STEREO CORRELATION FOR LARGE SCALE PHOTOGRAMMETRY

Gilbert Hobrough
Theodore Hobrough
Automatic Vision Corporation
USA
Commission II

1 INTRODUCTION

For the last three years the authors have been applying electronic stereo correlation to the design of 3D vision systems for industrial robots. Methods developed originally for photogrammetric purposes are being redeveloped to provide robots with stereopsis (binocular 3D vision). Our work has brought into sharp relief three performance limitations of existing stereo correlators: poor accommodation of object discontinuities such as the edges of buildings; poor utilization of the low contrast imagery that is characteristic of most surfaces; and loss of fine image data where coarse image data is absent.

All of the above problems have now been solved and the performance of stereo correlators can be raised to levels previously considered unattainable. The solutions result partly from the advance in digital devices over the past decade, partly from the advent of finite impulse response digital filters, and partly from the invention of new stereo correlation techniques. In particular, the techniques of 'video delay transformation', 'envelope correlation', and 'video processing' are involved. The purpose of this paper is to explain how the new technologies relate to the automation of photogrammetry and in particular to large scale urban and engineering applications.

2 GENERATIONS OF STEREO CORRELATORS

Given a single point in one image of a stereo pair it is impossible to find the homologous or corresponding point in the other image directly since the point in question is indistinguishable from the host of other points having the same density. Furthermore, the densities of homologous points are not necessarily equal.

If we enlarge the given point to a sample area that includes a number of resolution elements there is hope of finding the homologous area by pattern correlation. We could then regard the centres of the areas as homologous for the purpose of computing the position of the corresponding point in the object.

If stereo images were identical the size of the sample area could always be made large enough to ensure reliable correlation, assuming that the object presents resolvable detail to the cameras. Of course, stereo images are not identical since they represent views of the object from different points in space. The resulting relative displacement of homologous image elements in the X or baseline direction imposes a limit on what we call the coherence area of the images beyond which correlation is not possible.

In general coherence area is a function of spatial wavelength so that larger detail exhibits a larger coherence area than smaller detail. In rough terrain the coherence area of stereo images may fall to less than one spatial or video wave length.

The earliest stereo correlators, such as the Stereomat I, developed in 1960 by one of the authors at the Hunting Survey Corporation of Toronto Canada, did not address the coherence problem. As a result, such correlators could only utilize relatively coarse image detail in rough areas, and would fail completely where such detail was not present.

The second generation of correlators, such as Stereomat II performed first degree image transformation of the sample areas which allowed coherence to be extended substantially with a corresponding increase in the accuracy and reliability of the correlation process. The Ares correlator, developed by one of the authors at the Itek Corporation of Lexington Massachusetts USA, employed second degree transformation to rectify automatically a pair of convergent photographs into a viewable stereogram but without measurement facility. Polynomial transformation involves several degrees of freedom and correlation proceeds stepwise with the correction of each coefficient improving the visibility of the other coefficients.

First and second generation stereo correlators for photogrammetry were essentially single point correlators, the point being the center of the sample area. They were designed to guide the floating mark of a conventional plotting instrument in the production of planimetric and contour manuscripts and strip orthophotographs.

The third generation stereo correlator abandoned image matching by polynomial transformation in favor of transformation by direct element displacement. The resulting Gestalt Photomapper also departed from the single point concept by delivering a full window of dx values at a density of one dx value for every 64 pixels.

4 THE GESTALT PROCESS

The gestalt correlator is so called because the German word describes exactly the accumulating memory of form or shape that is essential to its operation. The method was invented and the name adopted by one of the authors in 1966 at the Itek Corporation. The original purpose of the invention was for the active shaping of large parabolic mirrors for use in an orbiting telescope.

The gestalt process depends upon synchronous raster scanning of the stereo images. Useful matching of raster and stereo anisotropies is achieved by making the scanning spot vector precisely parallel to the X axis or base line so that x disparity between the images becomes a simple time disparity between the video signals.

There are six essential features of the gestalt process:

1) The two stereo photos are correlated in homologous windows or patches about 500 pixels square.

2) The input video spectrum is divided into frequency bands each having a frequency ratio of 2:1. Typically six or seven bands are required.

