Stereo Verification in Aerial Image Analysis

Clifford McVay, Bruce D. Lucas, and
David M. McKeown

Department of Computer Science
Carnegie-Mellon University
Pittsburgh, PA. 15213

July 8, 1985

Abstract

This paper describes a flexible stereo verification system, STEREOSYS, and its application to the analysis of high resolution aerial photography. Stereo verification refers to the verification of hypotheses about a scene by stereo analysis of the scene. Unlike stereo interpretation, stereo verification requires only coarse indications of three-dimensional structure. In the case of aerial photography, this means coarse indications of the heights of objects above their surroundings. This requirement, together with requirements for robustness and for dense height measurements, shape the decision about the stereo system to use. This paper discusses these design issues and details the results of an implementation.

This research was partially sponsored by the Defense Advanced Research Projects Agency (DOD), ARPA Order No. 3597, monitored by the Air Force Avionics Laboratory Under Contract F33615-81-K-1539 and by the Defense Mapping Agency Under Contract DMA 800-85-C-0009. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency, of the Defense Mapping Agency, or the US Government.

Stereo Verification in Aerial Image Analysis

Clifford McVay[1], Bruce D. Lucas[2], and David M. McKeown[3]

Department of Computer Science
Carnegie-Mellon University
Pittsburgh, PA. 15213

July 8, 1985

[1] Clif McVay was a Visiting Scholar from the Defense Mapping Agency, Aerospace Center, St. Louis during the academic year 1984-85.

[2] Bruce Lucas is currently a System Designer with the Information Technology Center (ITC) at Carnegie-Mellon and works in the area of computer graphics for advanced personal workstations.

[3] David McKeown is a Senior Project Scientist in the Department of Computer Science

Table of Contents

List of Figures

List of Tables

1. Introduction

This paper describes a flexible stereo verification system, STEREOSYS, and its application to the analysis of high resolution aerial photography. Stereo verification refers to the verification of hypotheses about a scene by stereo analysis of the scene. Unlike stereo interpretation, stereo verification requires only coarse indications of three-dimensional structure. In the case of aerial photography, this means coarse indications of the heights of objects above their surroundings. This requirement, together with requirements for robustness and for dense height measurements, have shaped the decision about the stereo system to use.

In this research we have attempted to address stereo analysis in a very unconstrained environment. Rather than simply focusing on isolated image analysis where stereo pairs are carefully controlled, we have constructed a system that can automatically perform matching and analysis using arbitrarily selected images. We are motivated by the observation that if knowledge-based image understanding systems are to begin to perform analysis tasks at a level of performance required for mapping and photo interpretation, they must be able to accommodate a much broader range of task uncertainty and complexity than has been previously demonstrated in any research or development system.

Stereo verification deals with a variety of problems that are not ordinarily present in isolated experiments with stereo matching and analysis:

- The selection of an appropriate conjugate image pair from a database of overlapping images based on criteria that will maximize the likelihood for good correspondence.
- The image pairs must be dynamically resampled such that the epipolar assumption (ie., epipolars are scan lines) used in most stereo matching algorithms can be applied.
- The size of the areas to be matched varies greatly; the system design must be flexible and general.
- An initial coarse registration step is necessary because the quality of the correspondence between conjugate pairs varies greatly. In many cases the

magnitude of the initial misregistration is greater than the expected disparity shift.

- In addition to producing a depth map image, the system must analyze the stereo results and generate a symbolic description that provides an estimate of the actual height of the region in question, and the confidence of that estimate.

These requirements, in turn, raise a set of broader research issues:

1. How can an aerial image database be used to automatically generate a useful stereo pair containing an arbitrary region?
2. How can a stereo system handle the misregistration problems inherent in variable sourced image databases?
3. What kind of stereo results are appropriate for use in a verification process?
4. How can stereo results be analyzed so as to reflect not only the presence (or absence) of height but also the inherent reliability of the results?

The results of this research indicate that image/map database issues in stereo verification influence the utility of such an approach as much as the underlying stereo matching algorithm. In fact, they are intimately related. Current stereo matching algorithms require nearly perfectly aligned conjugate images, a situation that is unlikely to occur in outside of the laboratory. We believe that the ability to dynamically select conjugate image pairs from a database based upon the region of interest and knowledge of the requirements of the matching algorithm is required for a fully automated image analysis system. Our results also indicate that stereo analysis can function as a very powerful discriminator in an image understanding system without having to perform shape reconstruction. That is, coarse estimates of height, coupled with confidence in those estimates, can greatly constrain search during image interpretation.

This paper discusses these broader research issues as well as providing the reader with an analysis of the results of our experimentation and details of the actual implementation.

2. Stereo Verification in SPAM

STEREOSYS was developed as a knowledge source for SPAM[1], a rule-based system that uses knowledge from a variety of sources to interpret airport scenes in aerial imagery. Many of the requirements for flexibility in a stereo system arise directly from the fact that STEREOSYS must interact in a larger context, that of the image understanding system. As we move from isolated computer vision experiments to system integration, the performance of particular components must be evaluated within the constraints and context of the overall system. SPAM manages and invokes various specialized low-level image analysis processes that allow it to gather information about regions in the image. These processes include texture analysis, feature alignment and grouping[2], and depth cue generation. SPAM has developed along two lines:

- The addition and refinement of knowledge about airports and procedures for recognition and matching of image-based descriptions to the airport scene model.
- The addition and refinement of low-level image processes that support the SPAM control structures by providing primitive intermediate-level scene descriptions.

STEREOSYS falls into the latter category as it uses stereo to generate a depth map (disparity image) description given a hypothesis region in the image. The role of STEREOSYS in the overall system is to verify hypotheses such as *terminal building, access road, tarmac, parking apron,* and *hangar* by measuring the amount of disparity within a hypothesis region and thereby estimating the likelihood that the region is above or at the ground plane. Further, if the region is deemed to be mostly above the ground, STEREOSYS provides a coarse estimate of the absolute height above the ground. One may contrast this with methods for stereo reconstruction that use feature matching or segment-based techniques: STEREOSYS does not attempt to construct a precise three-dimensional model of the feature within the scene. For the tasks that SPAM requires, for example, the verification of a hangar hypothesis, it is not as important to determine the shape of the roof as much as to reliably determine whether a roof of some type is present. The issue of robustness and reliability in aerial image interpretation is of principal importance since most of the hypotheses generated by the system will not correspond to features in the scene having significant height. Therefore, the ability to refute incorrect hypotheses such as hangar and terminal building by determining there is no apparent height as well as to reliably confirm 'no height' hypotheses in areas such as tarmac and parking aprons puts performance expectations on the stereo system that transcend simple stereo matching.

SPAM invokes STEREOSYS as a result of recognizing one of two situations. First, as a part of low-level information gathering, we might want to test every region generated by the segmentation system[3] having certain shape and size properties to determine whether it has significant height above the ground plane. Second, as a part of high-level disambiguation, there are a variety of cases where spatial constraints derived from the rule-based airport model are unable to distinguish between two competing hypotheses. For example, assume SPAM has found a conflict between two interpretations, "terminal building" and "parking lot". Spatial knowledge would allow these hypotheses to occupy similar spots in the overall scene for a wide variety of airports and, therefore, would not be able alone to resolve the conflict. Another common example are compact two-dimensional regions, such as runup pads and the roofs of maintenance buildings. Shape and size metrics such as compactness and area provide only weak cues in this situation. SPAM specifically recognizes situations where competing hypotheses involve features that can be disambiguated based upon knowledge of their height relative to their surroundings. Since we may often be looking at regions that are primarily at the ground plane, the ability to reliably determine that there is no apparent height difference between the region and its neighborhood is equally important.

In either case, the stereo verification process can be characterized as follows:

1. Given a region *R*1 within a geographic area *A*1 from image *I*1, find an appropriate second image *I*2 that contains a geographic area *A*2 that is the same as geographic area *A*1. STEREOSYS has access to a database of images through primitives provided by the MAPS system[4, 5].
2. Image fragments *A*1 and *A*2 are rectified (warped) and registered (shifted/rotated) into a stereo pair of overlaying geographic rectangles *W*1 and

$W2$.

3. The $W1$-$W2$ stereo pair is processed and the result is analyzed in order to compute confidence values that measure the height of $R1$ relative to its surroundings along with the system's overall confidence in the stereo result.

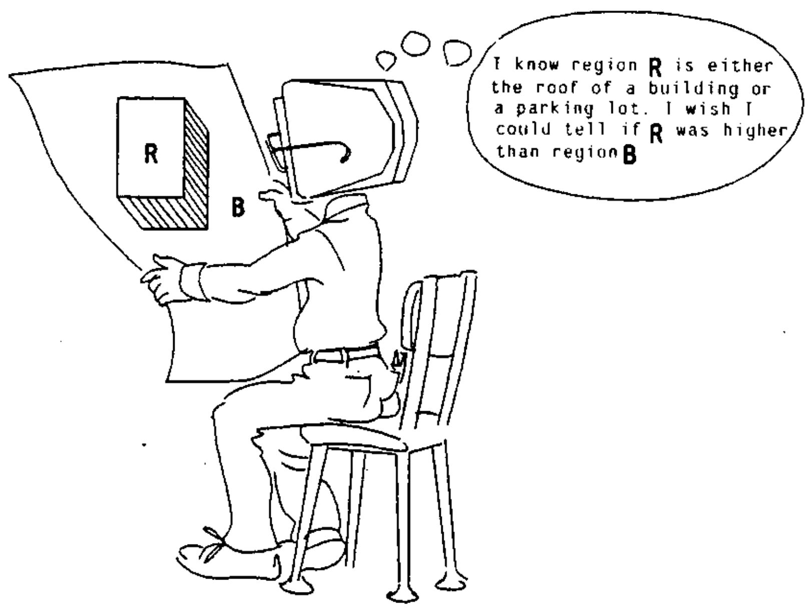


Figure 2-1: Stereo Verification

In the remainder of this paper we will discuss the stereo matching algorithm, how STEREOSYS uses this algorithm to perform stereo verification, and some experimental results that illustrate the strength of this technique as well as some of the more interesting pragmatic problems encountered in complex aerial imagery. Section 3 describes the basic stereo matching process used by STEREOSYS. Section 4 gives the sequence of steps necessary to apply the stereo algorithm to an arbitrarily selected region of an image. Section 5 shows examples of preliminary experiments with S1; the effects of good and poor initial correspondence estimates, the effect of the fine registration step on the subsequent matching, and the evaluation of STEREOSYS over many test regions. Section 6 overviews the strengths and limitations of this work, and suggests future research directions.

3. The Stereo Process

STEREOSYS uses a stereo matching program, S1, described in detail elsewhere[6]. In this Section we will review this stereo matching algorithm. S1 produces a *disparity image (map)* that is registered to the Left stereo pair image and whose pixel values indicate the film plane displacement of matched points in the stereo pair. The disparity value is in one-to-one correspondence with distance, or depth, from the camera and therefore indicates relative height in vertical aerial photography. The process, in effect, correlates neighborhoods about every pixel, but uses the method of differences to avoid costly exhaustive searches.

3.1. Method of Differences

Let $I_1(x,y)$ and $I_2(x,y)$ denote the two images of a stereo pair, and let $h(x,y)$ denote the disparity map. Then the values of the disparity map are a statement that the point (x,y) in I_1 matches the point $(x+h(x,y),y)$ in I_2, that is that

$$I_1(x,y)=I_2(x+h(x,y),y)$$

Let $\bar{h}(x,y)$ denote the correct disparity map. The process begins with a uniform disparity map $h_0(x,y)$, and successively updates the disparity map, yielding h_1, h_2, etc. Ideally, as successive refinements proceed, $h_k \rightarrow \bar{h}$.

Consider a point (x,y) in the left image of the stereo pair; the difference $\bar{h}(x,y)-h_0(x,y)$ between the correct disparity value and our initial estimate is the amount by which the stereo process must correct the disparity in going from h_0 to h_1. Initially this difference will be relatively large because the uniform disparity estimate is not particularly accurate. Because of this, the method of differences requires that we start out with smoothed images to accommodate these large differences. As the disparity estimate h_k improves, we can use less smoothed images because the error between h_k and $\bar{h}$ decreases.

