of the stability of the matching process than of the absolute match point accuracy. In most cases, instability and inaccuracy are well correlated. There are however, cases where match points that have been deemed unreliable during matching are very accurate positionally, and conversely, reliable points can exhibit some positional inaccuracy.

The reliability factor essentially serves two purposes. Internally, it acts as a decision table for the algorithm to trigger match point correction strategies; and externally, it functions as an analysis tool to aid the user in the selection of algorithm tuning parameters. In the former mode, two strategies have been employed. First, when a match point is unreliable with respect to the value of the correlation coefficient and the patch standard deviation is low, the algorithm holds the $\Delta u/\Delta x$ values for that block path constant until a more feature-rich area is encountered. The underlying idea here is that homogeneously gray areas generally represent planar terrain surfaces with constant slope. Second, when the $\Delta u/\Delta x$ value drifts outside its allowable range, a correction is made to place it within range and local match point corrections are made to stabilize subsequent predictions. This condition typically occurs when false correlation peaks and inaccurate shaping cause the predictor to oscillate. The inclusion of additional strategies in the algorithm must be preceded by an analysis of the specific problem area symptoms. The difficult aspect here is not so much the design of the strategies themselves, but rather the design of the detection algorithms that dictate when a specific strategy is to be applied.

An example of the use of the reliability factor as an external analysis tool is as follows: On a particular set of imagery, it became necessary to assess the applicability of line correlation as compared to area correlation. An image area containing 1,476 match points was chosen, and the matching grid intervals were 2 pixels in the y direction and 5 pixels in the x direction. The search segment length included 3 pixels on either side of a predicted point. For the line correlation case, the patch size was set to be 1 by 35

pixels. This corresponds to a single row of pixels on an epipolar line. For the area correlation case, the patch size was 3 by 35 pixels, giving three times as many samples. The reliability summaries printed at the end of each matching run appeared as follows:

Line Correlation		Area Correlation	
Criterion	Percent	Criterion	Percent
1	3.8	1	3.0
2	1.2	2	.3
3	21.4	3	11.9
4	13.5	4	11.9
5	8.0	5	9.6
Total Reliable	69.8		75.7

The percentages for the five reliability criteria represent the unreliable match points. The bottom line percentages represent total reliability. The conclusion here is that the averaging effect of the areal patch lends more stability to the process. Line correlation is rather jittery. This is borne out by the fact that the percentage of matches where the correlation maximum was found at the extremity of the search segment (criterion 3) halved itself when going from line correlation to area correlation. The increase in reliability of the values of the correlation coefficients and standard deviations (criteria 1 and 2) occurs primarily because of the increased number of samples in the area correlation. The increase in unreliability with respect to the slope of the correlation function can be attributed to the fact that larger patches typically have flatter correlation peaks.

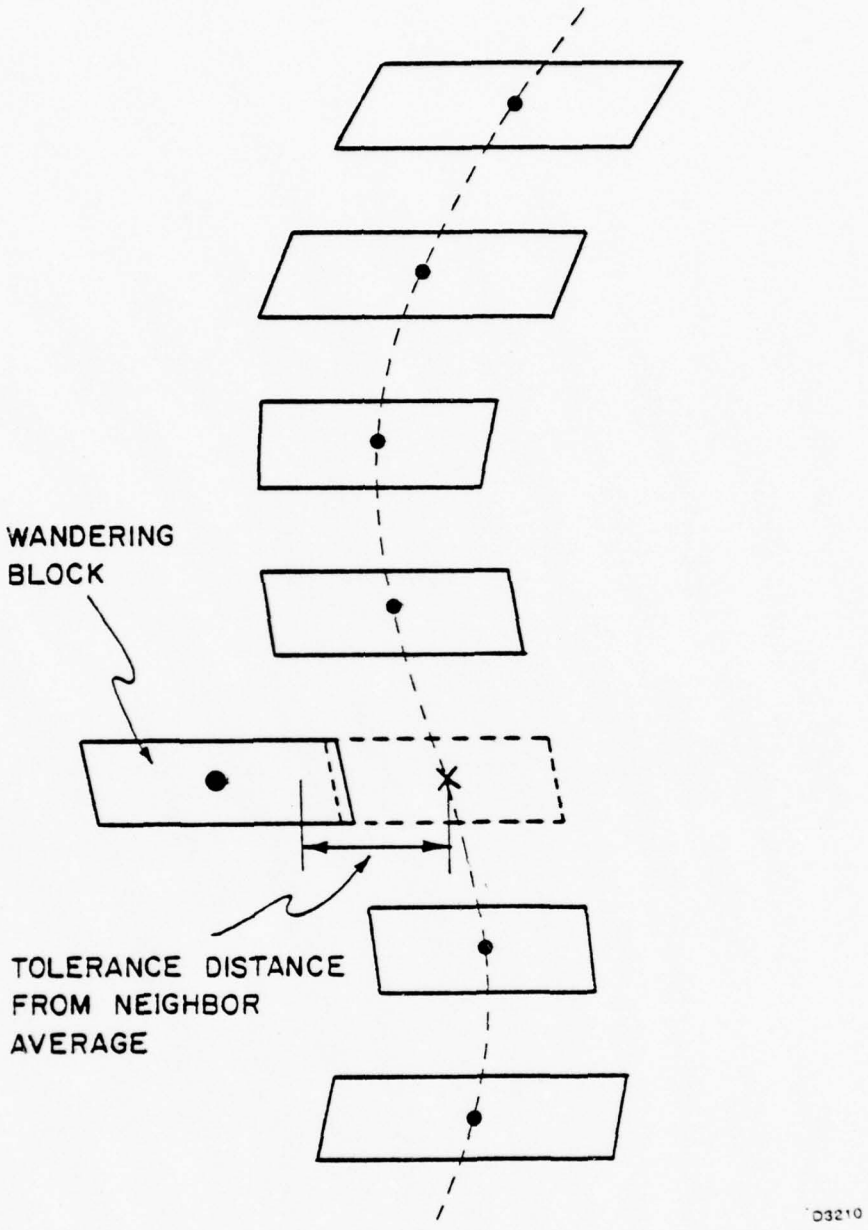
4.2 Wandering Block Tolerance

A control and correction mechanism that plays an important part in keeping the matching algorithm running through difficult image areas is the wandering block strategy. In hard-to-correlate image areas and in areas where false correlation peaks predominate, it is typical for blocks on some paths to lag behind or to jump ahead with respect to neighboring, more reliable block paths. The result is that the $\Delta u/\Delta x$ function produces artificial dips and bulges in the terrain surface at these match points.

The wandering block tolerance is a tuning parameter that has been set up to detect and correct these situations. When each column of blocks has been matched (blocks whose centers all have the same X coordinate on image A), the algorithm stops matching to analyze the entire column for wandering blocks. The u coordinate position of each block on image B is compared to the average u position of its two nearest neighboring blocks. These concepts are illustrated in Figure 4-2. If the block distance from the average position exceeds the preset tolerance, then the match point for that wandering block is moved toward the neighbor average by a weighted amount. This weight is also entered as a tuning parameter. For example, if the weight is 1.0, the match point is corrected to the neighbor average. If it is .5, the point is corrected halfway to the average, etc.