3) The iterative matching sequence of the sensing and clearing of local disparities starts with the lowest frequency band, representing large image areas, and proceeds from band to band until the highest frequency band is reached. In the Rastar correlators the highest frequency band represents data from individual pixels..

4) The resultant dx error signals for each band are combined into a total dx error stream which is accumulated synchronously in a serial gestalt memory running at frame rate.

5) The images are re-shaped, in response to the accumulating error stream, by a transformation means into exact congruence over the entire window area, (virtually pixel to pixel in our current designs).

6) After matching, the data stream from the gestalt accumulator becomes a matrix of the x disparity values (for every pixel in the window in the current design).

The gestalt process was first applied to stereo correlation in 1971 by the authors at at Hobrough Limited of Vancouver Canada. Gestalt Photomappers are are high speed machines producing orthophotographs and digital elevation models with little operator intervention. They are currently manufactured by Northway-Gestalt Ltd. of Toronto Canada.

3 TRANSFORMATION METHODS

Image transformation during correlation, either polynomial or gestalt, has been a characteristic of all stereo correlators developed by the authors after Stereomat I. Image transformation has been performed by changing the relative scanning velocity between the left and right image sensors. Difference in the arrival time of left and right image detail is sensed by an X disparity discriminator and the resultant error signal used to modify the local scanning velocity.

The technique is analog and works very well except when steep terrain slope or a vertical discontinuity is encountered. To execute a vertical step the scanning spot must move instantaneously from one point to another in one or both images. Such instantaneous change is not possible owing to the energy stored in the deflection fields of the image scanning devices. In fact smoothing circuits have been incorporated into Gestalt correlators to avoid the possibility of exceeding the acceleration limit of the scanning spot which could cause the feedback correlator to go open loop and lose contact with the terrain.

There are secondary problems associated with scanning velocity change as it affects the performance of image scanning devices. Vidicons and solid state array sensors owe their high sensitivity to their light integrating action. Charges liberated over the total frame period are collected at the instant of scanning each pixel giving a gain over non-integrating sensors equal to the number of pixels or about 200,000 times for a TV format. Unfortunately, the output current is equal to the rate of charge collection which is a linear function of scanning velocity. As a result velocity change induces level change or shading that must be compensated to obtain a useful video signal.

The above limitations have not proved to be serious for small scale topographic mapping. Errors introduced by the "rounding off" of buildings and terrain steps have not degraded map accuracy nor have they caused correlation failure except in the most rugged terrain.

A new transformation process developed by the authors eliminates the need for velocity modulation and its problems and permits Z steps of any size within the Z range of the system. We call the process Video Delay Transformation and it depends on a digital device that can delay individual pixels by an arbitrary number of video clock periods. The Rastar Mapping Correlator, developed jointly by the authors and the Institute for Photogrammetry of the University of Hannover, incorporates video delay transformation and has been described in some detail by Konecny and Pape (Photogrammetric Engineering and Remote Sensing, March 1981).

Since transformation by video delay requires video in digital form, digital filtering and image processing techniques can easily be applied. Non-recursive or Finite Impulse Response (FIR) digital filtering is ideally suited to feedback correlation since the time delay of such filters is independent of frequency. As a result synchronism is maintained for all components in each band thereby eliminating one of the problems in earlier machines.

Digital video also offers the opportunity for compensating resolution loss from lens aberration and defocussing. Compensation improves both the correlation of fine detail and the resolution of orthophoto printing. It is necessary for compensation that aberrations be constant only over the area of a patch so that compensation can be separately adjusted for central, corner and intermediate photo zones.
Video delay transformation can be readily applied to the gestalt process and offers the following advantages:

1. Elimination of the "ramping" of Z steps and vertical structures in the stereo model and of the consequent distortion in orthophotos and DEM's around such features. This resolves the first of the performance limitations set forth in the introduction.
2. Elimination of slope limits.
3. Elimination of sensitive analog circuitry for raster shaping and brightness compensation.
4. The use of array sensors in place of vidicon image sensors and the consequent elimination of deflection yokes and analog driving circuits.

