N87-26547

DEVELOPMENT OF A STEREO ANALYSIS ALGORITHM FOR GENERATING TOPOGRAPHIC MAPS USING INTERACTIVE TECHNIQUES ON THE MPP

James P. Strong
Space Data and Computing Division
NASA Goddard Space Flight Center
Greenbelt, Maryland 20771

ABSTRACT

The analysis of stereo images to determine depth or elevation is an example of the general problem of the detection of local movement or distortion when comparing two images. Examples of tasks that require solutions to this problem include topographical map generation, object tracking, autonomous vehicular guidance, and robotic vision systems. The key part of the problem is the matching of corresponding areas in the two images. While the human vision system is very good at this task, automated techniques have proven to be computationally expensive and not practical with standard computers. The thrust of the work presented in this paper is the development of a local area matching algorithm on the Massively Parallel Processor (MPP). It is an iterative technique that first matches coarse or low resolution areas and at each iteration performs matches of higher resolution. This is similar to what has been demonstrated to happen in human visual systems by Marr and Poggio, (1971). Results so far show that when good matches are possible in the two images, the MPP algorithm matches corresponding areas as well as a human observer. To aid in developing this algorithm, a control or shell program has been developed for the MPP that allows interactive experimentation with various parameters and procedures to be used in the matching process. (This would not be possible without the high speed of the MPP). With the system, optimal techniques can be developed for different types of matching problems.

INTRODUCTION

During October 1984, the Space Shuttle Challenger was flown with a Shuttle Imaging Radar instrument (SIR-B). One of the experiments during this mission was to obtain overlapping images of an area on the ground viewed from several different incidence angles. Any two of these images form "pseudo-stereo-pairs" which through a suitable geometric model can be used to compute surface elevations. The paper reports current results of an effort at the Goddard Space Flight Center to develop an automated algorithm for computing elevations from SIR-B image pairs using the Massively Parallel Processor.

Background

Historically, the derivation of elevations from stereo pairs has followed two general approaches: the contour and the profile. With the contour approach, the stereo pairs are adjusted in a viewer such that only objects at a certain height will overlap perfectly. The interpreter than traces the path of perfect overlap. In the profile approach, the spacing between the stereo pairs is adjusted until a given object overlaps perfectly in the two images. The height of the object is then obtained as a function of that spacing. Objects along a given line are usually matched with this technique and thus the term "profile". The second approach has been adopted for implementation on the MPP.

Difficulties in Stereo Matching

The major difficulties in detecting points or areas where perfect overlap occurs (i.e., matching of corresponding pixels in the two images) are:

1. Different brightness levels in the two images.
2. Local distortions of the image.
3. Low contrast areas and noise.

The first difficulty is often obviated by the use of normalized correlation functions for matching grey level or edge images. The second is inherent in stereo analysis because of the different viewing angles and makes automated matching more difficult than, for instance, in the case of matching control-point chips in Landsat images. It occurs most severely in regions of rapidly changing terrain and creates a horizontally stretched or compressed area surrounding corresponding pixels in one image relative to the other. Because of the large off-nadir viewing angle and difference in viewing angle required to achieve reasonable accuracy, this problem is particularly acute in radar images. Thus, the basic clue used to determine the elevation also makes the determination of that elevation more difficult.

Any techniques for correcting local distortions must take into account the fact that the distortion function can have a broad band of spatial frequencies. For example, the distortion function for a mountain range would have low frequencies, but added to these would be high frequencies caused by rock formations making up the surface. When a human observer fuses two images seen through a viewer, the low frequency information is used to obtain an initial fusion in which the eyes are brought into alignment (a technique used for automatic focusing of some cameras) and then high frequency information brings out a detailed perception of depth. The progression from low to high frequency suggests that a hierarchical approach for detecting matching pixels would be appropriate. With this approach, an initial match is performed on low frequency information in an image and then increasingly higher freqencies are incorporated to obtain the final matching of corresponding pixels.

Even with no local distortion, errors can occur due to noise, spatial periodicities and low contrast in the image. One way of reducing errors in general is to provide redundancy by computing matches at nearly every pixel in the image. Then continuity constraints on the ground surface can be used to correct local discontinuities in the elevation. Matching at nearly every pixel is a formidable task on standard serial computers. However, the architecture of the MPP is well suited to the local neighborhood operations required for matching pixels. The resulting speed allows iterations of the matching algorithm computed at every pixel in 512 x 512 images to be completed within seconds. The following sections discuss the matching technique developed for the MPP and results obtained using the MPP algorithm.

