Experiments on Target Location and Image Matching.

J.C. Trinder
School of Surveying
University of N.S.W.
Kensington NSW 2033
Australia

ABSTRACT

The development of software for the determination of object geometry from overlapping images involves the tasks of pointing to targets, standard bundle solution of camera parameters using selected object control points and the matching of corresponding details of the two images. Extensive investigations have been carried out on digital pointing to circular targets revealing the influence of image quality, pixel size in the image, quantization level and noise on the precision of pointing. Asymmetry of the target profile reveals significant systematic errors in the pointing accuracy. Under ideal circumstances, the precision of pointing can approach 0.01 pixel size. The process of image matching is currently being tested using interest operators such as edge detectors to assist in matching based on features.

INTRODUCTION

With the advancements in computer technology in speed and storage capacity, it has become possible to derive images in digital form and carry out the image processing for photogrammetric applications. Indeed real-time image processing for photogrammetric measurements has become a reality. There are many potential examples of close-range applications in photogrammetry where such digital methods could be used e.g. medicine, industry. This paper will describe components of a system being developed for the location of targets, and the computation of object geometry from digital images. The system is based on a video camera and an IBM-AT computer incorporating a Imaging Technology frame grabber for analogue/digital conversion of the data and also storage of the video data. Extensive studies have been undertaken of the precision of target location and the systematic errors brought about by asymmetry in the target intensity profile. Results of these studies will be given in this paper. In addition, image matching by least squares is being developed as part of the package.

TARGET LOCATION

Wong et al. (1986) have determined methods for the location of circular targets on digital images, while Mikhail et al. (1984) have studied the detection and location of edges and cross-targets. In this paper circular targets have been used, and therefore the method adopted is a modification of the method of Wong et al.

A window is approximately centred on the target and thresholding within the window is then carried out, initially based on Wong's value of:-

$$\text{Threshold} = (\text{Min pixel value} + \text{mean pixel value})/2 \qquad (1)$$

All values below the threshold are set to zero while those equal to or greater than the threshold are set to 1. This process isolates the pixels defining the targets, which should be used for precise location of its centre. Wong et al. chose to compute the position of the target by taking the centre of gravity using the formula:-

$$x=\frac{1}{M}\sum_{i=1}^{n}\sum_{j=1}^{m} j.g_{ij}$$

$$y=\frac{1}{M}\sum_{i=1}^{n}\sum_{j=1}^{m} i.g_{ij} \qquad (2)$$

$$M=\sum_{i=1}^{n}\sum_{j=1}^{m} g_{ij}$$

where g_{ij} is the value of each pixel, either 1 or 0 located in row i and column j.

However, investigations revealed that target location by this process led to variations in target position when the quotient in equation (1), or the location of the window was altered. This meant that low intensity pixels on the edge of the target had a disproportionately large influence on the location of the target. Therefore equation (2) has been changed to

$$x=\frac{1}{M}\sum_{i=1}^{n}\sum_{j=1}^{m} j.g_{ij}\ w_{ij}$$

and similarly for y

(3)

$$M=\sum_{i=1}^{n}\sum_{j=1}^{m} g_{ij}\ .\ w_{ij}$$

where i,j are the same as in equation (2) $w_{i,j}$ is a weighting factor applied to each pixel, being equal to the intensity of the pixel above threshold. This meant that the central high intensity pixels influence the pixel location more than the surrounding low intensity pixels.

PRECISION OF TARGET LOCATION

In an extensive series of tests, the accuracies of target location using equation (3) have been determined on artificially generated targets, with varying characteristics. Blurred targets, typical of those which would occur on photography were generated by convolution of circular targets of varying sizes and Gaussian spread functions with 2σ - widths ranging from 10 μm to 50 μm. Typical spread function 2σ - widths found on aerial photography range from 15-25 μm Trinder (1984). The result of this convolution was a profile across the target which would be symmetrical about the target centre, assuming that there was no asymmetry in the spread function forming the blurred image.

The process of digitizing involves the determination of the area contained under the target profile within the dimensions of the pixel as shown in Figure 1, for each pixel location, if a single profile is considered. If the target is assumed to be 2-dimensional, the volume contained within the dimensions of the square pixel, and the surface describing the intensity of the target over its dimensions, would have to be computed.