Suppose we have computed a disparity map h_k; that is, we estimate that the point (x,y) in

I_1 matches the point $(x+h_k(x,y),y)$ in I_2. To compute h_{k+1}, we wish to adjust the disparity at each point (x,y) by an amount $\delta(x,y)$ so that the difference between the images is made as small as possible, that is

$$I_1(x,y)-I_2(x+h_k(x,y)+\delta(x,y),y)$$

is minimized. Minimizing this quantity directly involves a costly search over the possible values of δ. Instead, the method of differences estimates this quantity by using derivatives:

$$I_1(x,y)-I_2(x+h_k(x,y),y)+\delta(x,y)D_xI_2(x+h_k(x,y),y)$$

where D_x denotes derivative w.r.t. x.

This quantity is linear in $\delta(x,y)$, as illustrated in Figure 3-1. It could be minimized directly, but we get better results by combining many such estimates from each point in the neighborhood of (x,y) using a least squares technique, and then minimizing. In any case, the estimate based on derivatives is valid only over a range around $x+h$ on the order of the size of the averaging window that has been used to smooth the image. But to be useful we require that this estimate be accurate over a range of at least δ, the discrepancy between the actual disparity and our disparity estimate. Thus because the initial disparity error is large, we must start with relatively smoothed images. For example, some of our images require an adjustment on the order of 15 pixels between the initial disparity estimate and the actual disparity, and so STEREOSYS begins with 32 by 32 smoothing windows.

3.2. Some Pragmatic Issues in Stereo Matching

S1 is also capable of computing a global registration shift between a stereo image pair, also by the method of differences. That is, a global offset can be obtained that indicates how much one image is translated, or shifted, relative to the other. This capability can often salvage the analysis of misregistered stereo pairs and is very attractive for use with SPAM since the underlying MAPS database does not have the image control necessary to

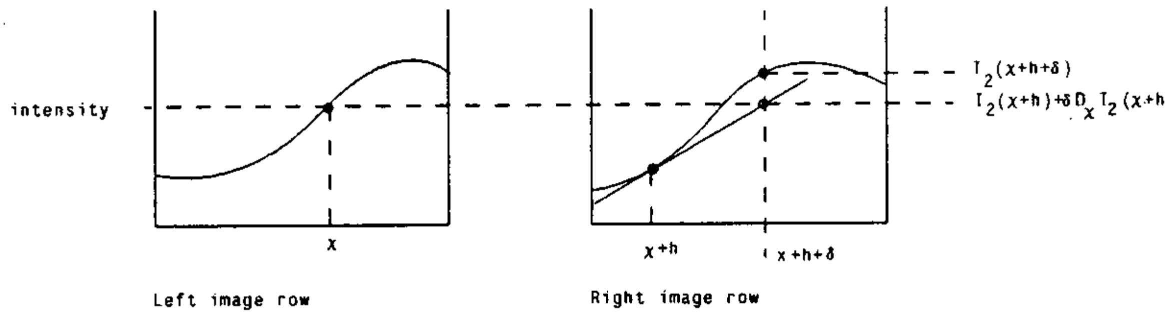


Figure 3-1: Estimating Disparity

guarantee accurately registered stereo pairs.

S1 does not involve the use of sensitive feature extraction thresholds. Stereo matching in S1 is accomplished for every pixel and is not restricted to selected image features such as interesting areas[7], edges[8, 9] or other extracted features[10]. Limiting a stereo procedure to matching extracted image features makes the process sensitive to the extraction technique and its associated thresholds. Since SPAM will be using a stereo process over a wide range of images and regions, such extraction thresholds should be avoided wherever possible.

Another issue in the selection of S1 for use by SPAM has to do with the fact that SPAM is not using stereo to recognize objects or build conceptual models from the stereo results. SPAM simply wants to know if the region of interest has height relative to its surroundings. A dense disparity image registered to the image containing the region of interest is an ideal source of data for the analysis necessary to do simple height verification. Almost all other stereo processes we are aware of produce sparse disparity results designed for purposes other than verification. Work by Panton[11] and Henderson[12] provide possible exceptions.

In summary, unlike many other stereo processes, S1 is not overly reliant on perfectly registered stereo pairs taken simultaneously by well parameterized cameras, nor does it require threshold tweaking to accommodate matching of edges or vertices. It produces an easily analyzed dense disparity image. S1 was chosen for use in stereo verification because these properties coincide well with the aerial image analysis domain that SPAM addresses.

4. Using Stereo Verification with an Aerial Image Database

Certain steps are necessary for a stereo process to work automatically as a verification procedure in association with a database of aerial imagery. A block diagram is given in Figure 4-1 that outlines the procedure and shows the interactions between the Image Analysis Process (SPAM) and the Image Database (MAPS). We can loosely organize these steps, beginning with the identification of a region of interest by the image analysis process as:

1. **Select Coverage**: Determine the available alternate images that cover the region of interest. Select the most appropriate alternative(s).
2. **Extract the Stereo Pair**: Extract a stereo pair of the region from the image coverage selected.
3. **Register the Stereo Pair**: Compensate for misalignment errors inherent in the aerial image database. Do any other processing necessary to assure the stereo pair meets any assumptions made by the stereo process.
4. **Run the Stereo Process**: Apply some stereo matching process (eg., S1).
5. **Analyze the Results**: Analyze the stereo results in order to verify if the region of interest has height relative to its surroundings.

The STEREOSYS process is initiated by SPAM with parameters identifying the region of interest, the database image that is being interpreted and contains the region, and an estimated height range (0-5 meters, 0-15 meters, 10-20 meters, etc) for the region.

Using the identity of the region of interest, STEREOSYS extracts the region's centroid, its boundary point list and an associated minimum bounding rectangle (MBR) from the MAPS database. This data is used in determining alternate imagery coverage, in extracting the stereo pair, and in analyzing the stereo results.

The MAPS database is used to produce an unsorted list of images, called a *coverage file*. Each image in the coverage file contains the region of interest. The image being

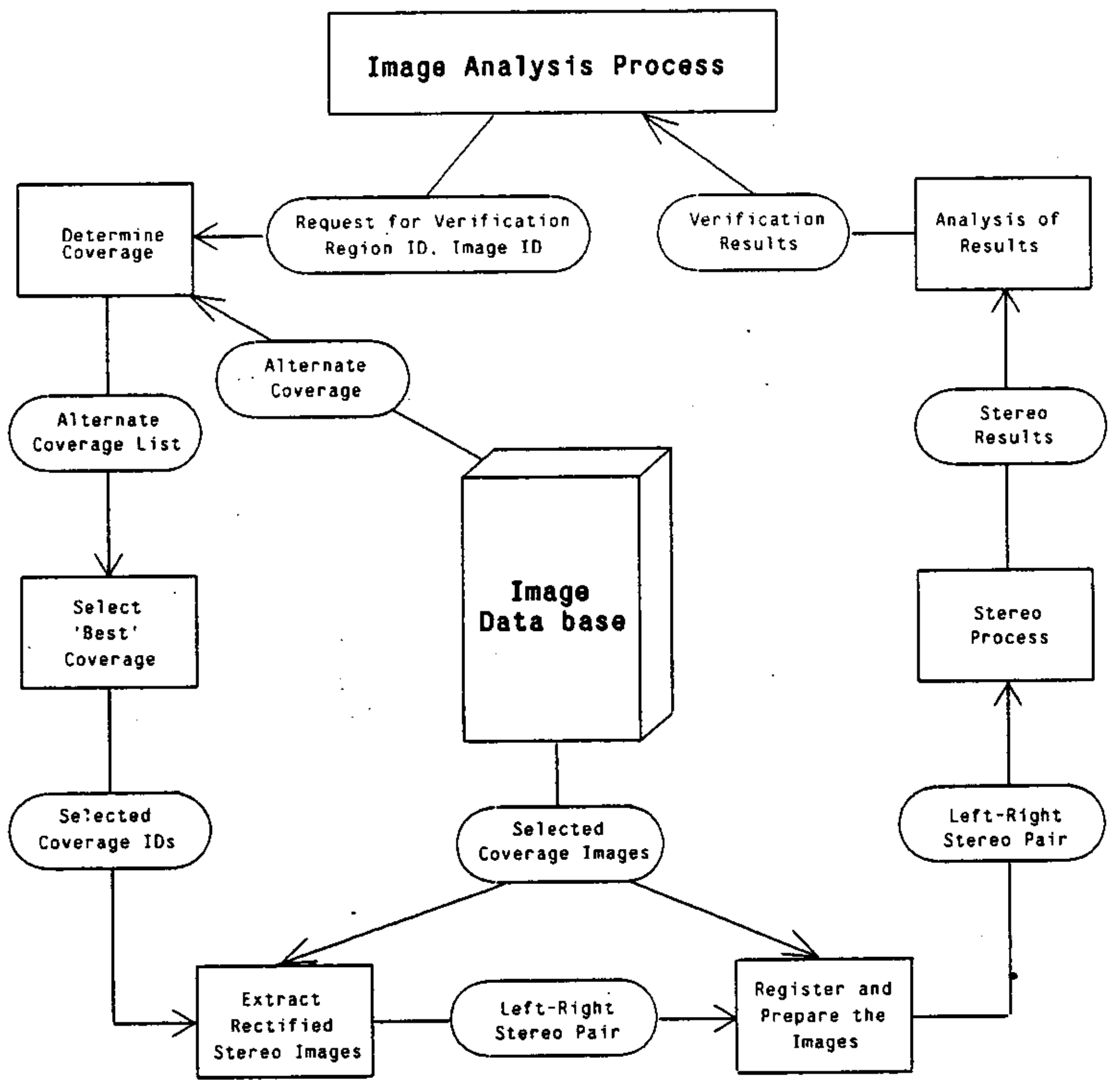


Figure 4-1: The Stereo Verification Process

interpretated by SPAM, and an image from the coverage file form the stereo pair.

The estimated height is used to select a *disparity range* that affects the contrast of the disparity image produced by the S1 algorithm. The resulting disparity image is quantized to 256 disparity levels. If this range is set too large, the disparity image will lack contrast and will be more difficult to analyze. If it is set too small, extremely large height disparities will occur outside the image range and will effectively be invisible. In other words, the initial disparity range determines the scaling of measured disparity into the disparity image. As in any linear scaling operation, one would like to utilize the full dynamic range of the output image while avoiding clipping at either end of the range. The selection of the disparity range constitutes the only external parameterization necessary in the implemented process. Our experience has shown that the disparity range need only be within a set of rather broad values to obtain useful results. For now, we use only three pre-selected ranges. Since SPAM actually selects the disparity range based on its region hypothesis, there is potential to add ranges to accommodate additional hypothesis types or to run the stereo process over a set of disparity ranges.

The following Sections will discuss these procedural steps in more detail and describe how STEREOSYS implements them. Some details are specific to the S1 matching algorithm used by STEREOSYS but are mentioned so that the reader may better understand our results.

4.1. Select Coverage

The MAPS database is used to produce an unsorted list of images, called a coverage file. Each image in the coverage file contains the region of interest. The interpretation image is used to create the Left stereo pair image since the S1 disparity image result overlays the Left image and, as will be seen, since there is no guarantee that the stereo image extracted from the alternate image will be properly registered to the region. The coverage file is used to select the database image from which the Right stereo pair image will be extracted. However, in most cases, the coverage file lists several images that contain the region in question. Several considerations enter into the choice of the best candidate. First, to minimize resampling extrapolation the candidate should be of the same or larger scale. Second, to reduce possible perspective distortion, the candidate should have the region of interest as near to its center as possible. In the case of vertical aerial photography this is the region's *nadir distance*. Third, if possible, the candidate should be from the same photographing mission, even flight line, as the original image to reduce temporal changes such as lighting, cloud cover, and ground movement. Figure 4-2 illustrates a pair of typical mapping aircraft flightlines that generate stereo coverage on successive frames of the same flightline as well as between adjacent flightlines. Figure 4-2 also illustrates that small changes in the aircraft platform position and direction can effect the actual area of overlap and must be accommodated; one cannot assume a constant direction or viewing position. This is discussed in Section 4.2.