The Rastar Mapping Correlator project saw the application of such a system to a standard analytical plotter using the single line array sensors available at that time (1976-78). The use of a single line sensor leads to difficulties in correlation and complexity in photo transport programming. These difficulties can now be avoided with area arrays currently available or projected for production in 1984. A GPM-type instrument using area array sensors could be developed at any time and would yield hitherto unavailable speed and performance at essentially no increase in cost.

The resulting ability to accommodate large vertical steps or discontinuities would enable stereo correlation to be used for the automatic printing of orthophotographs and digital terrain coordinates for large scale urban and engineering projects that seem to be the major market area for such map products.

The remaining difficulty in large scale mapping arises out of tall structures hiding the ground and is a problem of geometry not correlation. The use of longer focal length cameras eases the geometry problem and allows flying at a greater height, usually a convenience at very large scale. Also, the smaller field angle improves correlator performance still further.

Industrial robotic applications are similar to large scale mapping in that the surfaces of objects seen by the robot are not connected to the background nor to each other much the way the tops of buildings are disconnected from the ground. It was this similarity that led us into the robotic vision field and leads us to consider the application of our current technology once more to photogrammetry.

5 THE FEEDBACK STEREOPSIS SYSTEM

Figure 1 is a simplified diagram of a feedback stereopsis system for industrial robots. A system for photogrammetry would need a Y disparity system and computational elements for analytical functions not required in a robot stereopsis system.

The second performance limitation mentioned in the introduction (poor utilization of low contrast imagery) is corrected by the video processors. First, they equalize the effectiveness of highlight and shadow detail by taking the logarithm of the video signal. Second, they eliminate the fixed noise arising out of variations in the pixel to pixel sensitivity and dark current of the sensor arrays. Third, they adjust the brightness range of the video signals to match the capacity of the digital video channels. Another function of the video processors, not connected with contrast problems, is to detect and eliminate the disturbing effect of specularities on correlation.

The Transformation Delay Unit (TDU) transforms the images by shifting each pixel of each image in the X direction in response to a dx control signal, the pixels of the left and right images are shifted equally but in opposite directions.