The profile computed above therefore was rotated about the target centre and the pixel values for each pixel location computed over the 2-dimensions of the target. This was undertaken in scan-lines across a square window centred on the target. In order to obtain a meaningful estimate of the precision of pointing to targets, a displacement was introduced into the starting position of the pixels along the scan-lines. This meant that the distribution of pixels would not necessarily be symmetrical on the centre of the target, and the intensities of the pixels would therefore be asymmetric, as shown for comparison in Figures 2 (for symmetrically located pixels) and Figure 3 (for asymmetrically located pixels). A major task in this study was to determine the effects of the asymmetric distribution of the pixels on the target, on the precision of target location.

Having obtained the pixel values from this process of digitizing, the next step required was quantization into a selected number of grey scale values which would be typically used in the process of digitizing an image. Eight different quantization levels were selected equivalent to 2^8, 2^7.....2^1 grey scale values which means encoding into 8,7.....1 bits respectively.

Following quantization, the precision and accuracy of target location were derived. This process was repeated 50 times for each target with the commencement of the scanning by the pixels displaced for each new digitizing process by a random number varying up to ±1 pixel. The exact position of the target was naturally known on the artificially generated targets. A standard deviation of pointing and the systematic error in target location could therefore be computed from the 50 pointings.

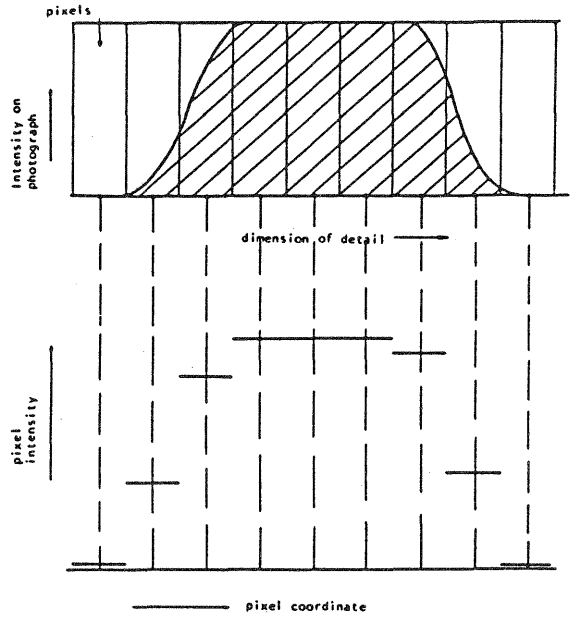

Figure 1 Process of digitizing by simulation.

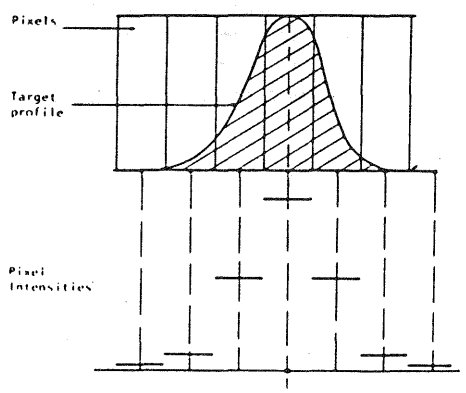

Figure 2 Distribution of pixel values obtained when the positions of the pixels are located symmetrically on the target.

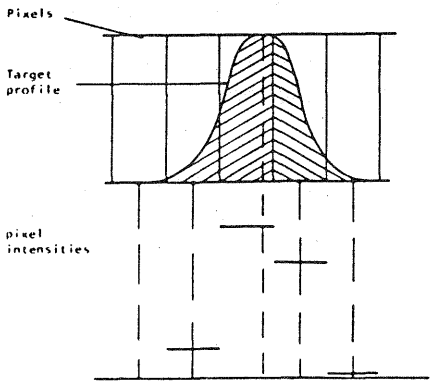

Figure 3 Asymmetric distribution of pixel values resulting from an asymmetric location of pixels on the target.