The strategy here is set up such that blocks which have been flagged as unreliable are corrected first. Then the remaining blocks of the column are analyzed. The value of the tolerance distance must typically be consistent with the allowable terrain slope limits set up for the run. When corrections are applied to wandering blocks, the matching history in terms of the stored $\Delta u/\Delta x$ values is also altered to account for the corrections.

It can be observed that the setting of the wandering block tolerance and the correction weight act as a control on the sensitivity of the matching algorithm. A large value of the tolerance makes the block paths more independent of one another and more responsive to small terrain fluctuations. On



- BLOCK IS CORRECTED TOWARD NEIGHBOR AVERAGE BY A WEIGHTED AMOUNT
- WEIGHT DETERMINED BY TUNING

Figure 4-2. Wandering Block Tolerance

the other hand, a small tolerance value increases the cross-coupling between block paths and has a smoothing effect on the resultant match points and terrain data. In this way, the algorithm can be tuned for a variety of situations.

4.3 Patch Center Shift

There is no hard and fast rule in stereo matching that dictates that a match point must be determined for the center of a correlation area. Referring to Figure 4-3 and the figures in Section 3.0 regarding prediction and shaping, there is always a certain percentage of the patch area whose shape is well-known and a percentage whose shape is estimated. The area of known shape is the area in which some match points have already been determined. These known and unknown patch areas can be increased or decreased with respect to one another by shifting the nominal correlation center. For example, shifting the center ahead or to the right of the patch, as pictured in Figure 4-3 increases the areas of known shape. Therefore, the correlation is biased to a greater extent by gray scale samples that are presumably known to correlate well.

Theoretically, this increase in known area should allow the matching algorithm to proceed through highly varying terrain and up steep slopes in a more cautious manner since the patch is well-grounded in familiar area. Conversely, if the unknown area is increased, the algorithm is more liberal in its perception of the terrain and the chances of mismatching increase.

The algorithm currently has the capability for shifting the nominal center in whole grid intervals anywhere within the patch area. However, a complete analysis has not been made to date to assess the effectiveness of the shift for various image events. For all the imagery that has been processed, the correlation center has coincided with the geometric center of the patch.

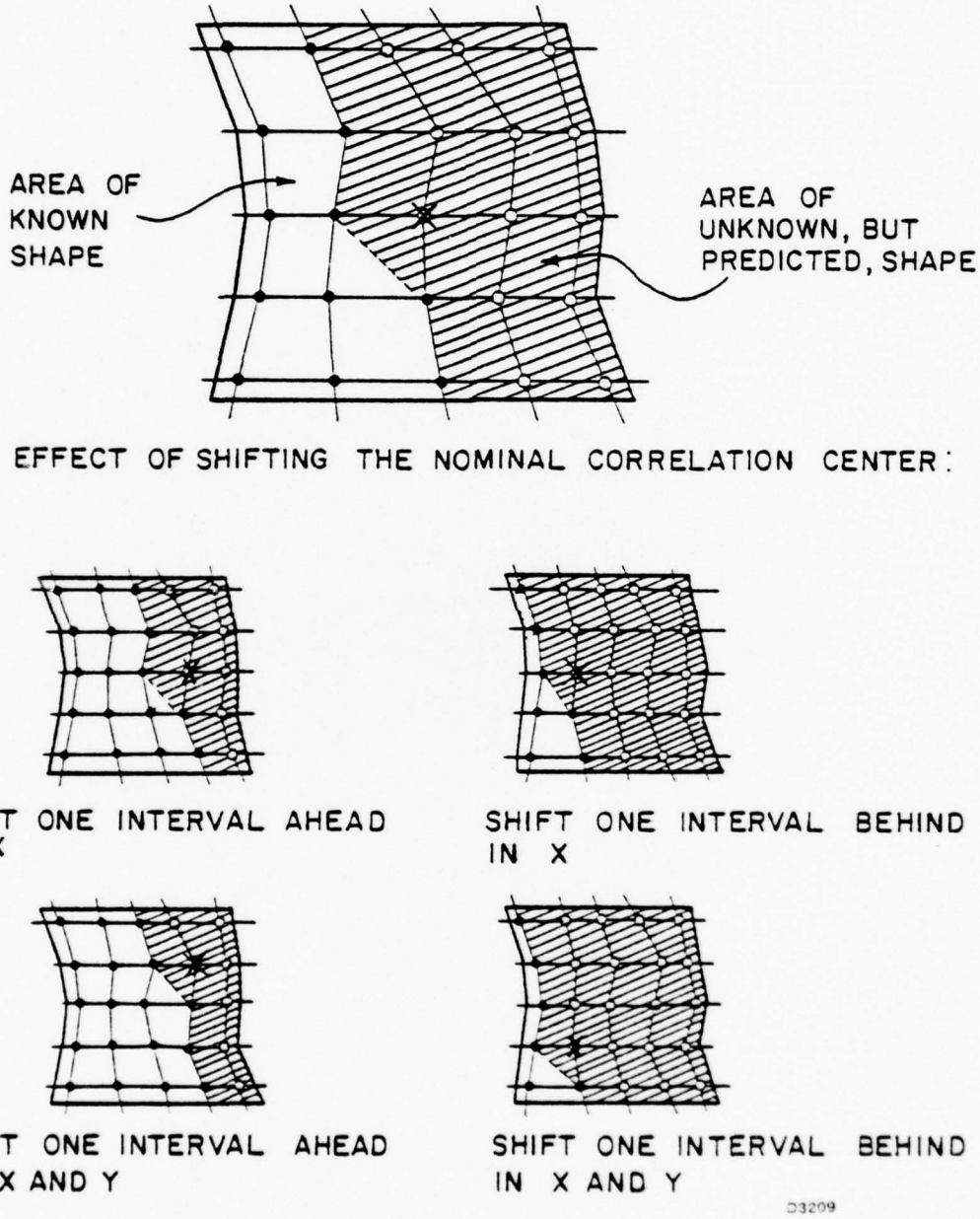


Figure 4-3. Nominal Correlation Center Shift

5.0 ADDITIONAL CAPABILITIES AND GENERALIZATIONS

When the matching algorithm was originally designed, a particular set of test data was used to verify the design concepts. But the complete range of stereo conditions is not exhibited in any one stereo pair. As a result, the algorithm remained untested for certain cases. Also, there are cases that fall completely out of the range of the algorithm's capability. This section of the report describes both the steps that were taken and the steps that need to be taken to provide extended algorithm capability over the range of possible cases.

5.1 Non-parallel Epipolar Lines

In the original design of the stereo matching algorithm, epipolar lines were assumed to be essentially coincident with the rows of pixels on the rasters of both image A and image B. This is usually not the case with real imagery. However, the imagery can always be digitally rectified by a preprocess such that this condition always holds. But the rectification requires image resampling which is rather costly in time, and which can produce a slight amount of image degradation which is magnified by the fact that the image must be resampled again in the algorithm shaping process. So for the current algorithm implementation, the decision was made to resample only once (during shaping) and to generalize the algorithm to take into account both non-parallel epipolar lines and epipolar lines that do not coincide with digital raster lines.