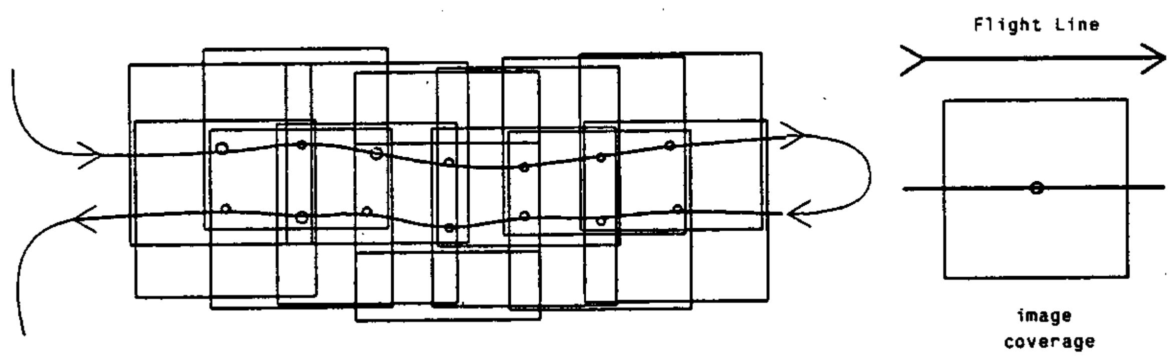


Figure 4-2: Mission Flight Lines

Other issues such as the source of the image, its recency, the processing and digitization history can enter into the selection of the images used to produce the stereo pair. For our purposes, STEREOSYS sorts the coverage file into a best stereo coverage order with respect to the hypothesis region's originating image as follows:

- Same Mission images (sorted by nadir distance)
- Same Scale images (sorted by nadir distance)
- All Other images (sorted by nadir distance)

The first image in the sorted coverage file best satisfies these criteria and is used to create the Right image.

4.2. Extract the Stereo Pair

The extraction of the stereo pair images is not a simple subimage cropping procedure. Like almost all stereo algorithms, S1 assumes image scanlines in the stereo pair are stereo epipolar lines. Without rotation this will not be the case with the selected Left and Right images. Photographic mission flight lines need not align with image digitization scanlines and, even if they did, sometimes the best coverage is found across mission flight lines or even from separate missions. For these reasons, a baseline orientation between the stereo pair is calculated so that the pair can be rotated to properly align the scanlines to meet the epipolar constraint.

However, this necessary rotation doesn't correct for distortions due to non-parallel camera axes. Even if the stereo process is sophisticated enough to account for large amounts of perspective distortions, chances are it will not be able to account for these distortions after they have been rotated. Therefore, the stereo pair Left-Right images are extracted through an orthographic rectification process before they are rotated. This method of subimage extraction removes perspective distortions by warping the subimage into a rectangular geographic box as well as establishing a common orientation for the image scanlines.

Several issues are considered in determining the size of the image area to be extracted. First, the area must contain the region's MBR plus a portion of the surrounding area since the S1 stereo results will only contain relative height information. In addition, the extracted area must be large enough so that the region of interest is contained in a rectangular sub-image cropped from the rotated image.

Specifically, to produce the necessary stereo pair, STEREOSYS extracts orthographically rectified areas identified as North-South oriented geographic rectangles by sub-pixel interpolation[13]. The corners of the extraction rectangle are calculated as a function of the region's centroid, the region's MBR, the Left-Right image scales, and the rotation necessary to make the extracted image East-West scanlines align with the baseline between the database coverage images.

4.3. Register the Stereo Pair

As mentioned in Section 3, S1 is capable of determining a global disparity or offset between stereo pairs. Using this S1 capability, the initially extracted stereo pair images are repeatedly processed by S1 to determine horizontal and vertical offset between the Left and Right images. With each pass over the image pair, S1 calculates a global offest value between the images. The process is repeated and the offset compounds until the offset stabilizes or begins to oscillate. Calculation of the registration offset is necessary because geodetic position correspondence control between images stored in the MAPS database is not sufficiently accurate to guarantee that the extracted Right image will overlay the Left image within the tolerances over which S1 can perform effective matching. As mentioned earlier, in many cases the initial registration errors may range from 5 to 30 pixels while the disparity shift is generally smaller than 10 pixels.

One can view the stereo matching process as first applying a coarse registration, followed by the actual calculation of disparity. It is interesting to note that the same technique, method of differences, appears to be effective for both global registration and local matching. A possible alternative to this registration step would be the addition of sufficient ground control to assure that images in the MAPS database could be registered within acceptable tolerances of 2 to 4 pixels. However, given that the ground sample distance for many of the images is approximately one meter, and that MAPS contains a wide variety of imagery with difference ground scales, projections, from multiple sources, it is unlikely that one would be able to totally eliminate the initial registration error.

The calculated registration offset is then used to extract the Right image for a second time. The orthographic extraction process is given a new geographic box that has been translated by the calculated offset. In this way, the new Right image will be more nearly registered to the Left image than if we had simply translated the original Right image. Originally we felt the offset could be handled entirely within the S1 stereo process and that resampling the Right image would be unnecessary. Experimentation showed this not to be the case, but since the internal offset capability was already added to S1, it is still in use. That is, even though we calculate an image offset and resample the Right image for a second time, we still later calculate any remaining offset between the Left and resampled Right images and use that value within the stereo process itself.

If necessary, the resulting Left-Right stereo pair images are rotated. The S1 stereo process assumes that scanlines are stereo epipolar lines. Until this point the stereo pair scanlines were East-West. Earlier a rotation value was calculated for use in determining the size of the extraction area. The rotation value is the amount the images must be rotated to make the epipolar lines become scanlines and assure that the Left-Right pair create a positive stereo image (ie. tall objects shift inward). The rotation value is the baseline orientation that was calculated earlier as the angle at the geographic center of the original image between East and the line to the alternate image geographic center. After rotation, the appropriate subimage rectangle of real data is cropped from the rotated image since the rotation leaves four right triangles of non-data at the corners.

4.4. Run the Stereo Process

At this point all constraints required by the S1 algorithm on the stereo pair have been met. The following few comments concern the specific use of the S1 process.

The Left-Right stereo pair images are repeatedly smoothed to form the coarse-fine hierarchy of images used by S1. As in Section 4.3, S1 again calculates a global registration offset value between the original Left image and the resampled Right image. This global offset is used internally by S1 during its calculation of the disparity image. The disparity image result is saved for analysis upon completion of the S1 disparity process.

4.5. Analyze the Results

In general the methods used in analyzing stereo results will depend on the stereo process used, the sensing method, and the type of disparity map produced by the process. Generally, one can characterize stereo matching results as one of the following:

- point correspondence(s)
- sparse depth map
- dense (complete) depth map

The objective of stereo verification is to determine if the region of interest has height relative to its surroundings. One of the major reasons for choosing S1 as our stereo process is that its dense disparity image simplifies this analysis step. Analysis of sparse feature based depth results like those produced by edge-based or interest area-based stereo processes would require careful determination of whether a feature belongs to the region of interest or to its surroundings. One obvious method would be to interpolate the sparse depth results into a dense map similar to the S1 disparity image. However it is not clear how reliable such a map would be, especially given the complex images presupposed in aerial interpretation, and techniques for doing such interpolation are still considered a topic for research[14]. The remainder of this Section describes how STEREOSYS analyzes S1 disparity images and is illustrated by several examples.

In order to analyze the dense S1 disparity image an overlaying bitmap of the region of interest is made. First the region's boundary point list is rectified to overlay the pre-rotated Left stereo pair image. The rectified boundary point list is then converted to a bitmap image of the region. Finally the bitmap is rotated to properly overlay the Left stereo pair image used in the disparity image calculation. The bitmap is used to distinguish the areas of the disparity image inside and outside the region of interest.

The disparity image and the overlaying region bitmap are used to calculate the mean and standard deviation for the disparity values of the areas within and without the region of interest. STEREOSYS uses a heuristic function that combines the standard deviations, S_{in} and S_{out}, and the difference in the means, D, to determine four confidence values:

1. Overall Confidence in the Stereo Results.
2. Confidence in the Region having Little to No Height.
3. Confidence in the Region having Moderate Height.
4. Confidence in the Region having Significant Height.

The first measure describes the overall confidence that can be placed on the stereo results. The disparity image results can vary from excellent to useless due to limits in correcting for misregistration and from noise caused by nondescript areas (Section 5.1). The confidence in the result is calculated as an empirically weighted sum of the mean difference and standard deviations.

$$0.1D + 0.5S_{in} + 0.4S_{out}$$

The D term is further influenced by the disparity image contrast which is related to the disparity range. A very small range can decrease this term by an empirical factor of 0.2. The S_{out} term is further influenced by an estimate of the amount of expected height clutter in the area. If the area is expected to be cluttered with tall objects this term increases by an empirical factor of 0.75. Both the disparity range and clutter values are provided by the processing context that caused SPAM to invoke STEREOSYS. These contexts include rules that recognize situations where height information can disambiguate competing hypotheses as well as supply likelihoods of clutter and height.

Confidence values (2-4) measure whether the region of interest was found to fall in one of three disparity or height ranges, provided by SPAM. These measures are relative to the hypothesized disparity range, rather than absolute statements about the regions height. For example, "Little to No Height" could mean about 5 meters high if a very large disparity range was selected but could mean less than one meter if a small range was used.

All three height confidence values are based on the difference in the means, D, but can be influenced by the disparity range in a manner similar to the D term in the results confidence described above. These values reflect where the height of the region falls within the height range supplied by SPAM. Confidence (2) is maximized as D goes to zero. Confidence (3) is maximized when D is approximately 1/7 of the full disparity range. Confidence (4) is maximized as D, the difference between the means, goes to maximum disparity. It should be remembered that very high objects can create disparity values beyond the range of maximum disparity in which case their extreme height would go unnoticed.

5. Experimental Results

This Section presents results produced by STEREOSYS that illustrate several of the important issues encountered during system development. We also amplify comments made in previous Sections concerning issues of registration, disparity estimates and automating the overall stereo process. Section 5.1 describes typical S1 results before minor revisions were made to the matching algorithm and STEREOSYS was implemented. Section 5.2 illustrates the problem caused by database image misregistration and results produced by STEREOSYS. Section 5.3 deals with stereo pair preparation processes as well as how the S1 results are analyzed. Section 5.4 describes test results from the automated use of STEREOSYS. Finally, Section 5.5 details a specific example from among the automated tests of Section 5.4.

Figure 5-1 shows one frame of aerial imagery containing National Airport in Washington, D.C. All of the examples in this paper come from various areas of this airport extracted from several stereo image pairs.