Noise was also introduced into the values of the pixels following digitizing and prior to quantization. The noise was computed as a random number within a positive and negative range of a certain percentage value of the maximum intensity value of the target. The percentages chosen were 10%, 20%, 40%, and 80%, equivalent to signal to noise ratios of 10:1, 5:1, 2.5:1 and 1.25:1 respectively.

Since this test was based on artificially generated targets, parameters of target size, image quality defined by the Gaussian spread function, pixel size and noise level, could be varied and the precision of pointing determined. In addition, it was possible to introduce targets whose profiles were subject to asymmetry in image quality, that is, different slope characteristics on one side of the target intensity profile than the other, and the effect on pointing accuracy and precision determined.

The results of these investigations are presented by Figures 4,5,6 and 7. For the cases where image quality was symmetrical, there was no effect on the precision of pointing as the 2σ - width of the Gaussian spread function varied from 10 to 50 μm. In Figure 4 precisions of pointing are shown for a target size of 100 μm, pixel sizes of 12.5 μm, 25 μm and 50 μm against the number of bits used for encoding the data. In Figure 5 are shown the variation in precisions for target sizes from 25 μm to 200 μm for a pixel size of 12.5 μm, and different levels of quantization. Larger targets were tested, but precisions of pointing to these targets were similar to those shown in Figure 5 for 200 μm target. In Figure 6 are shown the variations in precisions of pointing to 100 μm targets and pixel sizes of 12.5, 25, 50 μm for random noise for a quantization of 8 bits, expressed in terms of SNR, while in Figure 7 are shown the pointing precisions for targets 25 μm to 300 μm in size for a pixel of 12.5 μm,also in terms of SNR.

The precision obtainable by digital pointing is approximately 0.01 pixel size and slightly better on certain occasions as shown in Figure 4. This result has been obtained consistently for all target sizes and for all pixel sizes for quantization levels of 8 bits/pixel. There is a general deterioration in precision as quantization levels decrease especially below 4 bits/pixel, but also as pixel sizes increase. Figure 4 indicates that a pixel size of 12.5 or 25 μm would result in consistently high precisions for quantizations higher than 5 bits/pixel. The influence of the relative dimensions of the pixel in relation to the target size is shown in Figure 5, where precision is also influenced by quantization level.

The sections of graphs in Figure 4 below a quantization of 5 bits can be expressed by equations such as that shown in equation 4:-

$$\text{Precision} = (\text{QUANTIZATION})^{-1.4}\, K \text{ (pixels)} \qquad (4)$$

where K is non-linearly inversely proportional to pixel size, being 9 for 12.5μm pixels and 4 for 50 μm pixels. On this basis, the remainder of the investigations described in this paper will be limited to quantization of 8 bits.

The variation in precision in terms of SNR in Figure 6 demonstrates that for SNR greater than about 5:1, the precision is very similar to that with no noise, but below SNR of 5:1, clearly the precision deteriorates rapidly. This deterioration is even greater as the target size becomes smaller in relation to pixel size, as shown in Figure 7. It is significant that in Trinder (1984) the SNR below which visual pointing precisions for circular targets deteriorated was about 5:1. There is a striking similarity in the 2 cases of visual and digital pointing in the levels of noise which affect pointing precision.

The graphs in Figure 6 for SNR below 5:1 can be expressed by formulae such as equation 5

$$\text{Precision} = (\text{SNR}^{m}.K)^{-1} \text{ (pixels)} \qquad (5)$$

where m varies from 1.7 to 1.2 for targets of 50 to 200 μm in size and K varies linearly with target size. For target size of 300 μm, the relationship does not hold however.

Converting the formula to a similar form to that shown by Forstner (1982) for the precision of least squares image matching

$$\text{Variance of pointing} = K_1\, (\text{SNR}^{T} . N . \text{Pixel size})^{-1} \text{ (pixels)}^{2} \qquad (6)$$

where K_1 is a constant = 37.5 for Figure 7, T varies from 3.3 to 2.3 for target sizes of 50, to 200 μm and N is the number of pixels making up the dimension of the target. Though this formula is similar in form to that shown by Forstner (1982), the magnitude of the constants is inconsistent. This formula however, reveals the significance of the 3 parameters of SNR, target size (i.e. number of pixels) and pixel size on the precision.