Figure 5-1 illustrates a general situation. The epipolar lines are sloped with respect to the image A digital raster because of the image B platform translation. The digital raster on image B is not parallel to the image A raster or to the epipolar lines because of the platform kappa rotation. As expected by the current algorithm design, the evenly spaced match point grid is parallel to the raster of image A. The important point to consider is that the correlation subpatches that constitute a patch on image B are now no longer trapezoidal as in the previous discussions, but rather are generalized quadrilaterals. The slant of the subpatch sides is determined by the local $\Delta u/\Delta x$ values that are generated as match points slide along the slanted epipolar lines.

The generalized algorithm design handles this situation in the following way. Match point predictions are made along epipolar lines, although the correlation patch is oriented parallel to the digital raster on image A. In this way, the image A patch need not be resampled. This is a justifiable approach as long as all corresponding points within the image A and image B

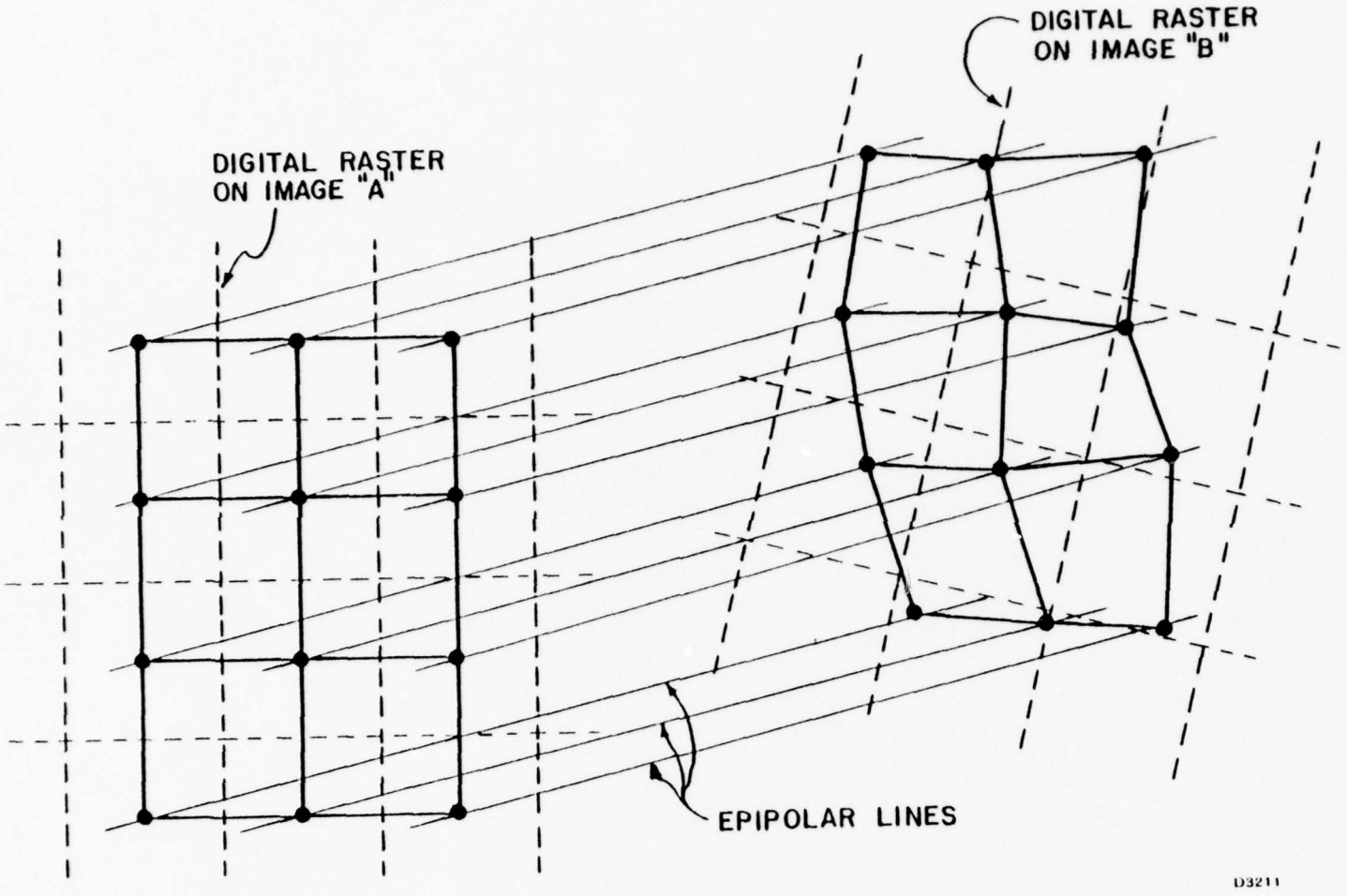


Figure 5-1. Match Point Configuration for Exposures that Are Relatively Translated in Y and Rotated in Z

patches lie on corresponding epipolar lines. The shaping of the image B patch can be performed by the bilinear resampling algorithm for any orientation of the patch with respect to the image B raster. $\Delta u/\Delta x$ values are computed with respect to the image A digital raster, but subpatch slant factors have been introduced to more completely shape the quadrilaterals.

The generalization of the algorithm as described has given rise to additional processing overhead to coordinate the various coordinate systems involved. The amount of overhead required for various geometric situations depends on where the evenly spaced matching grid is defined. For example; when the platform Y translation is large, it may be more advisable to define the matching grid along epipolar lines. This would require the image A patch to be resampled from the image A raster in addition to the image B resampling, but the prediction and shape determination using the $\Delta u/\Delta x$ values would be more straightforward. In addition, the problem of subpatch slant would be minimized.

There are numerous processing tradeoffs like this that exist for various stereo conditions. To account for possible algorithm changes, plugs have been designed into the modular software. In this way, the interplay between coordinate systems may be altered to meet the situation at hand.

It must be mentioned here that as a result of the algorithm generalization, the general-purpose sequential implementation that runs on the CDC 6400 is no longer the same as the parallel benchmark implementation of the algorithm. The capabilities not included in the benchmark are the full generalized shaping by subpatches, bilinear resampling of image B data, small additions to the wandering block strategy, the capability for sloped and non-parallel epipolar lines, and the capability for iterative processing which will be described next.

5.2 Iterative Processing

A new option has been designed into the matching algorithm to allow images to be matched by iterative refinement. That is, subsequent passes of the algorithm over the same image area make use of the match point information generated on a previous pass. The observed net result is that, on subsequent passes, reliable match points tend to stay the same but unreliable areas are improved.

When this iteration option is called for, the prediction mechanism of the algorithm is turned off. The match points from the previous iteration are the actual predictions. What this scheme does, then, is eliminate the patch areas of unknown shape that were described above in Section 3.3. Thus, match points are generated using neighboring match points that are presumably more locally valid in all directions from the desired point.

Using the reliability summary as a judgement factor, it has been observed that the total match point reliability can be increased by as much as ten percent under the iteration option. The largest increase in reliability occurs at the first iteration as is to be expected. After the first, subsequent iterations yield a few more percentage points. However, after the fourth iteration, the reliability stays the same or in some observed cases actually decreases slightly. In these cases, the algorithm tuning was the same for all iterations. A thorough analysis of this phenomenon has not been made, but it is reasonable to assume that after a certain point the process is oscillating on unreliability- merely changing the positions of unreliable match points- and unable to improve without additional information or different tuning.