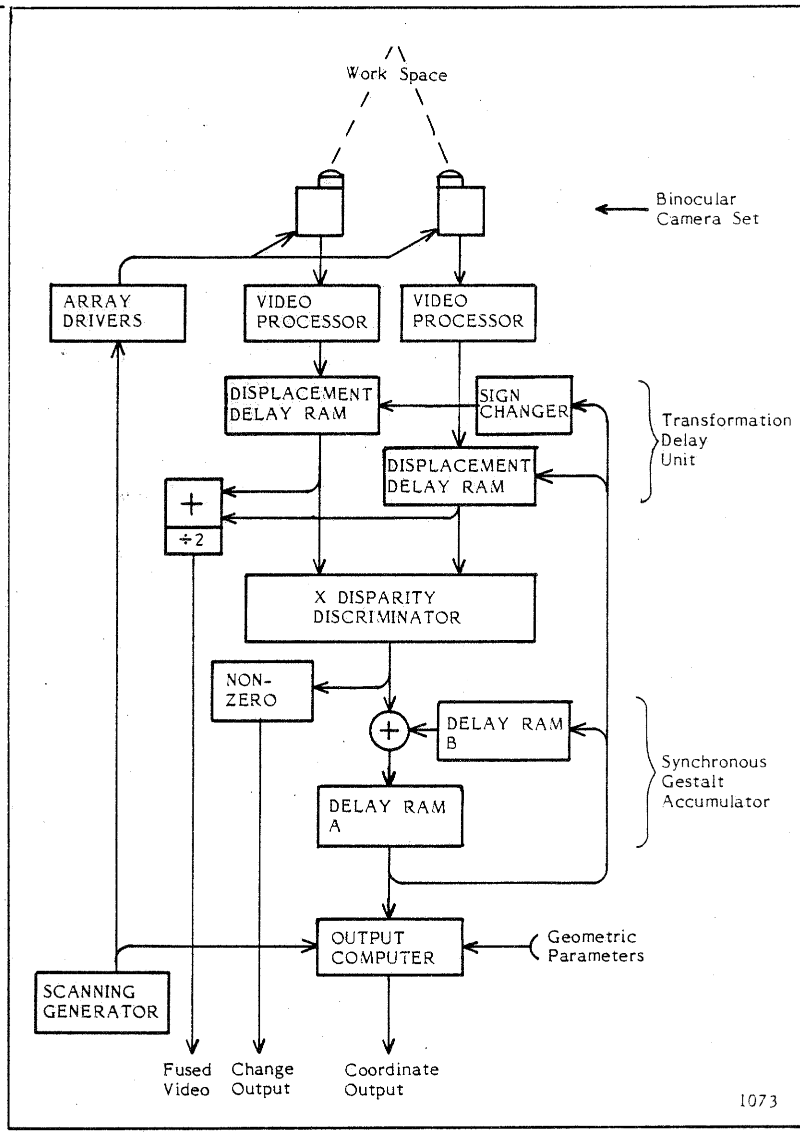

Figure 1 ROBOT STEREOPSIS SYSTEM

The X Disparity Discriminator (XDD) senses the magnitude and sign of local x disparity and delivers a string of local dx signals to the-

Synchronous Gestalt Accumulator (SGA) which stores the dx values for each pixel of one entire frame or field of view. During the scanning of the each frame the stored dx signal is fed back to the TDU as a control signal to shift the pixels in the X direction to reduce X disparity. Residual dx is then sensed by the XDD and the resultant dx signal is delivered to the SGA where it is added to the previous dx value for the appropriate pixel.

After a few iterations all X disparity between the inputs to the XDD is eliminated and the output of the XDD falls to zero. The output of the SGA then contains all of the dx data for the entire window matrix in serial form.

6 WHY FEEDBACK MATCHING AND BANDS?

1) To facilitate the coarse-to-fine iteration process.
2) To promote the spread of correlation.
3) To locate automatically all homologous pixels in the window.
4) To eliminate automatically detail present only in one image.

Stereo correlation involves multiplication of input data followed by integration or the summing of products over an area. The larger the area that can be integrated the less the noise or roughness in the Z signal. We define the relative coherence area as the largest circular portion of the stereo image areas that can be regarded as identical and over which dx is constant. Integrating beyond the limits of the coherence area sharply increases Z noise and reduces the reliability of the Z data.

A major problem in the design of stereo correlators is the small coherence area of stereo imagery before transformation, typically about one video cycle in both X and Y. Therefore the low frequency video components, representing large image detail, have a larger coherence area than higher frequency video components representing small image detail.

Dividing the video spectrum into frequency bands allows the integrating area to be separately matched to the coherence area for each band and also provides a basis for progressive feedback matching.

Figure 2 illustrates the progressive clearance of disparity for an object in the form of a layered pyramid. A large B/H ratio of 1/1 has been used to exaggerate the differences between the stereo images. The cameras are placed at infinity to simplify the drawing.

Unmatched parts of the images are shown diagonally hatched. Matched parts are shown clear and parts seen by only one camera are shown black. Four band iterations are shown and are sufficient to illustrate correlative matching on the simple object chosen. Six or seven bands are required to resolve fully the imagery of a typical 500 by 500 pixel patch.