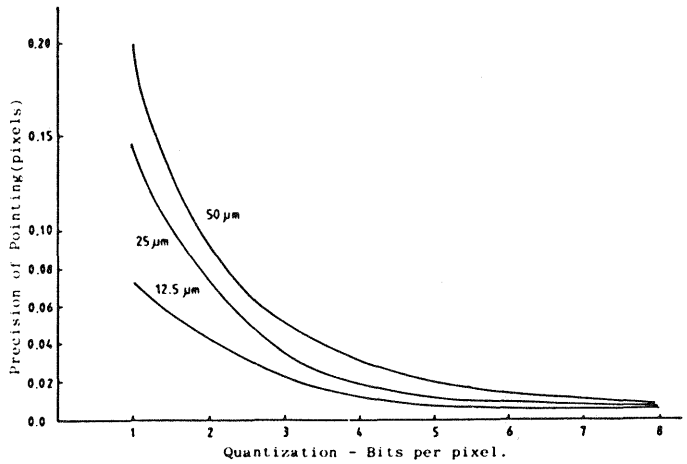

Figure 4 Relationship between pointing precision and quantization level for target size of 100 µm for 3 pixel sizes.

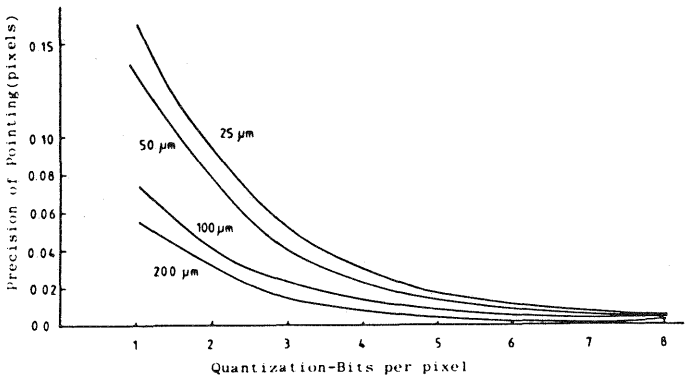

Figure 5 Relationship between pointing precision and quantization level for 4 target sizes using pixel size of 12.5 µm for digitization.

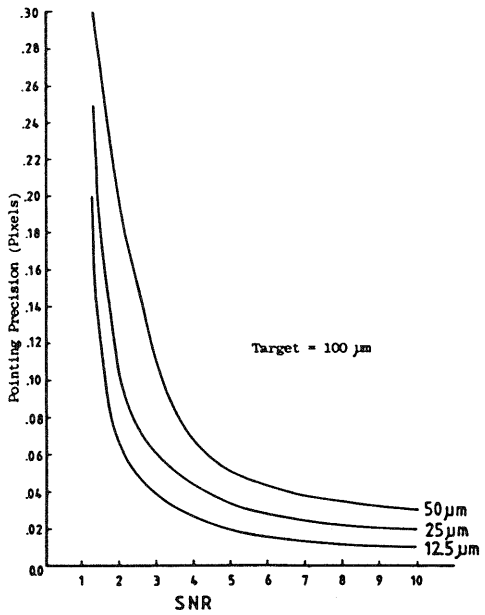

Figure 6 Relationship between pointing precision and SNR for target of 100 µm and 3 pixel sizes.

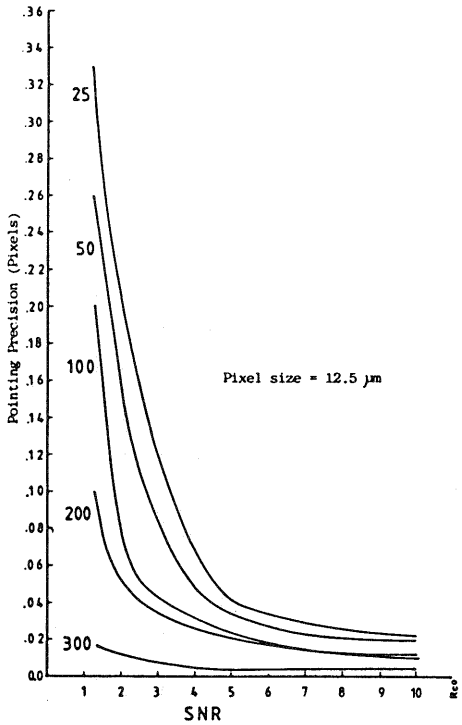

Figure 7 Pointing precision expressed against SNR for 5 target sizes using pixel size of 12,5 µm.