In attempting to improve the reliability the problem of tuning the algorithms for subsequent iterations to achieve maximum refinement is encountered. Again, this has not been thoroughly investigated. For example, should the correlation patch be smaller or larger on subsequent iterations? A smaller patch should theoretically offer more accuracy in the refinement,

but a larger patch would take into account more surrounding match point data, yet would have a smoothing effect. There are good arguments for either alternative that apply over the whole range of tuning parameters.

5.3 Algorithm Tuning

All previous discussions have hinted that the way the algorithm perceives imagery and terrain cay be changed by changing a set of control parameters that are referred to as tuning parameters. The motivation behind this design is that the algorithm must behave differently for different image and terrain events that can occur. The nature of these different events is determined by the type of sensor used, the stereo taking conditions, and the highly variable characteristics of the terrain that is imaged.

An experiment was conducted on a particular set of imagery to understand the effects of tuning changes. Fifty (50) matching runs were made on the same area of imagery that contained approximately 1,000 match points. The tuning parameters were varied one by one from one run to the next in an effort to constantly increase the matching reliability. The iteration option was not employed for this experiment. The result was that the matching reliability for all the tuning cases ranged from a minimum of 49% to a maximum of 82%.

Each tuning case changes the behavior of the algorithm to certain types of terrain. It was observed that small changes to the values of the tuning parameters sometimes made very large changes in algorithm behavior. The best matching seems to occur when the process is tuned such that it is on the brink of instability. In this way, the algorithm is most responsive to terrain variation, yet stable enough to track the image accurately from one end to the other.

A major factor in determining how the algorithm responds is how the algorithm starts out cold on an image. Experimental results show that the effects of inaccurate matching at the start can be observed as much as 100 scan lines down the image. There are essentially two alternatives to getting the algorithm started. The first, and most accurate, is to supply as input the starting positions and initial terrain slopes of each block path.

The second is to supply one or more manually measured points from which the algorithm can extrapolate the starting match point positions. The latter is a desirable approach from a user standpoint because the measurement of many closely spaced points and the estimation of terrain slope is a rather tedious and time-consuming operation. However, this automatic approach can be error prone depending on the number of points actually measured and the local terrain variation present in the vicinity of the starting block positions.

The automatic acquisition technique performs a matching iteration over extended search lengths. From the few measured points that are supplied, the block centers for the first column of match points are linearly extrapolated. A correlation is performed at each of these block centers using rather long search lengths to account for maximum parallax differences. There is no patch shaping involved here because the local $\Delta u / \Delta x$ values are unknown. It is as if flat terrain were being processed. Then predictions are made for the next column of match points, and they are correlated similarly. At this point, the $\Delta u / \Delta x$ values are computed between these first two columns and the process backs up to start over -this time shaping as it goes. This second pass acts as a refinement so that at its completion, all the starting positions and shapes are available and control is transferred to the full matching algorithm which starts once again at the first column. In the acquisition, as each match point is correlated repeatedly, the search segment size is decreased progressively.

For summary purposes, the following list of matching algorithm input values is given. The distinction is made between parameters that are constant for a particular stereo pair and ones that are part of the tuning parameter set.

Input Parameters

- Relative orientation elements - these specify the baseline distance and relative tilt angles that relate exposure station B to exposure station A.

- Interior orientation transformations and camera focal length - these specify the conversion between digital scan coordinates and photo coordinates.

- Image buffer parameters - a characterization of available buffer sizes and starting line addresses.

Tuning Parameters

- Match point grid limits and interval sizes.

- Correlation patch size in two dimensions.

- Number of correlation sites along search segment.

- Amount of nominal patch center shift.

- Block path prediction weights.

- Wandering block tolerance and distance correction factor.

- Reliability thresholds for the correlation coefficient, standard deviation, $\Delta u/\Delta x$ range, and slope of the correlation function.

- Prediction option selector.

- Manually measured start-up points.

5.4 The Matching of Non-central Perspective and Dissimilar Imagery

All of the algorithm development and implementation described thus far has taken place for central perspective photography. The epipolar prediction geometry is well-defined and straightforward with respect to this kind of imagery. If the matching algorithm is to be applied to non-central perspective imagery, then the module of the algorithm that requires redesign is the prediction mechanism. All other modules can essentially remain the same. In order to achieve accurate predictions, the dynamics of the sensor of interest must be modeled in such a way that conjugate points on the images lie on pseudo epipolar lines. The motive here is to keep the correlation search one-dimensional. The modeling problem involves finding the direction on the imagery in which terrain relief displacement occurs. Unlike epipolar lines, this direction may be different for the two images of the stereo pair and also for different regions in the same image. But if the terrain relief is analytically predictable in small localized areas, there is a good chance for matching success. To date, processing of non-central perspective imagery has not taken place with the current matching algorithm.

Between two images that have the same ground coverage, image dissimilarity can occur in two domains: the geometric domain, and the radiometric or intensity domain. In the geometric domain, image dissimilarity between images to be matched is attributable either to an overall scale difference or to differing sensor dynamics. Scale dissimilarity can be handled by the current algorithm. As long as the relative orientation elements and focal lengths reflect the scale change, correct scale matching is a straightforward by-product of the resampling and shaping schemes. The larger scale image may have to be filtered in some cases to reduce the high frequency noise and to make its intensity more similar to that of the smaller scale image, but this is the only additional processing that would be required. When image dissimilarity occurs in an overlapping pair because different sensors are used, the remarks of the previous paragraph apply. The matching algorithm must model the sensors such that accurate geometric predictions can be made.

In the intensity domain, two alternatives exist for solving the problem of image dissimilarity. The first alternative follows the assumption that all the matchable features are contained in both images, but under differing gray-scale renditions. The approach to take here is a form of intensity shaping which is analogous to the geometric patch shaping contained in the present algorithm. The gray-scale differences are modeled according to a gray-scale mapping function such that when resampling occurs for image B, for instance, the pixel gray values are transformed by the mapping function into the equivalent image A values. Thus, the correlation algorithm can operate on seemingly similar imagery. But when this intensity modeling is not possible or when features are not common to both images, as in the case of radar vs photography, then a second alternative must be employed. Both images must be transformed either by preprocessing or through resampling to a third, common intensity spectrum which enhances the similarities and suppresses the dissimilarities. One such image transformation is the feature edge or gradient operation. Again, the general principle is that automatic image matching can only occur if the two dissimilar images can be made to appear similar to the matching metric through some sort of modeling or intensity transformation.

6.0 PROCESSING EXAMPLE

To illustrate the behavior of the stereo matching algorithm described in this report, a stereo pair taken over the Phoenix - South Mountain area in Arizona was processed. The scale of the photography is 1:48000, the focal length of the mapping camera being nominally six inches. Two by two inch sections from the original photos were digitized at ETL using a 35 micrometer scanning spot size over a 24 micrometer interval. This resulted in two digital images that are 2,048 X 2,048 pixels. A pixel side corresponds roughly to four feet on the ground, but a mismatch or parallax error of one pixel results in about 7.7 feet of vertical error on the ground. The mensuration and interior orientation of the photos were also performed at ETL.