MATCHING TECHNIQUE

The matching algorithm developed for the MPP is an example of what has been termed the Hierarchical Warp Stereo (HWS) technique. Initial work on stereo analysis using a hierarchical approach was done by Marr and Poggio (1979). The Marr-Poggio algorithm performs low pass filtering and edge detection on the two stereo images and then matches the edges. Filters of several bandwidths are used from 1/100 to 1/4 the highest frequency. An edge in one image is said to match an edge in the other if 1) that edge appears within a given search area, 2) the slopes of the two edges match and 3) the direction of the change in brightness of the two filtered images is the same across the edge. The relative location of edges in the most highly (narrowest bandwidth) filtered image determines the relative positions of large objects in a scene (For instance, mountains or mountain ranges.) These displacements are then used to define search areas for corresponding edges in the second most highly filtered image. The procedure is repeated until corresponding edges are matched in the least filtered image.

More recent work has been reported by Quam (1984) who processes multiple resolution versions of both images (by sub-sampling by powers of 2 in both directions). Starting at the lowest resolution, matches are calculated using a normalized correlation measure applied to neighborhoods or windows surrounding reference and test pixels. The disparities between corresponding pixels (i.e., difference in their location in the two images) are then used as a one dimensional distortion function to warp one image (the test image) so that its matched pixels will be in the same location as in the other (the reference image). The warped image is then resampled at the next higher resolution. This cycle is repeated until the highest resolution images are matched. The warping operation at each iteration reduces the local distortion so that at the next iteration with the next higher resolution there is a higher probability of obtaining a good match between pixels. At the end of the process, the sum of the distortion functions from all iterations forms the disparity function used to compute elevations. Quam's algorithm also eliminates potentially bad matches at each iteration by interpolating across pixels with low values of maximum "match scores" obtained for the reference neighborhoods over the corresponding search areas in the test image.

MATCHING ALGORITHM ON THE MPP

The algorithm developed for the MPP is similar to the Quam algorithm with the following exceptions:

1. Instead of applying equal sized windows to versions of the input images at increasing resolution, decreasing sized windows are applied to each iteration to the original input images. This eliminates the need sub-sampling.

2. Iterations repeat until the neighborhoods are of a size within which there is no useful information for correlation.

3. At each iteration, the net amount of warping (i.e., an updated disparity function) is computed. This net disparity is always applied as the warping function to the original test image which eliminates loss of information at each warp iteration.

4. At each iteration, areas where bad matches occur are detected and interpolated over. Then the disparity function is smoothed before being used for warping.

The matching algorithm consists of the following steps:

1. Preprocessing of the test image,

2. Determination of matches,

3. Removal of "bad match" areas in the disparity function,

4. Smoothing the resulting disparity function,

5. Warping the test image.

The steps 2 through 5 are repeated for each iteration. The following subsections discuss each of these steps in detail.

Preprocessing of the Test Image

Because of the viewing geometry, the resolutions of the two SIR-B images are different in the stereopsis direction. Thus, a linear scale change is applied to the test image so that its resolution is the same as that of the reference image. This operation is implemented using a linear warping function in the stereopsis direction. Second, the test image can be translated to reduce the absolute value of the maximum disparity between the two images. Translation is effected by making the warping function constant over the image. When this is done, the size of the initial search area can be reduced. The amount of initial warping can be determined mathematically based on the different synthetic aperture radar incidence angles or it can be determined interactively by displaying the reference and test images. In either case, the initial warping function is incorporated into the net disparity function when determining elevations.

Determination of Matches

For each reference image pixel, a match is performed between a neighborhood surrounding that pixel (the "template") and neighborhoods within a search area in the test image. The location of the center pixel within each neighborhood in the test image relative to the reference pixel is the disparity value associated with that neighborhood.

The measure used for matching neighborhoods is the normalized mean and variance correlation given by:

Where X_i and Y_{i-k} are grey levels of the i^{th} pixels within the template neighborhood and the k-th in the search area respectively. The values $\bar{X}$ and $\bar{Y}_k$ are the mean values computed over the template and K-th search area neighborhoods.

For each pixel in the reference image, the match score for all neighborhoods within the search area is computed. The pixel at the center of the neighborhood with the highest correlation value or match score is selected as the matching pixel. The resulting disparity function is, therefore, made up of integer values.