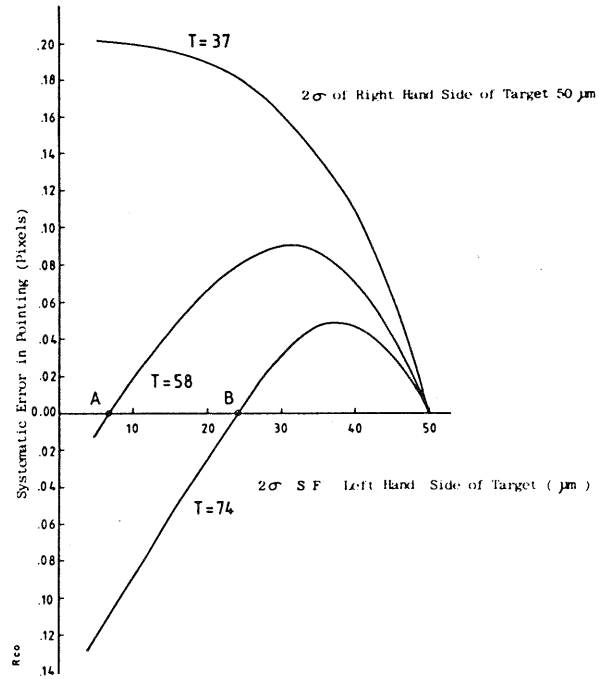

Figure 8 Relationship between asymmetry in image quality of target 100 µm wide and systematic error in target location. Abscissa scale describes the 2 σ - width of spread function defining the left hand size of the target, while the right hand side of the target is derived by a spread function with 2 σ-width of 50 µm. Quantization is 8 bits. T defines the threshold chosen in the pointing computation.

ACCURACY OF TARGET LOCATION

The accuracy of target locations derived in the above studies was of the same order as the precision. That is, for repeated pointings on symmetrical targets the mean position of the target was the correct centre within the precision of pointing. However, as the target profile becomes asymmetric, a simple pointing operation using the threshold in equation (1) will result in significant errors in pointing. This phenomenon was investigated thoroughly by convolving the target with different spread functions on each side. Three nominal spread functions were adopted for the right hand side of the target viz 10,25 and 50 μm. Then the spread function defining the profile on the left hand side of the target was varied in steps of 5 μm from 5 μm to 50 μm. The systematic errors in target location were then derived. In addition, the threshold values were varied to investigate its effect on target location.

The deterioration in pointing accuracy as the target profile increased in asymmetry is demonstrated in Figure 8. For threshold levels set at 0.14, 0.23, and 0.29 of the maximum pixel value, (i.e. grey scale values of 37, 58 and 74 respectively) the target location varied significantly for different levels of asymmetry in target. The investigations in this study revealed that for a particular target size and level of asymmetry a threshold can be chosen to obtain the most accurate target location as shown by A and B on Figure 8. Similar graphs can be drawn for other sizes of targets, but it was found that the relationship is inversely proportional to target size.

Considering practical applications of this work it can be safely assumed that for most circumstances the asymmetry in the target profile will not be greater than 20% of the assumed value. It can also be normally assumed that the approximate value of the spread function will be known. If it is not then the spread function it should be determined approximately because it influences the accuracy of the pointing process. A simple algorithm for the computation of the threshold has been derived assuming quantization to 5-8 bits, as follows:-

$$\text{Threshold} = 74.\,(\text{SF})^{1.3}\,(\text{Target size})^{-1} \qquad (7)$$

where SF refers to the 2σ - width of the Gaussion spread function.

Tests of this algorithm demonstrate that the choice of threshold is efficent in reducing the systematic errors to approximately 0.01 pixel with and without noise introduced into the data, except for cases where the target is small in relation to size of the spread function. Indeed, if the target size is very small, and the image quality poor, it is impossible to locate the target accurately because the target profile becomes so distorted.