Figure 6-1 is image A of the stereo pair. The size of this digital image is approximately 1,260 pixels by 2,032 scan lines. It is considerably enlarged for illustrative purposes and has been digitally enhanced since the gray-scale range of the original digital data was rather narrow. The black lines on the picture are the evenly spaced matching grid lines in the parallax direction. These lines are spaced eight scan lines apart. Along these lines matching occurred at every tenth pixel, so for this image section, 53,910 match points were generated. A majority of these points were matched using a 21 X 21 pixel correlation patch over a seven-site search segment.

Figure 6-2 shows the conjugate grid superimposed on image B. This grid plot was produced by plotting straight lines between all the match points lying in the same column. The plot, then, is actually a picture of the match points and also of the parallax function that relates image B to image A. If Figure 6-1 and Figure 6-2 were viewed under a stereoscope, the corresponding lines would fuse stereoscopically and the grid would appear to lie on the terrain in three-dimensional space.

The match points were photogrammetrically intersected using the absolute orientation elements to produce a file of digital terrain data. The elevations were then contoured at 20-foot intervals. The resulting digital contour

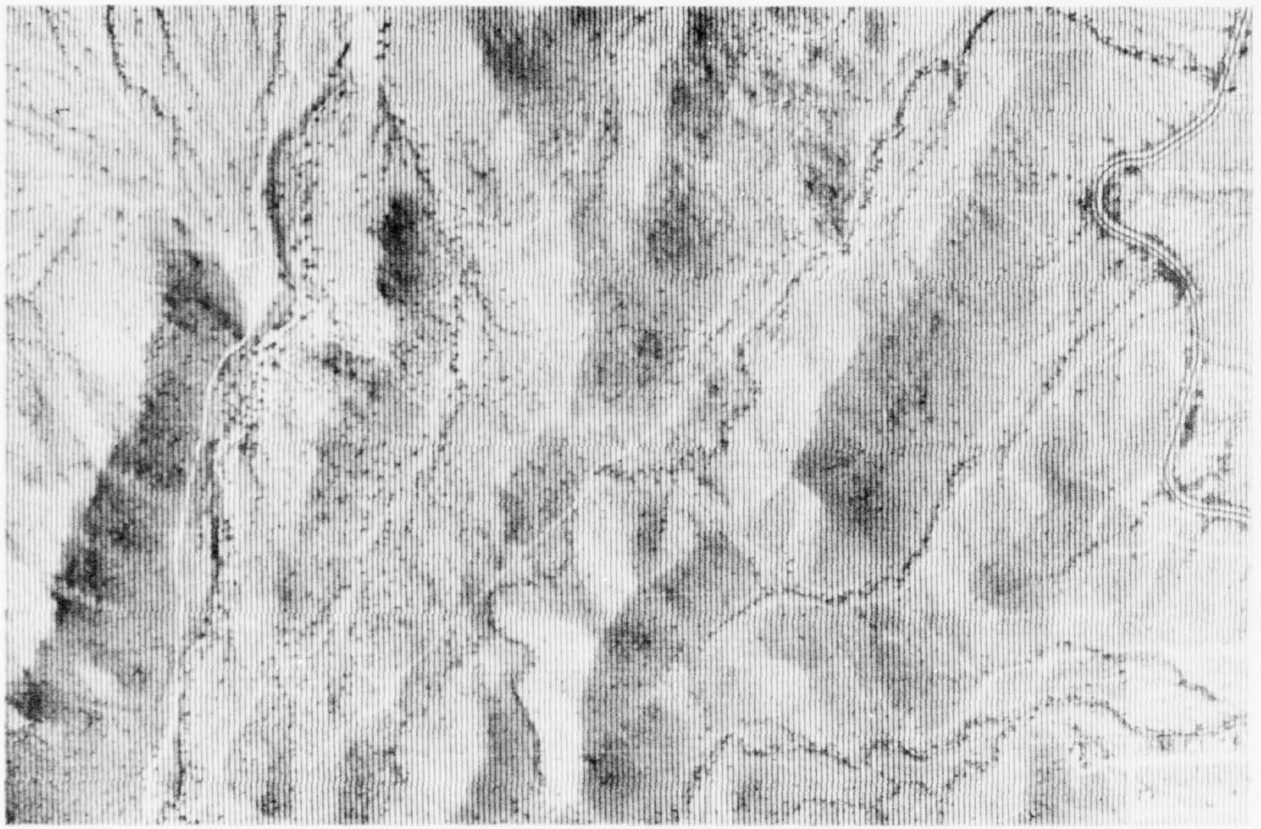


Figure 6-1 Image A with Evenly Spaced Grid Superimposed

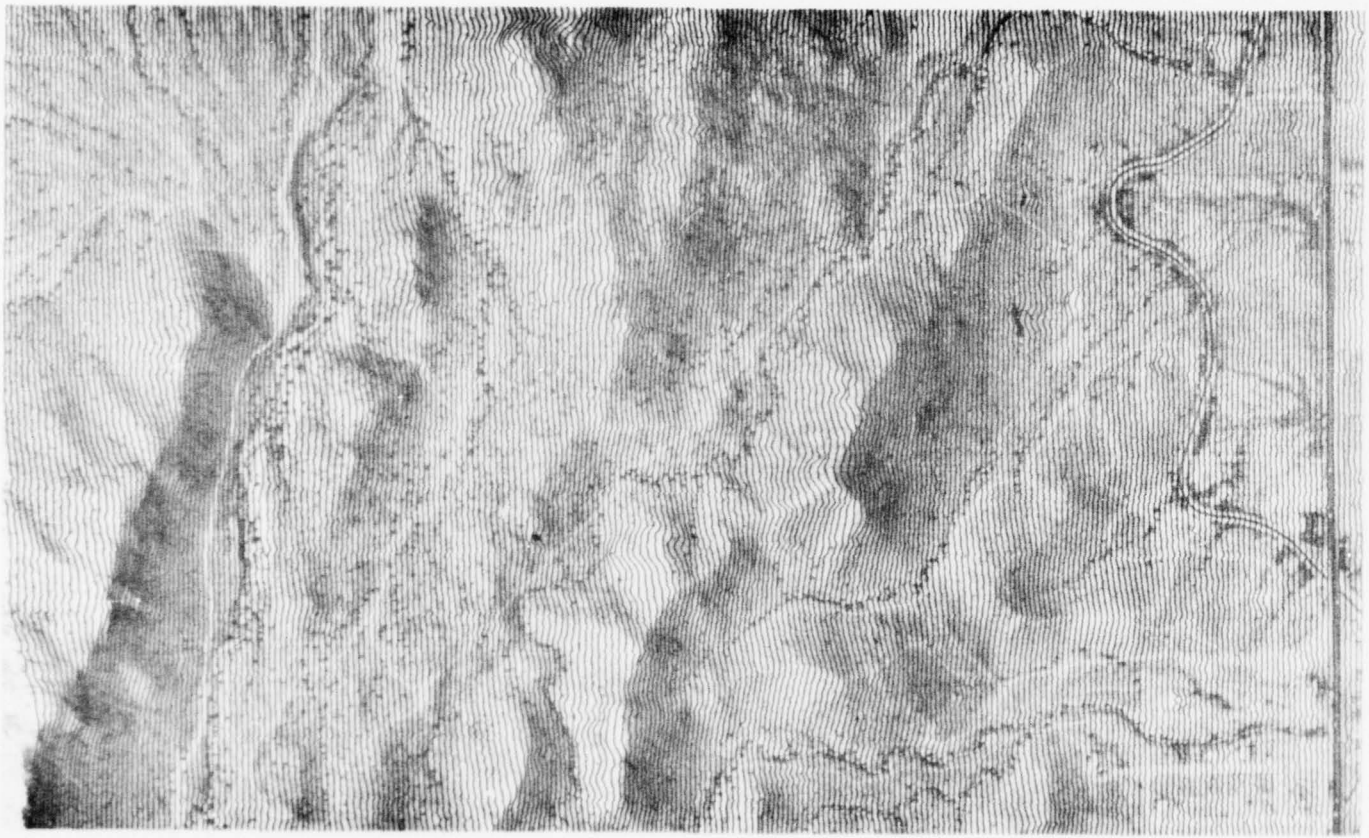


Figure 6-2 Image B with Conjugate Grid Superimposed

image superimposed on image A appears in Figure 6-3. Neither the terrain data nor the contour lines here have been smoothed. The contouring procedure makes use of local surfaces defined by bicubic polynomials that fit the data exactly; least squares techniques are not used. Since this contour image is generated in image A space, it is unrectified with respect to model or object space. However, the contour labels on the edges do represent actual feet above sea level on the ground.

Figure 6-4 is a plot of the reliability factor for each match point superimposed on a lightened version of image A. The more unreliable a match point is with respect to the five reliability criteria, the darker is its gray-scale value in this picture. For this processing example, of the 53,910 match points processed, 72% are reliable. Most of the unreliable areas observed in the picture are due to a low standard deviation of the image intensity in those areas. In these pictures, north is to the left. A large number of unreliable areas consistently fall on the north slopes of the mountain ranges. These slopes are highly illuminated and notably lacking in feature content. A specific example is the circular area at the top edge of the picture. Looking back at Figure 6-1, it is evident that this unreliable area lies on a steeply sloping, conical peak whose entire northern exposure is rather featureless. Matching was very difficult in this area, and the mismatching is manifested in the false contouring of this peak in Figure 6-3. However, the matching process did not break down completely after encountering this area. The match point reliability increased as the feature content improved on the other side of the peak. Other unreliable areas were not as dramatically inaccurate; in fact, some- though flagged as unreliable-were remarkably accurate.