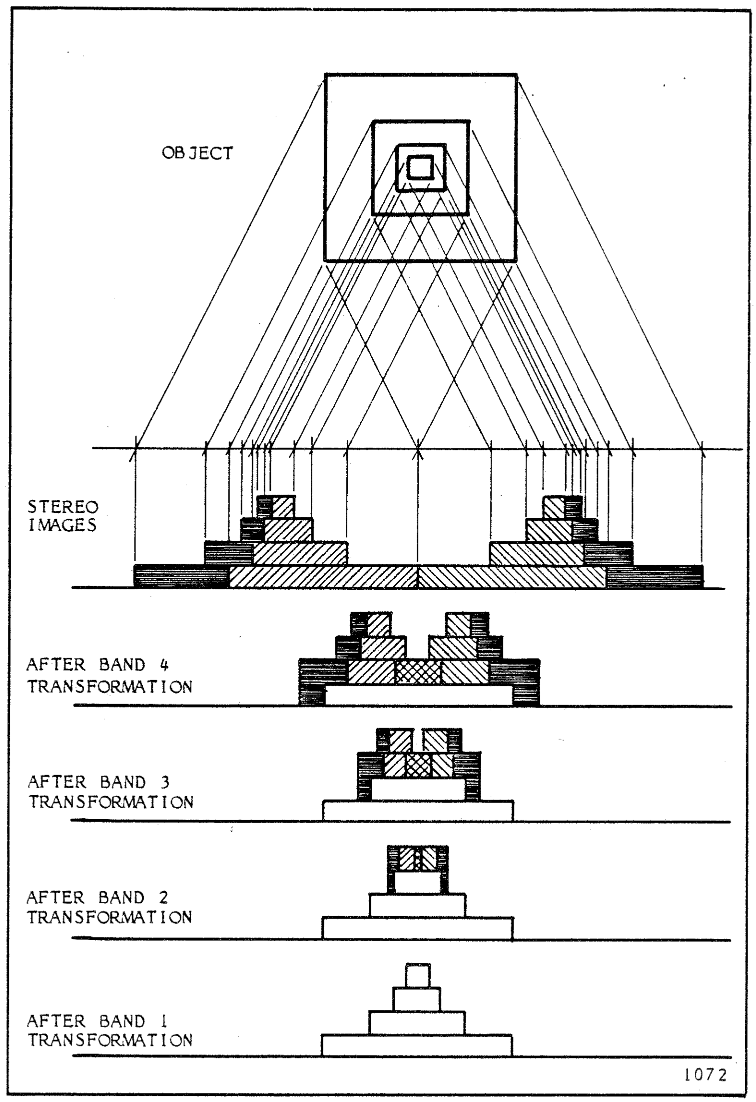

Figure 2
PROGRESSIVE DISPARITY CLEARANCE

7 ENVELOPE CORRELATION

In this section we deal with the third performance limitation mentioned in the introduction -- the loss of fine image detail where no coarse image data is present.

Progressive matching of the images from coarse-to-fine detail is only possible if all of the bands contain correlatable image data. If data is missing from one band then the discriminator for the next higher band may not be within its capture range and the process may come to a halt. While such occurrences are not frequent with terrain imagery, a human operator must monitor the very fast Gestalt Photomappers and take corrective action when necessary. The function of Envelope Correlation is to eliminate the problem by providing correlatable envelope data in bands where direct signal data is missing.

Figure 3 illustrates the difference between a composite signal and a modulation envelope.

The left column of figures shows how band pass filters separate low and high frequency signals prior to correlation. The right column shows how to extract a low frequency envelope signal from a modulated high frequency signal that contains no direct low frequency data.

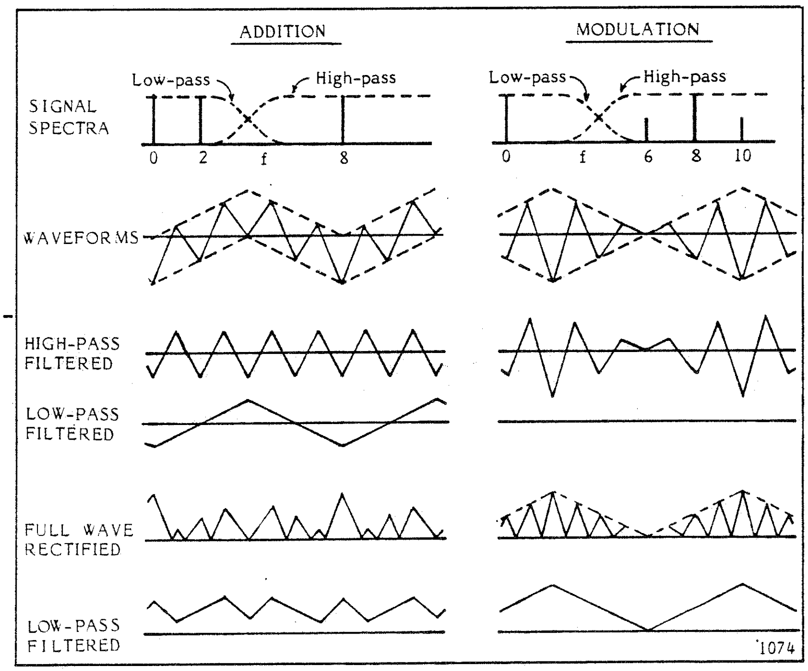

Figure 3
SIGNAL ADDITION vs. MODULATION

Figure 4 shows how frequency changes in a constant amplitude signal can give rise to low frequency envelopes in several higher frequency bands.