5.1. Preliminary S1 Experiments

Before trying to build a stereo verification system using S1, we experimented with the overall process in order to get a feel for how S1 might perform with MAPS images. Several issues arose: how to automatically set S1's initial disparity range values; deciding on modifications to provide the flexibility necessary to accommodate SPAM's requirements for a verification process; and how to analyze S1 results. These first experiments were performed on stereo pair images registered by hand and extracted from the database using

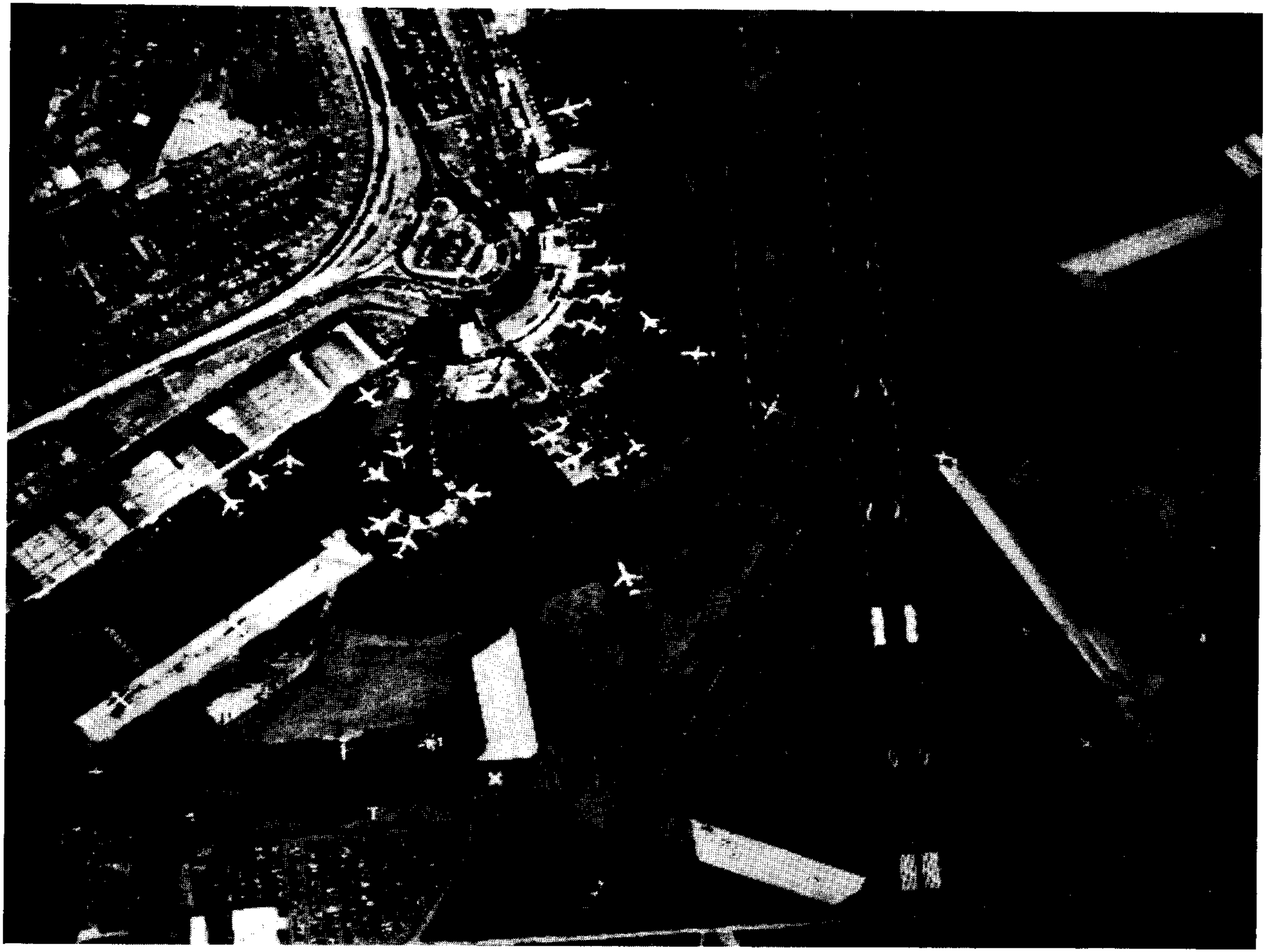


Figure 5-1: National Airport, Washington, D.C.

the same orthographic rectification process in STEREOSYS.

Figure 5-2 shows the Left and Right stereo images of a long hanger building running diagonally from the top right to bottom left of each image. Below the Left image is the S1 disparity image result. Within the disparity image dark areas are closer to the camera (higher) than are the light areas. The hanger is clearly shown to be higher than its surroundings. Some points of interest concerning the disparity image are:

- The speckled areas are caused by the loss of correspondence in large nondescript areas such as the large solidly shaded areas of pavement to the right and below the hanger. Such nondescript areas are characterized by the lack of edges or texture.
- Boundary edge effects show up as errors all around the disparity image. These effects are caused by lack of data outside the image and have been alleviated somewhat in the modified versions of S1.
- Stereo aliasing effects probably caused the problem with the curved hanger roof in the lower left corner. The white area in the roof indicates a concave section where none exists.
- Temporal changes in the stereo pair images can cause unpredictable results. An example is the white area along the right side caused by the moving truck.

Figure 5-3 shows a taxiway/runway area of the airport. This area contains very little variation in height and contains large variations in image intensity. The disparity image shows no significant height for any image region but again illustrates the problems with large nondescript areas and edge effects. Note also that the edge effects are propagated into nondescript areas. The statistical analysis method described in Section 4.5 was chosen partially because of its ability to recognize these situations as not being a significant indication of elevation.

5.2. Registration Problem and Solution

The results of the previous Section were produced from stereo pairs that were registered by hand. That is, the identification of the extraction areas was not done automatically and any misregistration in the stereo pair was kept to less than two pixels. This can be contrast with the 6 to 15 pixel disparities we normally experienced in the images used of

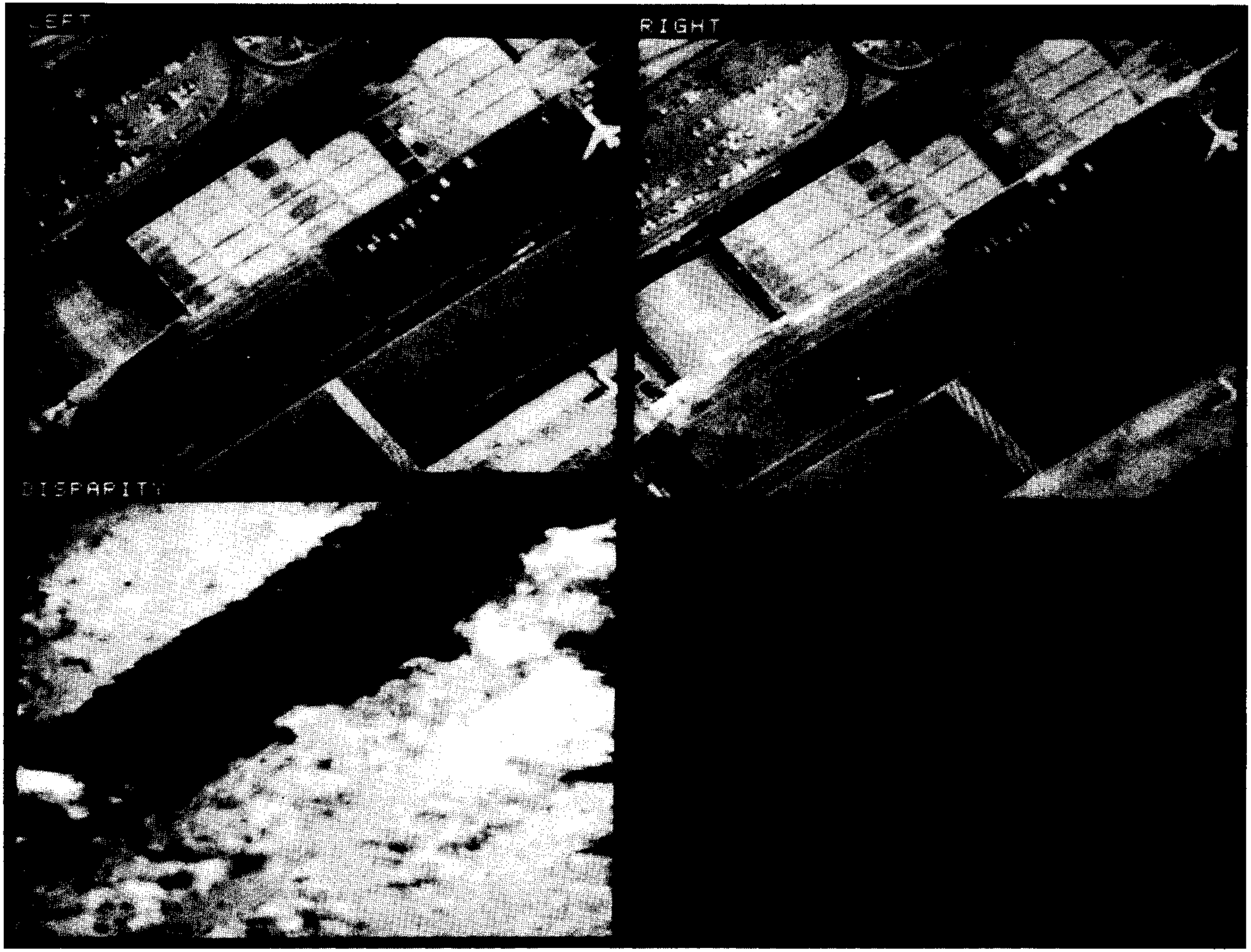


Figure 5-2: Early S1 Showing Height

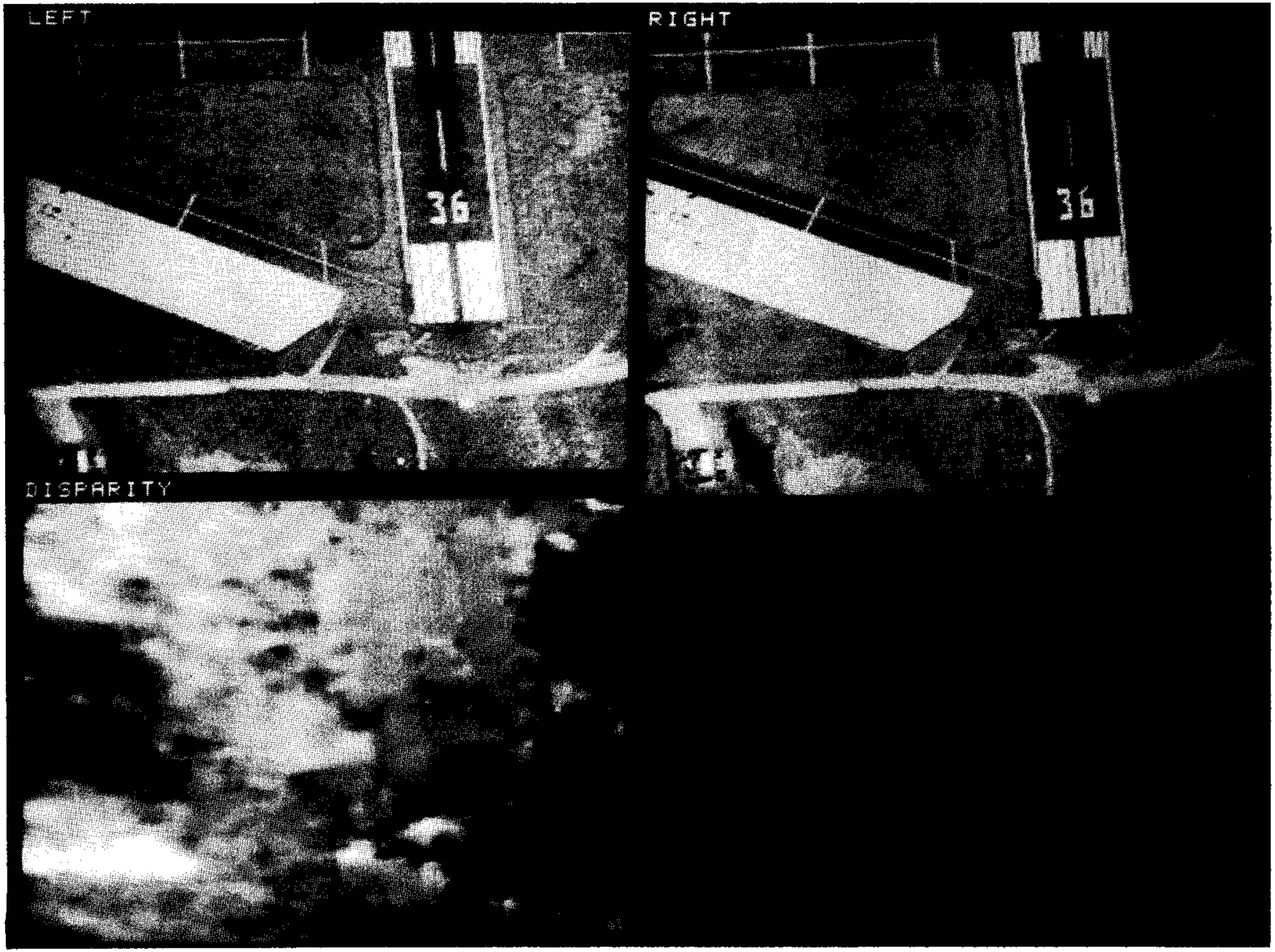


Figure 5-3: Early S1 Showing No Height

the Washington D.C. National Airport.