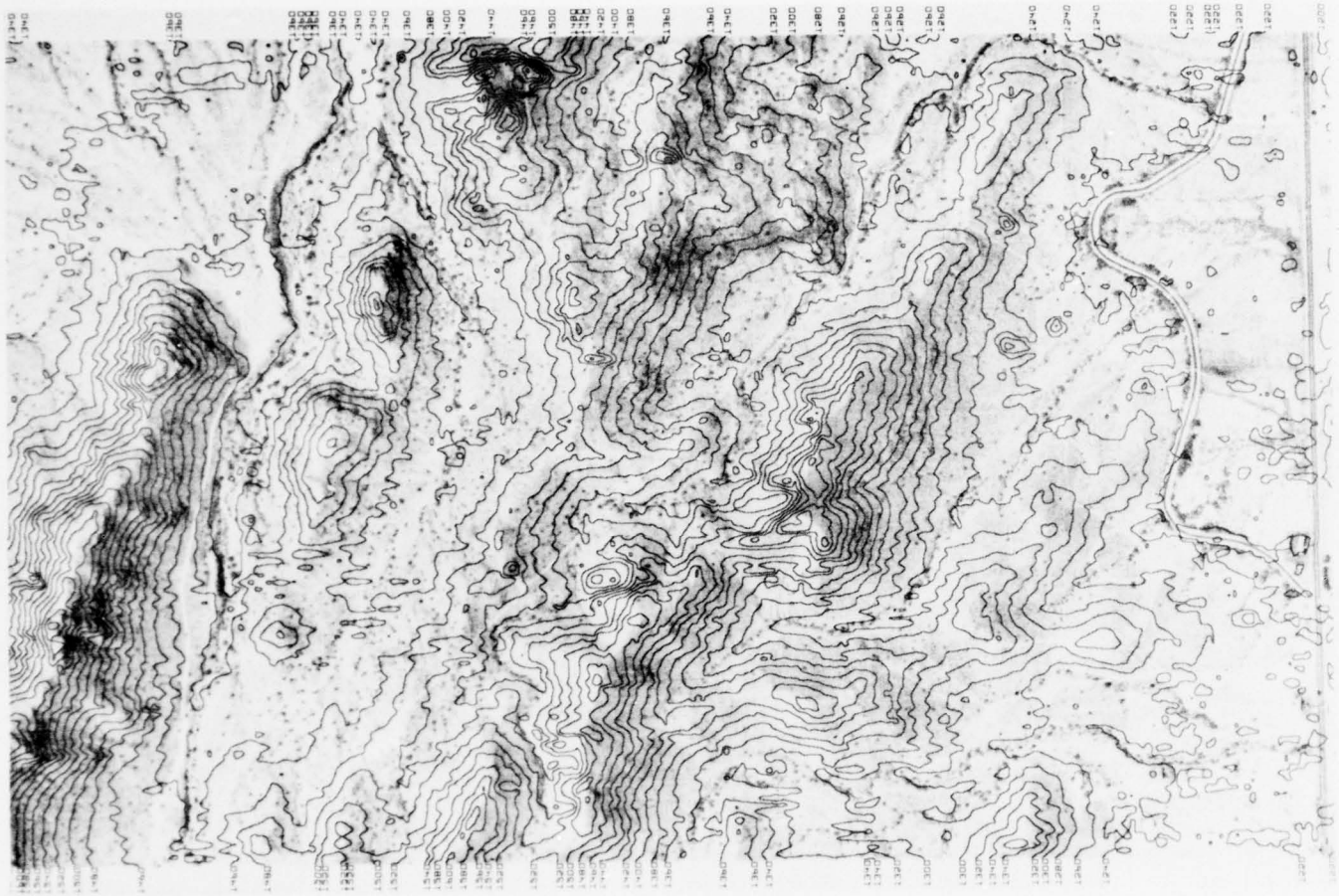


Figure 6-3 Image A with Digital Contours Superimposed

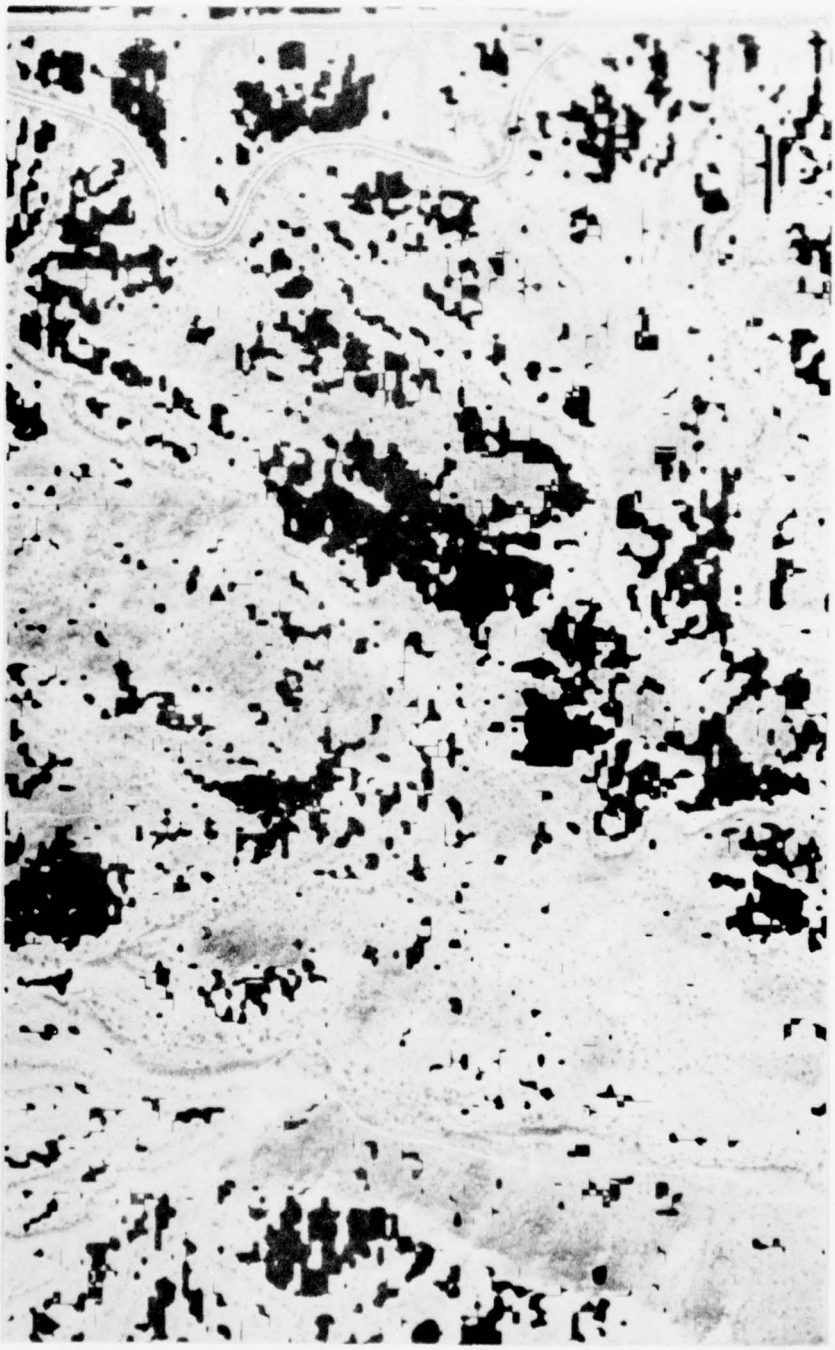


Figure 6-4 Image A with Reliability Plot Superimposed

7.0 CONCLUSION

All of the strategies that collectively make up the stereo matching algorithm described in this report have analytical justification and are theoretically predictable. However, in the face of real imagery of various kinds, there are so many variables involved relating sensor geometry, exposure conditions, and terrain and image characteristics that at times the algorithm appears to have a mind of its own. Algorithm tuning, at this stage, is rather an art than a science or engineering exercise. Trial and error procedures based on the accumulated experience of the user-analyst are required to obtain maximum algorithm performance.

If it is at all feasible or practical to tabularize and parameterize image, sensor, and terrain events, then the tuning of the algorithm to the conditions can be made more automatic and reside more in the realm of engineering than art. This is the subject of study in the next phase of the contract.

APPENDIX A

DIGITAL INTERIOR ORIENTATION: A PROCEDURE FOR REDUCING DIGITAL SCAN COORDINATES TO CALIBRATED PHOTO COORDINATES

DIGITAL INTERIOR ORIENTATION

A PROCEDURE FOR REDUCING DIGITAL SCAN
COORDINATES TO CALIBRATED PHOTO COORDINATES

A procedure for reducing digital scan data coordinates to calibrated photo coordinates is necessary when photogrammetric computations are to be performed with digital image data.

Three distinct coordinate systems are to be considered:

- digital scan coordinate system
- reseau coordinate system
- fiducial coordinate system

Basic image manipulations are typically performed in the digital scan coordinate system; that is, in terms of scan lines and pixels. A reseau system is typically deposited on mapping images to remove film distortion. The fiducial system defines the optical axes of the mapping sensor and forms the base of all photogrammetric computation. The task then of digital interior orientation is to transform the digital scan coordinate system through the reseau system to the calibrated fiducial system. The effect of this transformation is to remove any film distortion that exists on the original film image and also to correct for scanner induced distortions.