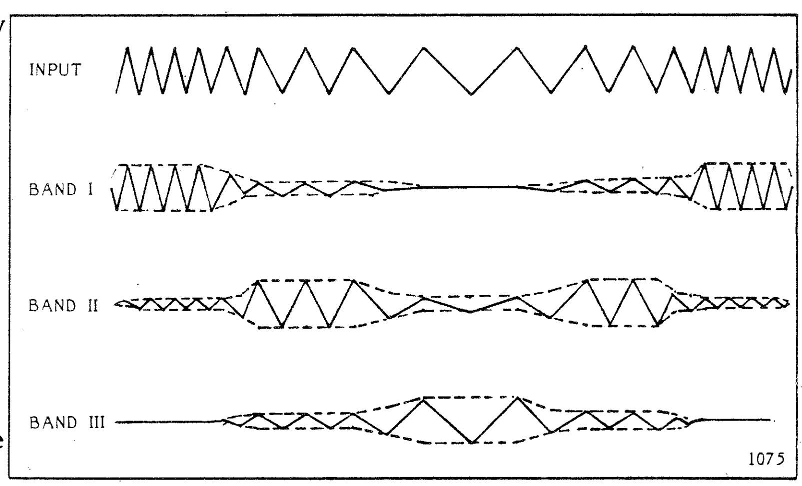

Figure 4
ENVELOPES FROM FREQUENCY CHANGE

Envelope Correlation is an easy addition to multiband Gestalt type stereo correlators and increases greatly the reliability and speed of the correlation process. In photogrammetry the advantage of envelope correlation would be particularly useful in finely textured areas without shading and in tree covered areas under high sun conditions where leaf detail is strong but shading is weak.

Envelope data may seem to come from nowhere. Actually, it is data present in the images that has not been used by previous correlators. Envelope data is particularly useful under certain conditions. For example repeating patterns, such as a plowed field or an orchard, tend to confuse stereo correlators even more than they confuse the eye. A false match between one furrow and its neighbor can introduce height errors of multiples of the furrow spacing times B/H. With envelope correlation low frequency data in the form of contrast or spacing differences provide a lower band error that can override the false match error regardless of the clarity and contrast of the furrows or trees.

Another application for envelope correlation, perhaps without any direct signal correlation, could be the matching of side looking radar imagery taken from different angles. Such material is not normally correlatable owing to the incoherence of the scatter patterns. One would expect the scatter pattern envelopes to be coherent however and the correlative matching of side looking radar imagery should be possible.

8 CONCLUSION

The automation of photogrammetric plotting has developed at a slow and sometimes unsteady pace for many years. Stereo correlation has presented a series of technical problems that have required unique solutions having virtually no application except in topographic mapping. It has been difficult to justify the investment required to develop solutions in view of the small number of instruments likely to be sold. In commercial terms the two most successful automatic plotters to date have been the B-8 Stereomat and the Gestalt Photo Mapper and only about a dozen of each have been built.

There has been little activity in other industries contributing directly to the cause of photogrammetric automation except for the tremendous advance in small computer technology. The situation is rapidly changing as robotization sweeps through the world's manufacturing industries. The needs of robotic vision creates a huge demand for new systems and methods, some of which will be usable directly in photogrammetric applications. For the first time, a wide variety of sophisticated yet inexpensive tools are becoming available to the designers of automatic stereo plotters and interpreters. We believe that the swelling wave of robotization will affect the mapping industry as profoundly as it is now affecting the automobile industry.

Technology now exists that can extend the performance of stereo image correlators for photogrammetric plotting to the physical limits of the input material and at a very high productivity. Very large scale integration chips and the newer semi-custom devices greatly reduce the design time and equipment cost compared with earlier methods. The result will be a cost effectiveness in the next generation of automated photogrammetric instrumentation enormously higher than we have come to expect from the current generation.