Removal of "Bad Match" Areas in the Disparity Function

In order for stereo analysis to produce correct topographic results, there must be a one to one correspondence between pixels in the test image and those in the reference image (at least down to the resolution required to produce the desired elevation accuracy). For synthetic aperture radar images, this means that, 1) both images must be taken from the same side of the spacecraft (or aircraft) , and 2) both incidence angles must be such that there is no "layover" due to large surface slopes, and that there are no "shadows". If the image having the larger incidence angle is used as the reference image and both images meet the above two requirements, it can be proven that the disparity function will always have a gradient (or slope in the stereopsis direction) between 0 and 1. If "ground range" images are used (where both images have the same pixel resolution), the slope of the

$$\text{Match Score }(k) = \frac{\sum_i (X_i - \bar{X}) \times (Y_{i-k} - \bar{Y}_k)}{\text{SQRT}(\sum_i (X_i - \bar{X})^2) \times \text{SQRT}(\sum_i (Y_{i-k} - \bar{Y}_k)^2)}$$

disparity function will always be between ± 1. This result is necessary if there is to be a one to one mapping between the test image and its warped version when the disparity is used as the warping function. Since the slope of the disparity function must be between ±1, a simple test for a bad match is to observe adjacent values of the disparity function and determine if there is a jump of more than 1 pixel.

The human visual system appears to have the capability of interpolating surfaces over areas where bad matches occur. This process is emulated in the MPP algorithm by interpolating the disparity function across all areas where a bad match has been detected. The detection of bad matches and the interpolation are accomplished with the following operations:

1. Detection of "bad match pixels" (pixels where discontinuities occur)

Sudden jumps in the disparity function are detected by examining a 3 x 3 neighborhood surrounding each pixel. If the disparity value at the pixel differs from that of any of its neighbors by more than one, the pixel is identified as having a bad match.

2. Expansion or growth of each bad match pixel to form a neighborhood

If the maximum difference between the disparity value at a pixel and those of its adjacent neighbors is "d", then one must interpolate the disparity function over a neighborhood surrounding that pixel whose diameter is at least "d" to satisfy the constraint that the slope of the disparity function be less than +1. The expansion of each "bad match pixel" to form a neighborhood is done for this purpose. This is accomplished on the MPP by alternately expanding each pixel into 4 and 8 element neighborhoods. The resulting neighborhood is octagonally shaped. A diameter of 2N is achieved with N iterations. As each pixel is expanded, the resulting overlapping neighborhoods form bad match regions.

3. Interpolation of the disparity function over resulting bad match neighborhoods

Interpolation is performed using heat flow equations. The architecture of the MPP is well suited to the iterative solution of the boundary value partial differential equations typical of heat flow problems. To perform the interpolation, two dimensional heat flow partial differential equations are applied to solve for the steady state "temperature" or disparity in the bad match regions assuming that the bordering pixels surrounding the bad match area are held at a constant "temperature" or disparity. The equations used for obtaining the interpolated disparity at pixel [i,j] at iteration t+1 from the values at iteration t are:

$$D(i,j,t+1) = D(i,j,t) + \frac{d(D(i,j,t))}{dt} \quad \text{where}$$

$$\frac{d(D(i,j,t))}{dt} = \frac{\partial^2(D(i,j,t))}{\partial i^2} = \frac{\partial^2(D(i,j,t))}{\partial j^2}$$

The second partial differential equations reduce to

$$\frac{d(D(i,j,t))}{dt} = D(i,j+1,t) + D(i,j-1,t) + D(i+1,j,t) + D(i-1,i,t) - 4D(i,j,t)$$

when is is assumed that the two dimensional grid increments are unity. The number of iterations required to reach "steady state" is dependent on the size of the bad match regions.

Smoothing the Resulting Disparity Function

After interpolation, a smoothing operation is applied over the whole disparity function in order to obtain a smoother warping function with fractional pixel values rather than integer values. Smoothing is performed on the MPP by averaging over a neighborhood proportional in size of the neighborhood used to obtain the disparity function.

Warping the Test Image

The smoothed disparity function is used as a one dimensional distortion function to "geometrically correct" the test image in the stereopsis direction. The brightness values in the warped image are obtained by applying a linear interpolation function to the test image data in the resampling process.

INTERACTIVE OPERATIONS ON THE MPP

The matching algorithm on the MPP has been implemented to be run in an interactive mode where parameters such as neighborhood sizes, search area, and discontinuity thresholds can be input before starting each step of the matching algorithm. At the end of a given step, the results (such as the disparity function, or the warped version of the test image) can be immediately displayed or saved on disk. This interaction provides the ability to experiment with various parameter values and quickly observe the results. In addition, the control or shell program is designed so that operations can be easily modified or added.