Through experimentation it was found that the S1 process could sometimes produce fair results if the stereo pair was up to 6 pixels misregistered, but this was found to be far too restrictive for automatic purposes since the MAPS correspondence between database images can be off by as much as 30 pixels or more in areas with little ground control. Figure 5-4 shows early S1 results from a stereo pair created automatically from the database.

The misregistration problem is handled by S1's ability to calculate a global disparity shift between pairs of images. STEREOSYS uses S1 to calculate the shift between the originally extracted stereo pair then uses the shift value to re-extract the Right image. Figure 5-5 illustrates this process. The top two images are the original Left-Right stereo pair. The lower right contains the Right image after a calculated shift of 7 pixels vertical and 13 pixels horizontal has been eliminated.

Since the shift is an inexact statistical value S1 was also modified to calculate and compensate for any remaining small misregistrations. The lower left of Figure 5-5 contains the disparity image that results from the combination of these techniques. This approach has demonstrated the ability to properly compensate for original misregistrations of up to 25 pixels. Beyond that point the global shift calculation normally fails. However this shortfall can be properly overcome by adding enough control to the image database to assure misregistrations will not exceed the limits of the registration process.

5.3. Analysis of Results

If the reader looked carefully at the stereo pair used in the last Section she might have noticed that the pair forms a negative stereo image. That is, objects with height lean away from one another and, if viewed in stereo, would form a reversed stereo image. In such an image buildings would appear to go down into the ground. To correct this, either

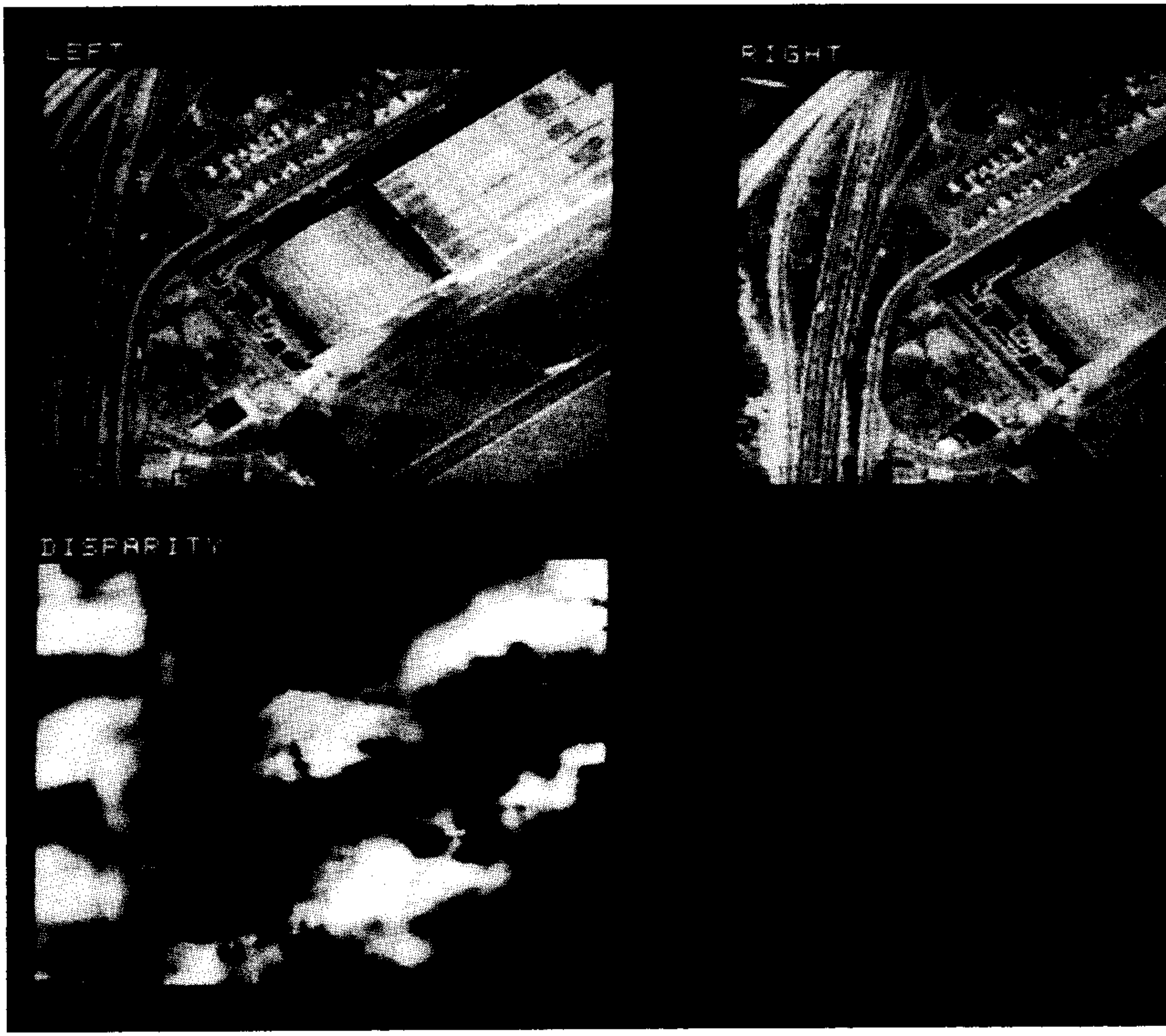


Figure 5-4: Early S1 Against Poor Registration

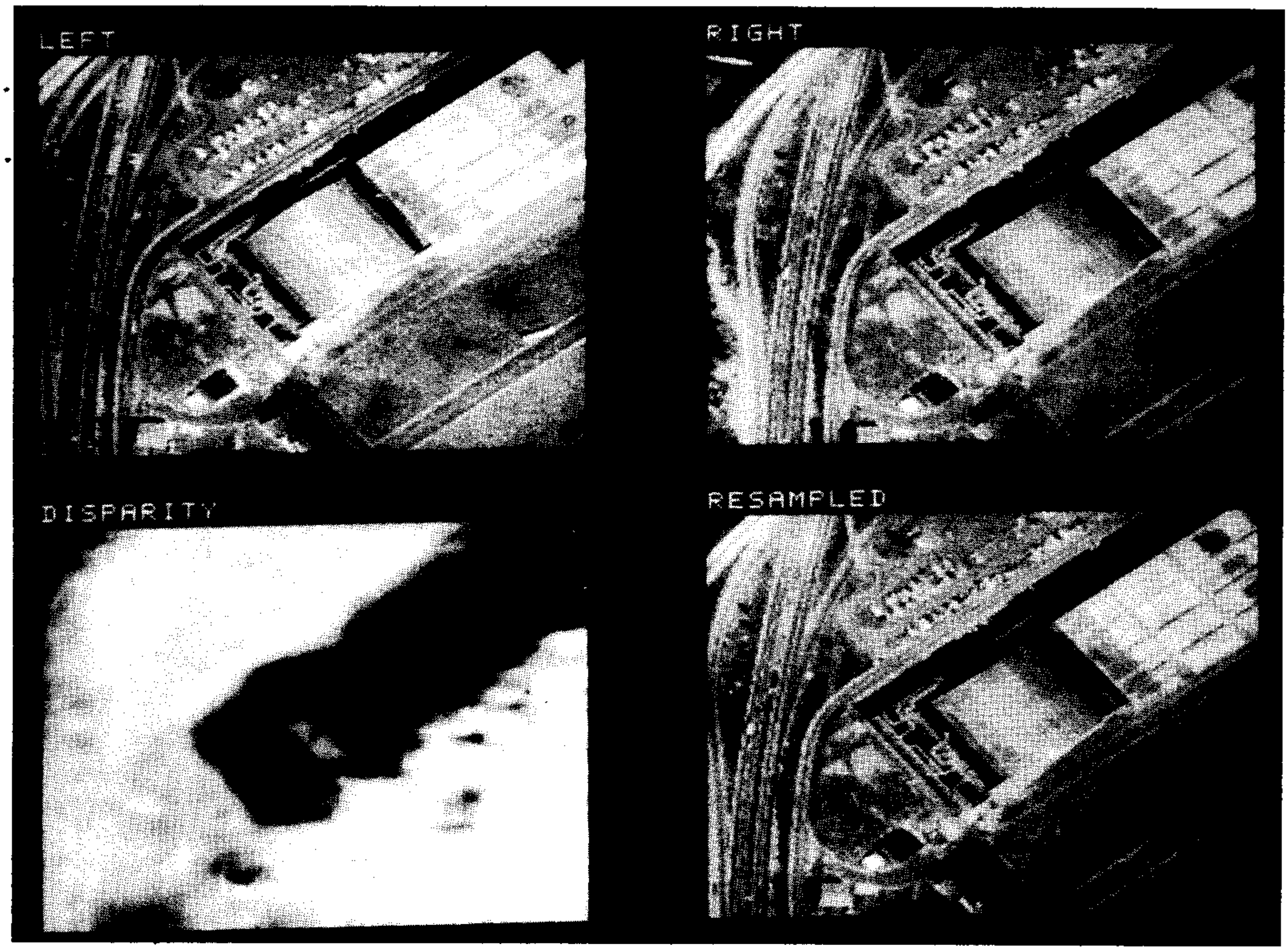


Figure 5-5: Misregistration Solution

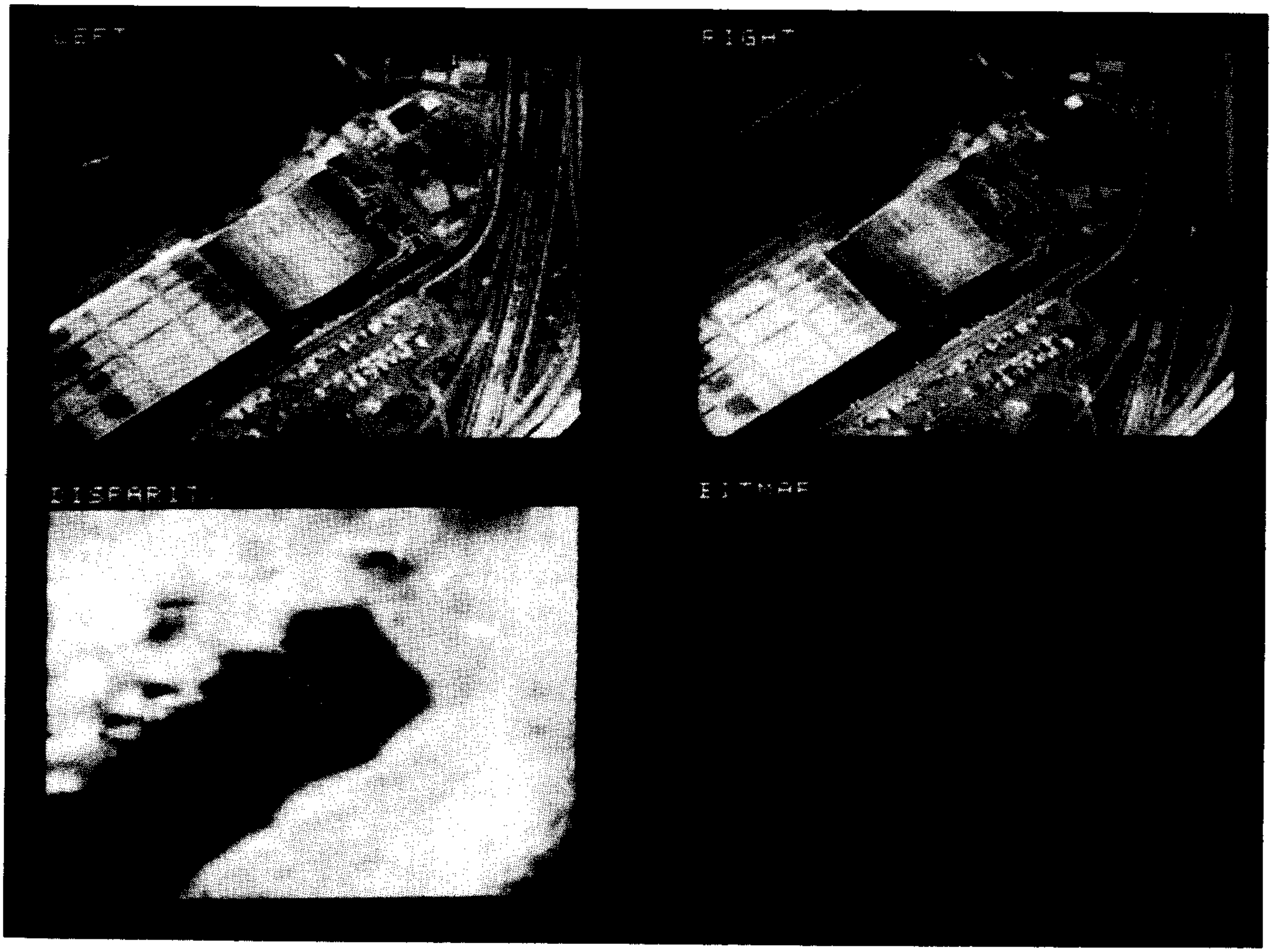
the Left and Right images could be exchanged or both could be rotated 180 degrees. We choose to rotate the stereo pair images since this can be combined with the arbitrary rotations necessary to align scanlines and epipolar lines. The results in Figure 5-6 show the rotated stereo pair of Section 5.2. Note that the disparity images shown in Section 5.2 were produced after exchanging the images to from a positive stereo pair or else the disparity image would have shown negative height for the buildings.