Figure 1 describes the marks to be used in the transformation procedure. Fiducial marks A, B, C, and D are primary fiducials, whose intersection defines the principal point of the photo coordinate system. Fiducials E, F, G, and H are secondary fiducials.

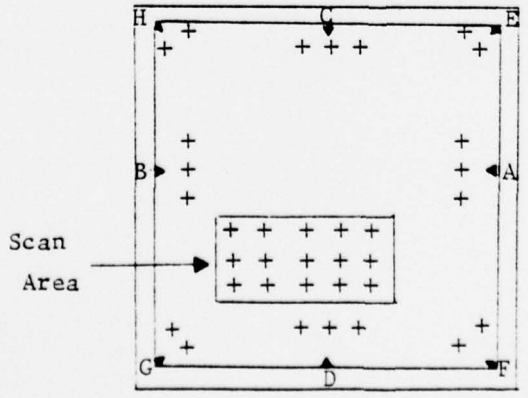


Figure 1

The calibrated coordinates of all of these fiducial marks are generally supplied in the calibration certificate for a particular mapping camera and lens. The shape of the fiducial marks and the reseau pattern varies with the type of camera. The reseau marks are generally etched on a glass plate in front of the film and deposited on the film at the time of exposure.

The calibration procedure for digital scan coordinates is as follows:

1. On the original film image measure a series of reseau marks for each local neighborhood around the fiducial marks.

   Using these measurements construct for each neighborhood a transformation T_R from the measured reseau marks to the calibrated reseau marks:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} T_R \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

where x',y' are the coordinates of the calibrated marks and x,y the coordinates of the measured marks.

This reseau transformation can take on one of the following forms:

$$x' = a_1 + a_2x + a_3y \qquad x' = a_1 + a_2x + a_3y + a_4xy \tag{1-1}$$

$$y' = b_1 + b_2x + b_3y \qquad y' = b_1 + b_2x + b_3y + b_4xy \tag{1-2}$$

$$x' = a_1 + a_2x + a_3y \qquad x' = \frac{a_1 + a_2x + a_3y}{1 + c_1x + c_2y} \tag{1-3}$$

$$y' = b_1 - a_3x + a_2y \qquad y' = \frac{b_1 + b_2x + b_3y}{1 + c_1x + c_2y} \tag{1-4}$$

The type of transformation selected is dependent on the number of marks measured and the nature of film distortion anticipated.

2. Measure the fiducial marks on the film and transform them to the calibrated reseau system using the previously derived transformations:

$$\begin{bmatrix} x'_F \\ y'_F \end{bmatrix} = \begin{bmatrix} T_R \end{bmatrix} \begin{bmatrix} x_F \\ y_F \end{bmatrix}$$

where x_F, y_F are the measured coordinates of a fiducial mark and x'_F, y'_F are the coordinates of the mark with respect to the calibrated reseau system.