The operations presently implemented in the matching algorithm which can be run interactively are shown below along with the input parameters which can be selected:

1. INITIAL WARP (left edge movement, right edge movement)

2. MATCH (neighborhood size, search area size)

3. "BAD MATCH DETECTOR" (discontinuity threshold for bad match, neighborhood diameter for expansion)

4. INTERPOLATE (number of iterations)

5. SMOOTH (neighborhood size)

6. WARP

The left edge and right edge movement in the INITIAL WARP operation define the linear warp function in the stereopsis direction. If they are equal, only a translation is applied to the test image.

Interactive Turn Around Time

The following table shows turn around times for the six stereo matching operations. Where times are dependent on parameters, some example parameters and the corresponding times are shown. The times are measured from the time a key is pressed on the terminal to start the task to the time the next prompt is displayed indicating that the task is finished. Times less than about a half second could not easily be measured.

The matching operation which is the most computationally expensive task has been optimized to require a time proportional to the length of the sides of the neighborhoods as opposed to the area. The smoothing operation is unoptimized and requires a time proportional to the area of the neighborhood. However, since the time required for smoothing is not prohibitively long for interactive purposes, this optimization has not yet been implemented.

INTERACTIVE TURN AROUND TIME FOR STEREO ANALYSIS OPERATIONS

Operation	Parameter		Time in Seconds
Initial Warp			0.5
Match	11 x 11	Nbh. size	5
	21 x 21		10
	41 x 41		20
Detect Bad Match	2	Radius	< 0.5
	4		0.5
	8		1
Interpolate	250	No. of Iterations	2
	500		4
	1000		8
Smooth	11 x 11		2
	21 x 21		6
	41 x 41		21
Warp			3

RESULTS AND CONCLUSIONS

The matching algorithm has been tested on overlapping SIR-B images with incidence angles of 25 and 42 degrees taken over a plateau region in Northern India at its border with Bangladesh. The signal to noise ratio is high almost everywhere in these two images. The results of the matching algorithm on these imaqes are illustrated in Color Plates II & III. Figures 1a and 1b show the reference and test image. Viewing figures 1a and 1b stereoptically, one can observe the plateau region and the river valley around it. Figure 2a is the reference image again and figure 2b is the test image after it has been warped during the matching process. If one views figures 2a and 2b steroptically, it will be seen that there is virtually no depth since the warped test image matches the reference image very well. Two iterations of matching and warping were required to obtain this image. The first iteration used a 25 pixel square correlation neighborhood and the second, a 13 pixel square neighborhood. Further iterations with smaller neighborhoods 10 pixels square or less yielded too many discontuities. This indicates that for many areas in the particular SIR-B images used, there is insufficient information in neighborhoods smaller than about 10 pixels square to produce good correlation. Figure 3 is the two dimensional disparity function derived during the matching process. Dark areas in this image are where pixels in the reference image lie to the right of their correspondings pixel in the test image. In the light areas, the opposite is the case. The disparity function is approximately linearly proportional to the actual elevations in the images with the dark areas in figure 3 corresponding to low elevations and the lighter areas to higher elevations. In figure 4, a three-dimensional perspective view is presented of the disparity function. In generating this view, the disparity function was treated as a two dimensional surface illuminated by a light source at approximately the location of the shuttle radar sensor about 42 from vertical. One can easily identify the plateau region and river valleys corresponding to those seen in the original stereo pairs.

A close comparison of figures 2a and 2b shows that the matching portion of stereo image analysis for determining elevation can be accomplished with the MPP stereo matching algorithm as well as human observers in most areas where the local distortion is not too severe. A human capability which can possibly increase the resolution of the matching algorithm is the ability to discern and match edges practically to the nearest pixel. In areas where there are significant edges, one may be able to use smaller neighborhoods for correlation and thus, in these areas, increase the spatial resolution of the elevation data. The inclusion of edge detection and the matching of edges into the algorithm is one of the current changes being implemented on the MPP.

ACKNOWLEDGEMENTS

The author would like to acknowledge the help of Mr. William Haynes and Mr. Edward Seiler of Scientific Applications Research Corporation for aiding in the implementation of the matching algorithm on the MPP as well as Dr. H. K. Ramapriyan for his editorial contributions.

REFERENCES

1. Marr, D. and Poggio, T., "A Computational Theory of Human Stereo Vision", Proceedings, Royal Society of London, 1979, pp. 301-328.

2. Quam, L., "Heirarchical Warp Stereo", Proceedings, Image Understanding Workshop, 1984, pp. 149-155.