In order to analyze the S1 disparity images an overlaying bitmap of the region of interest is produced as described in Section 4.5. The bitmap is used to distinguish areas of the disparity image as being either inside or outside the region of interest. Based on this separation, the mean and standard deviation of disparity values within and without the region are calculated. STEREOSYS uses the standard deviations and the difference in the means in its heuristics that determine the stereo verification confidence values also described in Section 4.5. These values reflect confidence in the stereo result and confidence in the region of interest having little height, moderate height or significant height. The values are such that 0.0 signifies no confidence while 1.0 signifies "perfect" confidence. An example of the bitmap and the confidence values are also shown in Figure 5-6 .

Figure 5-7 is an example of STEREOSYS results where the region of interest has no height.

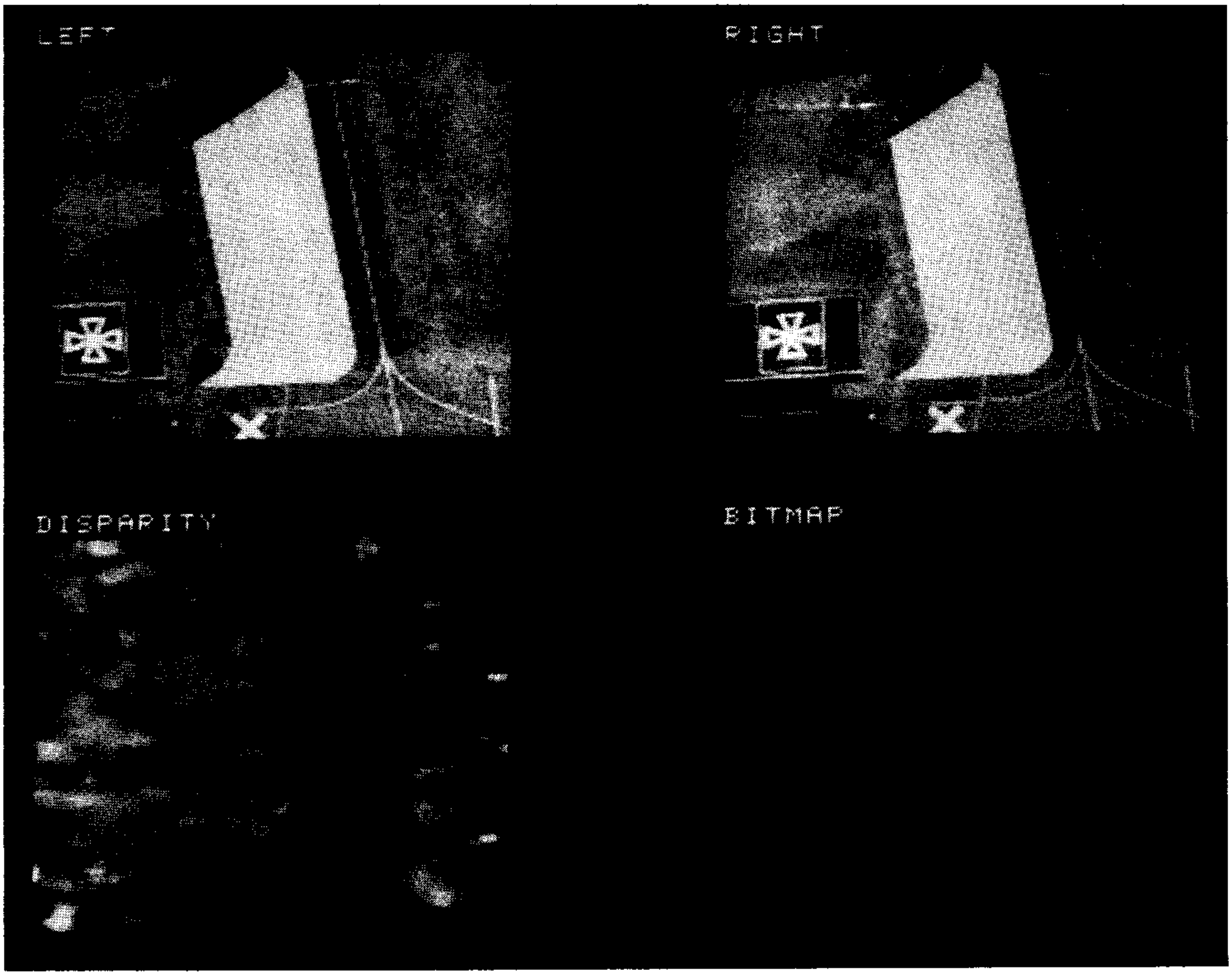
5.4. Fully Automatic Use

One important objective for STEREOSYS was that it be flexible enough to work reliably with all sorts of regions and in concert with SPAM. To test STEREOSYS against these goals, SPAM was given access to STEREOSYS for the purpose of stereo verification while trying to interpret the Washington D.C. National Airport area. STEREOSYS was called upon to give a verification analysis of 70 regions. Table 5-1 lists the confidence results for these regions of interest. The *Human Interpretation* column gives the correct interpretation for each region. The *SPAM Hypothesis* column gives the SPAM hypothesis used in invoking STEREOSYS. Exact interpretation of the *Low*, *Med* and *High* columns depends on what



Result Confidence	0.651
Low Height Confidence	0.051
Moderate Height	0.159
Significant Height	0.791

Figure 5-6: Analysis Results Indicating Height



Result Confidence	0.994
Low Height Confidence	0.963
Moderate Height	0.039
Significant Height	0.000

Figure 5-7: Analysis Results Indicating No Height

hypothesis SPAM had for the region when it invoked STEREOSYS. For example, if the hypothesis was for a low object like tarmac then *Low* would indicate a range in heights of 0-1 meters; *Med* 1-5 meters; and *High* 5-infinity. But, if the hypothesis was for a moderately tall object such as a hanger then *Low* would indicate a range of 0-5 meters; *Med* 5-12 meters; and *High* 12-infinity. A similar broadening of ranges would hold for very tall hypothesis but in this case of airport analysis, such hypothesis are not used.

Close examination of Table 5-1 reveals that as the *Result* confidence decreases height confidences tends to move toward *Low*. This is because disparity images with low *Result* confidences are random noisy messes which cause the mean values for the areas within and without the region of interest to become nearly equal. The heuristics calculating height confidences rely mostly on the difference in these means; no difference indicates no height. The very few cases where poor results cause confidence in the region being tall happen when the region of interest is very small and happens to lie on a random dark area of the disparity image.

Table 5-2 summarizes the test by categorizing result confidence values. This data primarily reflects how often the system was able to properly register the stereo pair. Result confidences of over 0.6 (out of 1.0) reflect good registration. Result confidences below 0.4 reflect cases where the system was probably unable to determine the shifts necessary to bring the stereo pair into registration. Values 0.4 - 0.6 can be caused by areas cluttered with high objects, highly nondescript areas or registration problems. Poor results due to bad registration can be alleviated through the addition of correspondence control between data base images. The remaining problems like nondescript areas and moving objects are inherent in the stereo process itself and are not dealt with in this work. Table 5-2 also summarizes how well the confidence results agreed with human height evaluation for the regions being verified. For the purposes of this evaluation a "winner take all" strategy is used. That is, the height confidence range having the highest confidence was deemed to be the height assigned to the region by STEREOSYS.

Region ID	Human Interpretation	SPAM Hypothesis	Confidences: Result	Low	Med	High
R01	Runway	Runway	0.86	0.98	0.02	0.00
R02	Runway	Runway	0.42	0.98	0.02	0.00
R03	Runway	Runway	0.33	0.98	0.02	0.00
R04	Parking-apron	Hanger	0.34	0.98	0.02	0.00
R05	Grassy-area	Parking-apron	0.28	0.42	0.48	0.10
R06	Runway	Runway	0.44	0.21	0.60	0.19
R07	Runway	Runway	0.40	0.82	0.16	0.02
R08	Taxiway	Parking-apron	0.35	0.30	0.57	0.13
R09	Taxiway	Hanger	0.56	0.13	0.57	0.30
R10	Taxiway	Hanger	0.29	0.98	0.02	0.00
R12	Taxiway	Parking-apron	0.69	0.70	0.25	0.05
R13	Taxiway	Hanger	0.58	0.51	0.41	0.08
R14	Taxiway	Hanger	0.23	0.75	0.22	0.03
R15	Taxiway	Hanger	0.30	0.98	0.02	0.00
R16	Taxiway	Hanger	0.37	0.98	0.02	0.00
R17	Taxiway	Hanger	0.78	0.22	0.60	0.18
R18	Taxiway	Parking-lot	0.32	0.98	0.02	0.00
R19	Taxiway	Terminal	0.27	0.98	0.02	0.00
R20	Grassy-area	Hanger	0.32	0.67	0.28	0.05
R22	Grassy-area	Parking-apron	0.46	0.43	0.48	0.09
R23	Grassy-area	Hanger	0.21	0.19	0.60	0.21
R25	Grassy-area	Hanger	0.31	0.98	0.02	0.00
R26	Grassy-area	Parking-apron	0.40	0.98	0.02	0.00
R27	Grassy-area	Hanger	0.94	0.55	0.38	0.07
R28	Grassy-area	Parking-apron	0.40	0.50	0.42	0.08
R29	Grassy-area	Hanger	0.30	0.27	0.58	0.15
R30	Grassy-area	Hanger	0.67	0.98	0.02	0.00
R32	Grassy-area	Hanger	0.73	0.98	0.02	0.00
R35	Grassy-area	Parking-apron	0.47	0.98	0.02	0.00
R36	Grassy-area	Hanger	0.57	0.98	0.02	0.00
R37	Grassy-area	Hanger	0.66	0.57	0.36	0.07
R38	Grassy-area	Parking-apron	0.26	0.39	0.51	0.10
R40	Grassy-area	Hanger	0.62	0.89	0.10	0.01
R41	Parking-lot	Parking-apron	0.20	0.98	0.02	0.00
R43	Parking-lot	Runway	0.48	0.98	0.02	0.00
R44	Parking-lot	Hanger	0.86	0.98	0.02	0.00
R45	Parking-lot	Parking-apron	0.48	0.36	0.53	0.11
R46	Parking-lot	Parking-apron	0.75	0.16	0.59	0.25
R47	Parking-lot	Parking-apron	0.29	0.98	0.02	0.00
R48	Parking-lot	Parking-apron	0.26	0.24	0.59	0.17
R49	Parking-lot	Hanger	0.61	0.50	0.42	0.08
R50	Grassy-area	Grassy-area	0.21	0.98	0.02	0.00
R51	Hanger	Hanger	0.65	0.05	0.16	0.79
R52	Terminal	Hanger	0.42	0.05	0.16	0.79
R53	Hanger	Hanger	0.99	0.09	0.47	0.44
R54	Grassy-area	Parking-apron	0.75	0.60	0.34	0.06
R55	Taxiway	Hanger	0.22	0.57	0.37	0.06
R56	Taxiway	Terminal	0.30	0.41	0.49	0.10
R57	Hanger	Parking-apron	0.58	0.05	0.19	0.76
R58	Hanger	Hanger	0.25	0.74	0.23	0.03
R62	Parking-lot	Hanger	0.85	0.98	0.02	0.00
R64	Grassy-area	Parking-apron	0.40	0.12	0.55	0.32
R65	Grassy-area	Parking-apron	0.39	0.51	0.41	0.08
R66	Grassy-area	Grassy-area	0.36	0.27	0.58	0.15
R67	Grassy-area	Parking-apron	0.49	0.98	0.02	0.00
R68	Grassy-area	Parking-area	0.39	0.65	0.29	0.06
R73	Terminal	Hanger	0.82	0.10	0.52	0.38
R75	Terminal	Hanger	0.65	0.09	0.50	0.41