This transformation removes the effect of film distortion on the individual fiducial marks.

3. Using the above-derived fiducial coordinates calibrated with respect to the reseau, next construct the transformation T_{RF} from the calibrated reseau system to the calibrated fiducial system:

$$\begin{bmatrix} x''_F \\ y''_F \end{bmatrix} = \begin{bmatrix} T_{RF} \end{bmatrix} \begin{bmatrix} x'_F \\ y'_F \end{bmatrix}$$

where x'_F, y'_F are the fiducial coordinates derived above and x''_F, y''_F are the calibrated fiducial coordinates supplied in the camera calibration certificate.

Again, the coefficients of the transformation T_{RF} can take the forms (1-1) to (1-4) mentioned above.

4. On the digital image corresponding to the scan area of Figure 1 measure a series of reseau marks and express the coordinates in the digital scan coordinate system.

Using these measurements, construct a transformation T_{SR} from the digital scan coordinate system to the calibrated reseau system:

$$\begin{bmatrix} x_{CR} \\ y_{CR} \end{bmatrix} = \begin{bmatrix} T_{SR} \end{bmatrix} \begin{bmatrix} I \\ J \end{bmatrix}$$

where I,J are the measured digital coordinates of the selected reseau marks and x_{CR}, y_{CR} are the corresponding coordinates of the calibrated reseau marks.

Again, the transformation T_{SR} can take the forms (1-1) to (1-4).

The effect of this transformation is to remove both the original film distortion local to the desired scan area and any geometric distortion introduced in the scanning and digitization process.

5. Construct the composite transformation T_{SF} from the digital scan coordinate system to the calibrated fiducial system:

$$\left[T_{SF}\right] = \left[T_{RF}\right] \boxdot \left[T_{SR}\right]$$

where $\boxdot$ is a transformation composition operation (typically a matrix multiply).

The overall result of this calibration procedure is that any pixel (or point) whose coordinates are I,J in the digital scan coordinate system can be related to the calibrated photo coordinate system for subsequent photogrammetric computations. This relation is:

$$\begin{bmatrix} x_P \\ y_P \end{bmatrix} = \left[T_{SF}\right] \begin{bmatrix} I_P \\ J_P \end{bmatrix}$$

where x_p, y_p are the calibrated photo coordinates of point P.

If a particular photogrammetric application requires the transformation from photo coordinates to digital scan coordinates, then the operation may be inverted, such that

$$\begin{bmatrix} I_P \\ J_P \end{bmatrix} = \begin{bmatrix} T_{FS} \end{bmatrix} \begin{bmatrix} x_P \\ y_P \end{bmatrix}$$

where

$$\begin{bmatrix} T_{FS} \end{bmatrix} = \begin{bmatrix} T_{SF} \end{bmatrix}^{-1}$$

APPENDIX B

EPIPOLAR LINE DETERMINATION

EPIPOLAR LINE DETERMINATION

Consider the following relative orientation elements that relate exposure #1 to exposure #2.

(B_x, B_y, B_z), the coordinates of exposure station #2 with respect to exposure station #1.

$(W_{21}, \phi_{21}, K_{21})$, the coordinate rotation angles that relate photo coordinate system #2 to photo coordinate system #1.

$(x_1, y_1, -f)$, the coordinates of a point in photo coordinate system #1 through which an epipolar plane is to be determined.

It is assumed that the focal length, f, is the same for both exposures.

The equation of the epipolar plane in system #1 is:

$$A_x + B_y + C_z = 0$$

where $A = -(f\, B_y + y_1\, B_z)$

$$B = f\, B_x + x_1\, B_z$$

$$C = y_1\, B_x - x_1\, B_y$$

Then, the equation of the epipolar line in system #1 passing through $(x_1, y_1, -f)$ is:

$$y = -\frac{A}{B}\, x + \frac{C}{B}\, f$$

Now, let $$\begin{bmatrix} B_x^* \\ B_y^* \\ B_z^* \end{bmatrix} = \begin{bmatrix} A_{21} \end{bmatrix}^T \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix}$$

be the baseline vector expressed in a coordinate system parallel to photo coordinate system #2, where A_{21} is the 3 X 3 rotation matrix derived from W_{21}, ϕ_{21}, K_{21}.

Let $$\begin{bmatrix} x_1^* \\ y_1^* \\ z_1^* \end{bmatrix} = \begin{bmatrix} A_{21} \end{bmatrix}^T \begin{bmatrix} x_1 \\ y_1 \\ -f \end{bmatrix}$$

be the point of interest system #1 expressed in a coordinate system parallel to system #2.

Then, the equation of the epipolar plane in system #2 is:

$$D_x + E_y + F_z = 0$$

where $$D = z_1^* B_y^* - y_1^* B_z^*$$

$$E = -(z_1^* B_1^* - x_1^* B_z^*)$$

$$F = y_1^* B_x^* - x_1^* B_y^*$$

The equation of the conjugate epipolar line in system #2 is:

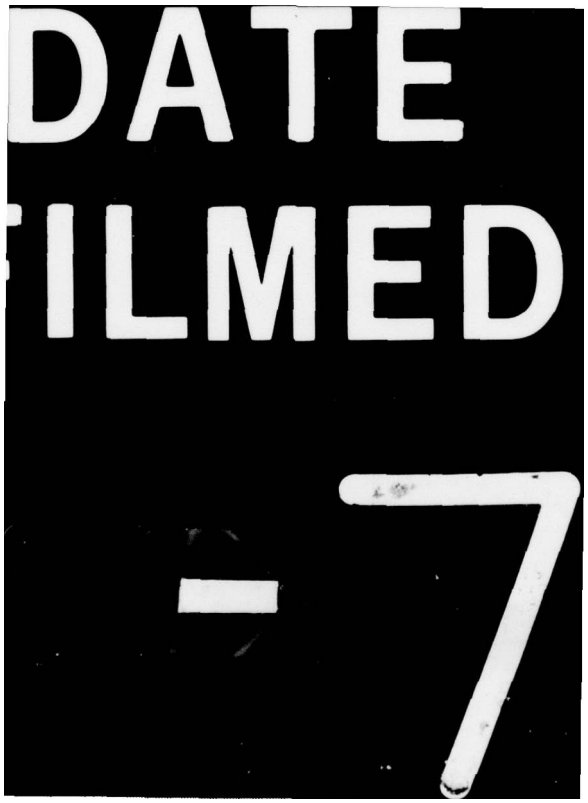
$$y = -\frac{D}{E} x + \frac{F}{E} f$$

REFERENCES

1. Panton, D. J. and Murphy, M. E., "Digital Cartographic Study and Benchmark: First Interim Technical Report," to USAETL, October, 1975.

2. Panton, D. J. and Murphy, M. E., "Digital Cartographic Study and Benchmark: Second Interim Technical Report," to USAETL, December, 1975.

3. Panton, D. J. and Murphy, M. E., "Digital Cartographic Study and Benchmark: Third Interim Technical Report," to USAETL, September, 1976.


DATE
FILMED
-7