Table 5-1: Test Results

Region ID	Human Interpretation	SPAM Hypothesis	Confidences: Result	Low	Med	High
R77	Terminal	Hanger	0.60	0.07	0.39	0.54
R78	Tarmac	Parking-apron	0.69	0.98	0.02	0.00
R79	Parking-apron	Parking-apron	0.43	0.95	0.05	0.00
R80	Parking-apron	Parking-apron	0.36	0.54	0.38	0.08
R81	Tarmac	Parking-apron	0.24	0.98	0.02	0.00
R82	Parking-apron	Grassy-area	0.50	0.98	0.02	0.00
R83	Tarmac	Parking-apron	0.55	0.20	0.59	0.21
R84	Parking-apron	Hanger	0.39	0.98	0.02	0.00
R85	Tarmac	Parking-apron	0.17	0.91	0.08	0.01
R87	Road	Taxiway	0.29	0.07	0.36	0.57
R94	Road	Road	0.48	0.14	0.57	0.29
R95	Road	Runway	0.27	0.98	0.02	0.00

Table 5-1: Test Results, continued

Confidence	% of Tests	% Human Agreement
0.6-1.00	27.1	89.5
0.5-.599	8.6	66.7
0.4-.499	20.0	64.3
0.0-.399	44.3	67.8

Table 5-2: Test Summary

The careful reader will notice that the "% Human Agreement" value in Table 4-2 does not decrease when the "Confidence" value is below 0.4. This is because disparity image results for regions with such low confidence are usually randomly noisy. Statistically, this causes the difference in mean disparity values between the areas inside and outside the region to approach zero which in turn causes the STEREOSYS height confidence heuristic to favor a low height interpretation. Of the 70 test cases presented, only eight (8) are of regions with real height and of these, only one created a result confidence below 0.4. Since the remaining poor result confidences were caused by regions without height, this somewhat inflates the percent of human agreement within the low confidence range.

Tables 5-3 through 5-6 are confusion matrices showing the performance of STEREOSYS over several result confidence ranges. The table columns are the number of times a SPAM hypothesis was correct or incorrect, with respect to height, as compared to a human interpretation. For example, if SPAM hypothesized a low height region, say grassy area, and the region was any other low height region, say tarmac, then the hypothesis was deemed correct. The table rows indicate the number of times STEREOSYS confirmed or rejected the SPAM hypothesis. A perfect result would find STEREOSYS always confirming correct hypotheses and rejecting the incorrect hypotheses, ie., zeros in the lower left and upper right elements of the confusion matrix.

Hypothesis	Correct	Incorrect
Confirmed	31	5
Rejected	14	20

Table 5-3: All Result Confidences [0.0 - 1.0]

Hypothesis	Correct	Incorrect
Confirmed	19	1
Rejected	7	12

Table 5-4: Result Confidences [0.4 - 1.0]

Hypothesis	Correct	Incorrect
Confirmed	10	1
Rejected	2	12

Table 5-5: Result Confidences [0.5 - 1.0]

Hypothesis	Correct	Incorrect
Confirmed	9	0
Rejected	1	9

Table 5-6: Result Confidences [0.6 - 1.0]

Table 5-7 indicates that STEREOSYS performs well with objects having height. One initial concern with the S1 stereo process is that often, when it is initiated with too small a disparity range, the S1 method will not converge to a useful result. This could be the case when SPAM hypothesizes an object with no height and in reality the object has significant height. To lessen the chance of this problem occurring we tried to be generous in the size of our three standard height ranges. In the one case where this situation actually occurred, STEREOSYS produced the correct response.

Hypothesis	Correct	Incorrect
Confirmed	6	0
Rejected	0	1

Table 5-7: Regions with Actual Height [0.4 - 1.0]

5.5. A Detailed Example

As an example of how stereo verification can aid image analysis, this Section describes one of the 70 invocations of STEREOSYS by SPAM from Section 5.4. This Section is included also to give the reader a flavor for how SPAM, a rule-based production system, utilizes stereo verification. As mentioned in the introduction, SPAM may invoke STEROSYS in one of two modes. During early stages of scene interpretation SPAM gathers low-level information by testing newly generated regions for height in order to develop an initial set of hypotheses for the region. During later processing, as collections of regions begin to be combined into components of the airport model, STEREOSYS is employed to disambiguate between two or more plausible but conflicting hypotheses. This Section describes the former situation by showing extracts from the SPAM and STEREOSYS execution traces. These extracts have been edited slightly to enhance their readability.

Figure 5-8 contains several of the SPAM OPS5[15] rules that lead to the invocation of STEREOSYS. The first rule, region-to-fragment::get-depth, is used to recognize an appropriate point for the invocation of STEREOSYS. The firing of this rule causes SPAM to change its operating *context* to the *generate-depth-info task*. The next two rules are examples of rules activated by this context. They will set up the STEREOSYS parameters appropriate for the region of interest's current best hypothesis- based on an assigned confidence value. These parameters are an indication of the expected height range and height clutter for the region of interest. The rules for setting STEREOSYS parameters appropriate for a runway or a hanger hypothesis are shown. The rule applicable to the hypothesis with the highest confidence value will fire. Along with setting the necessary parameters, the rule firing will change the context to *get-depth* in order to fire the next rule, specific::get-region-depth, which actually invokes the STEREOSYS process.

Figure 5-9 is an excerpt from the SPAM trace just before it invoked STEREOSYS. The region of interest is Hand36809-N.37—0 ("R37" for short). Rule firings 853 to 856 step through the development of hypothesis confidences for region R37. By the end of this sequence of rules the region had a 0.94 confidence of being a hanger and a 0.68 confidence of being a grassy area. These interpretations were based on weak heuristics and measurements such as 2D shape, texture, ect. Rule firing 857 changed the operation context to the get-depth task because the hangar hypothesis had the highest confidence of any interpretation for this region. This caused the STEREOSYS parameters for height-range and clutter to be set appropriately during rule firing 858. Finally, using parameters best set for finding height information about typical hangers, STEREOSYS was invoked.

Figure 5-10 gives extracts from the STEREOSYS trace of region R37. An explanation of the trace, coded by the bold capital letters, follows:

- **A**: The input parameters are listed. The parameter *Height-range* determines what disparity range S1 will use. In this case it was set to *10-inf* because SPAM thought R37 might be a hanger which is usually 10 or more meters high.
- **B**: The database is used to find the boundary list file and centroid for R37.
- **C**: Again the database is used to create an unsorted ".ec" coverage file of images containing R37.
- **D**: The coverage file is sorted; the best images are selected; and the extraction regions are calculated. Notice that since the rotation value is so near zero the later rotation steps are skipped over and replaced by simple UNIX moves (mv).
- **E**: The **Ortho** process extracts an orthographically rectified stereo pair.
- **F**: S1 is invoked to calculate any misregistration between the pair. The calculated offset is not shown in the trace, but for this particular region it was 13 pixels vertically and 14 pixels horizontally.
- **G**: **Ortho** is called to extract a new Right image based on the calculated offset.
- **H**: S1 is again invoked to calculate any remaining offset and produce a disparity image.
- **I**: The boundary list for R37 is warped, or rectified, to overlay the Left image extraction area. The result is a ".seg" file which is converted to a bitmap.

```
(p region-to-fragment::get-depth
    { <context> (context ^task region-to-fragment ^datum <token>) }
    (region ^token <token> ^region-status interpreted)
  - (context ^task <> region-to-fragment)
  - (store-results ^result-one class-match)
  - (store-results ^result-one subclass-match)
  -->
    (remove <context>)
    (make context ^task generate-depth-info ^datum <token>))


(p depth::get-runway-depth
    { (context ^task generate-depth-info ^datum <token>) <context> }
    (fragment ^region-token <token> ^hypothesis runway ^confidence <c>)
  - (fragment ^region-token <token> ^hypothesis <> unknown
        ^confidence > <c>)
  - (context ^task <> generate-depth-info)
  -->
    (remove <context>)
    (bind <height-estimate> 0-5)
    (bind <cluttering> isolated)
    (make context ^task get-depth
          ^datum <token> <height-estimate> <cluttering> runway))


(p depth::get-hangar-depth
    { (context ^task generate-depth-info ^datum <token>) <context> }
    (fragment ^region-token <token> ^hypothesis hangar ^confidence <c>)
  - (fragment ^region-token <token> ^hypothesis <> unknown
        ^confidence > <c>)
  - (context ^task <> generate-depth-info)
  -->
    (remove <context>)
    (bind <height-estimate> 10-inf)
    (bind <cluttering> cluttered)
    (make context ^task get-depth
          ^datum <token> <height-estimate> <cluttering> hangar))


(p specific::get-region-depth
    { (context ^task get-depth
          ^datum <regtok> <height-est> <regcontext> <hyp>) <context> }
    (global-status ^current-image <img>)
    (interp-constants
        ^output-file <outfile>)
    { <region> (region
          ^token <regtok>
          ^symbolic-name <symname>) }
-->
    (call depth <symname> <img> AREAL <height-est>
          <regcontext> -o <outfile>)
    (remove <context>)
    (modify <region> ^depth-low <lowdepth>
          ^depth-moderate <moddepth> ^depth-high <highdepth>))
```

Figure 5-8: OPS5 Production Rules

```
853. region-to-fragment::generate-subclass-match 2939 2912 2938
      CONFID-LIST: (1.0 0.909238 0.0 0.85332)
Match of hangar for region Hand36809-N.37_0 = 0.940639
854. interpret-as-hangar 2946 12 2912 2935 2938
      Interpreting region Hand36809-N.37_0
855. region-to-fragment::generate-subclass-match 2923 2912 2976
      CONFID-LIST: (1.0 0.918012 0.124357)
Match of grassy-area for region Hand36809-N.37_0 = 0.68079
856. interpret-as-grassy-area 2978 12 2912 2974 2976
      Interpreting region Hand36809-N.37_0
857. region-to-fragment::get-depth 2912 2984
858. depth::get-hangar-depth 2986 2984
859. specific::get-region-depth 2988 2982 2 2984
```

Figure 5-9: SPAM Execution

- J: The disparity image is analyzed; the statistics are converted into confidence values and the results are sent back to SPAM.

These particular results are interesting in that SPAM sent R37 to STEREOSYS with a current hypothesis that R37 was a hanger but got back a fairly confident indication that there was no appreciable height present. That is, the result confidence was 0.66 with a 0.57 confidence that R37 had little to no height. Figure 5-11 shows R37's originally extracted stereo pair, the disparity result and region bitmap.

6. Conclusions

We believe that using height information in verification of aerial image analysis is an important approach and that the general stereo verification steps of Section 4 are minimal and applicable to all image analysis supported by an image database. In this context, our work with STEREOSYS has explored the pertinent issues and found viable solutions to the following important questions:

1. How can an aerial image database automatically generate a useful stereo pair containing an arbitrary region?
2. How can a stereo system handle the misregistration problems inherent in variable sourced image databases?
3. What kind of stereo results are appropriate for use in a verification process?
4. How can stereo results be analyzed so as to reflect not only the presence (or absence) of height but also the inherent reliability of the results?

```
A:  STEREOSYS:  Region_id    = Hand36809-N.37_0
                Generic      = dc36809
                Region_type  = AREAL
                Height_range = 10-inf
                Clutter      = cluttered

    STEREOSYS:  Region Temp file key is R09-37
B:    D3_INFO:  D3_file = /visf/airport/approxfeas/Hand36809-N/37A0.d3
                    X = 139841.616519   Y = 277348.279196
C:  STEREOSYS:  d3entcor /visf/airport/approxfeas/Hand36809-N/37A0.d3 R09-37.ec
D:  STEREOSYS:  sorting EC file by stereo coverage
    STEREOSYS:  selecting best coverage:
         Left       = dc36809     ScaleL = 0.000083
         Right      = dc36808     ScaleR = 0.000083
         HalfHeight = 82   HalfWidth = 44
         Rotation = 0.782102
         Del_lat  = 0.040685  Del_lon  = 0.053419

E:  STEREOSYS:  ortho dc36809 38 50 34 902 77 2 23 472 38 50 48 330 77 2 33 85
                                -m 1.056762 R09-37.l.tmp
    Mapping lbw image of dc36809 to the box formed by:
          lat N38 50 34 (902)    lon W77 2 23 (472)
    and   lat N38 50 48 (330)    lon W77 2 33 (85)
    Requested gridsize:  1.06 meters
    Actual gridsize:  0.0407 X 0.0525 seconds (1.05 X 1.06 meters)
    Size of result: 60390 bytes (330 rows X 183 columns)

    STEREOSYS:  ortho dc36808 38 50 34 902 77 2 23 472 38 50 48 330 77 2 33 85
                                -m 1.056762 R09-37.r.pre

F:  STEREOSYS:  Off Set cmd file created = b25825.off.tmp
    STEREOSYS:  s1 b25825.off.tmp

G:  STEREOSYS:  ortho dc36808 38 50 34 374 77 2 22 741 38 50 47 802 77 2 32 354
                                -m 1.056762 R09-37.r.tmp

    STEREOSYS:  mv R09-37.r.tmp R09-37.right
    STEREOSYS:  mv R09-37.l.tmp R09-37.left

H:  STEREOSYS:  S1 command file created = c25825.cmd.tmp
    STEREOSYS:  s1 c25825.cmd.tmp

I:  STEREOSYS:  Created warped SEG file  R09-37.w.seg
    STEREOSYS:  segtoimg R09-37.w.seg -o R09-37.b.tmp -m -s -I R09-37.left

    STEREOSYS:  mv R09-37.b.tmp R09-37.bitmap

J:  STEREOSYS:  Stereo Statistics for Hand36809-N.37_0

            Mean difference:  5.41722
            Region stddev:  26.9966
            Backgnd stddev:  34.3314

            Result confidence:  0.661235
            Low depth confidence:  0.573924
            Moderate depth confidence:  0.360413
            High depth confidence:  0.065663
```

Figure 5-10: STEREOSYS Execution

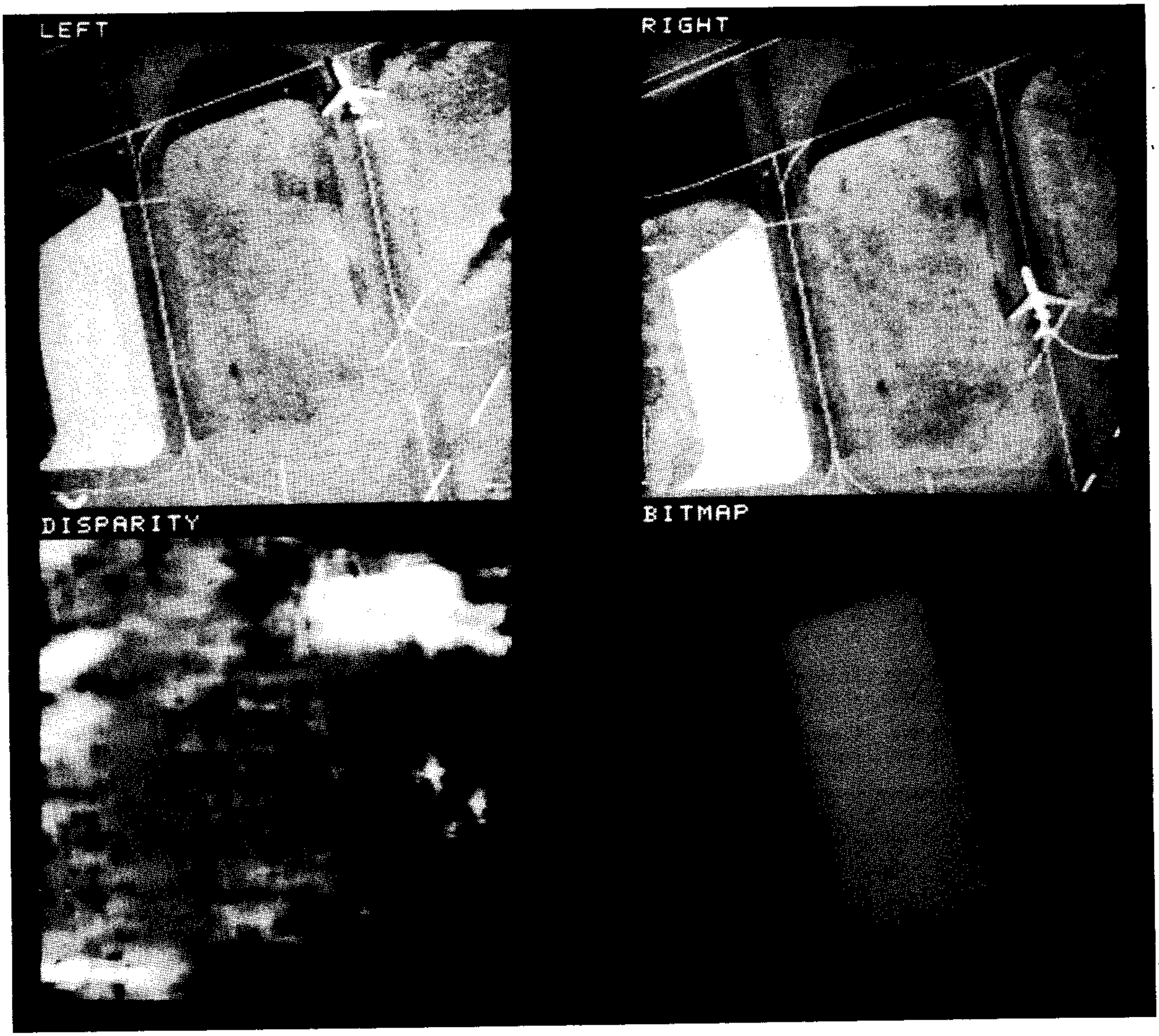


Figure 5-11: Region R37

STEREOSYS is not an infallible stereo verification system as indicated by the experimental results presented in Section 5. However, STEREOSYS is a highly flexible system that accomplishes the *entire* stereo process *automatically* from selecting image coverage to analyzing the stereo results while using an image database that has less than perfect image correspondence capabilities. From this viewpoint, we feel STEREOSYS has demonstrated the potential use of stereo verification in aerial image analysis.

If one defines stereo verification, as we do, to be a process whose purpose is to give a simple indication of the depth of one region in an image relative to the rest of the image, then stereo verification can be seen to be applicable to any domain where the identification of regions with significant differences in depth is important. For example, stereo verification could be useful for collision avoidance in mobile robotics or for the initial locating of tall objects in aerial photographs. This is especially true if an emphasis is placed on the use of fast and flexible processes. STEREOSYS has shown itself to be flexible but lacking in speed primarily due to the necessity for subimage rectification during the extraction of the stereo pair images. The registration step, needed to determine the offset in the originally extracted pair, is also time consuming*. However, we believe stereo verification can be done far more efficiently and, if so, can greatly benefit aerial analysis and other domains.

Many different approaches to performing passive photographic stereo have been studied[16] and several have been implemented but few have been incorporated into systems that accomplish anything useful beyond producing pretty results if given a tightly controlled stereo pair. Flexible stereo verification is a useful application of stereo processes. Our work has outlined the general process of stereo verification and has studied how one stereo process, S1, can do useful verification work. We believe that one immediate direction of study in stereo verification should be in the testing of other known stereo

---

*Approximate time for each experiment is about 20 cpu minutes using a VAX 11/780 under UNIX processes in stereo verification systems.

7. Acknowledgements

Wilson Harvey integrated STEREOSYS into the SPAM rule based system and enabled us to run experiments using realistic system hypotheses and machine generated segmentations. Steven Shafer and John McDermott commented extensively on an earlier draft of this report.

8. References

1. McKeown, D.M., Harvey, W.A. and McDermott, J., "Rule Based Interpretation of Aerial Imagery", *Proceedings of IEEE Workshop on Principles of Knowledge-Based Systems*, Dec 1984, pp. 145-157.

2. McKeown, D.M. and Pane, J. F., "Alignment and Connection of Fragmented Linear Features in Aerial Imagery", *Proceedings IEEE Computer Vision and Pattern Recognition Conference*, June 1985, Also available as Technical Report CMU-CS-85-122

3. McKeown, D.M., Denlinger, J.L., "Map-Guided Feature Extraction from Aerial Imagery", *Proceedings of Second IEEE Computer Society Workshop on Computer Vision: Representation and Control*, May 1984, Also available as Technical Report CMU-CS-84-117

4. McKeown, D.M., "MAPS: The Organization of a Spatial Database System Using Imagery, Terrain, and Map Data", *Proceedings: DARPA Image Understanding Workshop*, June 1983, pp. 105-127, Also available as Technical Report CMU-CS-83-136

5. McKeown, D.M.,, "Digital Cartography and Photo Interpretation from a Database Viewpoint", in *New Applications of Databases*, Gargarin, G. and Golembe, E., ed., Academic Press, New York, N. Y., 1984, pp. 19-42.

6. Lucas, B. D., *Generalized Image Matching by the Method of Differences*, PhD dissertation, Computer Science Department, Carnegie-Mellon University, July 1984.

7. Moravec, H. P., *Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover*, PhD dissertation, Computer Science Department, Stanford University, September 1980.

8. Baker, H. H., "Depth from Edge and Intensity Based Stereo", AIM 347, Stanford Artif. Intell. Lab, 1981.

9. Ohta, Y and Kanade, T., "Stereo by Intra- and Inter-scanline Search Using Dynamic Programming", *IEEE PAMI*, Vol. PAMI-7, No. 2, March 1985, pp. 139-154.

10. Herman, M., Kanade, T., and Kuroe, S., "Incremental Acquisition of a Three-Dimensional Scene Model from Images", *IEEE-PAMI*, Vol. PAMI-6, No. 3, May 1984, pp. 331-340.

11. Panton, D. J., "A Flexible Approach to Digital Stereo Mapping", *Photogramm. Eng. Remote Sensing 44*, December 1978, pp. 1499-1512.

12. Henderson, R. L., Miller, W. J., Grosch, C. B., "Automatic Stereo Reconstruction of Man-made Targets", *Computer Vision, Graphics, and Image Processing*, Vol. 1861979, pp. .

13. Vocar, J.M. and Faiss, R. O., "Image Magnification On STARAN", Tech. report GER-16342, Goodyear Aerospace Corporation, August 1976.

14. Terzopoulos, D.,, "Multi-Resolution Computation of Visible-Surface Representations", Tech. report, Dept. of EE & CS., MIT, 1983, Ph.D Thesis

15. Forgy, C. L., "The OPS5 User's Manual", Tech. report, Carnegie-Mellon University, Department of Computer Science, 1981.

16. Barnard, S. T. and Fischler, M. A., "Computational Stereo", *Computing Surveys*, Vol. 14, No. 4, December 1982, pp. 553-572.