In Trinder (1984), the recommended optimum target sizes in relation to spread function width are given for visual observations. Such rules cannot be applied directly to digital pointing because different phenomena affect pointing precision on both cases. From the study of systematic errors in this paper, it is recommended that targets sizes should be 6-8 times the σ - width of the spread function of the system. Provided this rule is followed, systematic errors in pointing will be less than 0.02 pixel, if the correct threshold value is incorporated in the computation.

FEATURE BASED IMAGE MATCHING

Feature based matching as opposed to image matching based on grey scale intensities, is a method used in pattern recognition, and has recently been investigated in photogrammetric applications by Forstner (1986), Piechel (1986) Lohman et al. (1986) amongst others. Feature based matching uses extracted information in the image as a basis of image matching, such as edges or other distinct features. The method therefore is based on a higher level image data than the raw pixel values.

Currently, the main efforts of this research are directed towards extracting edges or localized sharp changes in intensity in the image and their use in the image matching process. These edges define regions in an object which should provide information for extracting three dimensional information of the object. It is assumed that epipolar lines can be computed on each image from the camera orientation parameters obtained from the bundle solution. Knowledge of the epipolar lines is an important aid in the solution of the image matching problem, since within the accuracy of the bundle solution, it reduces the problem to primarily one dimension.

The detection of edges in images has been shown to be a characteristic task in vision. Indeed stereopsis involves the extraction and matching of contours to enable observers to achieve good depth perception. Significant advances have been made by Marr and Poggio (see Marr and Poggio 1979) in studying this phenomenon. Attempts to copy the visual system by some researchers studying computer vision have therefore incorporated edge extraction procedures.

A simple edge detector algorithm for digital images can be derived from a 2x2 pixel window, which is used to scan an image.

$$E_{i,j+1} \quad E_{i+1,j+1}$$

$$E_{i,j} \quad E_{i+1,j}$$

FIGURE 9

The derivatives or slopes at the centre of this group of pixels is:

$$\frac{\partial E}{\partial x} \approx \frac{1}{2e}\left[(E_{i+1,j+1} - E_{i,j+1}) + (E_{i+1,j} - E_{i,j})\right]$$

$$\frac{\partial E}{\partial y} \approx \frac{1}{2e}\left[(E_{i+1,j+1} - E_{i+1,j}) + (E_{i,j+1} - E_{i,j})\right] \qquad (8)$$

where e is the spacing between the pixels and $E_{i,j}$ etc. are pixel values (Horn, 1986).

Hence, the square gradient is computed from

$$\left(\frac{\partial E}{\partial x}\right)^2 + \left(\frac{\partial E}{\partial y}\right)^2 \approx \left[(E_{i+1,j+1} - E_{i,j})^2 + (E_{i,j+1} - E_{i+1,j})^2\right] \qquad (9)$$

A large number of edge operators or kernels have been used by various researchers to detect edges. These operators which act as digital filters on the data, are convolved with the image pixels to reveal the changes in slope of the intensities of the pixels. Orientation of the edge as well as location can be derived. Burns et al. (1986) have made an extensive investigation of edge detection operators and found the simple 2x2 shown in Figure 10 as the most suitable for x and y directions. However, many more complex operators exist. In this study, a combination of median filter followed by a Roberts filter (similar to that shown in Figure 9) have been used for edge detection.

-1	-1	-1	1
1	1	-1	1

FIGURE 10

The extracted edges provide the basic information for the first step of image matching. Epipolar lines should be intersected on the edges on the target image in selected positions to commence the search for edges. Connectivity of neighbouring pixels is then determined within a window by analysing which of the neighbouring pixels around each pixel contains information about the edge. Pixel positions of the edge in the window will be recorded in lists which must be used for the image matching. Having extracted segments of edges on the target image, a similar search will be carried out on the search image with approximations for positions of corresponding edges being obtained from epipolar lines and the nearest control point or the nearest matched edge. The initial test of matching the two edge segments will be based on length and direction of the two segments. Then matching will be based on the least squares method to precisely located the matching edges. Once edge matching is completed, image matching on a grid over the whole object can be carried out with the information derived from the nearest matched edge as the initial information for the localized grid point. As this method is currently still being developed no results are available yet.

CONCLUSIONS

(i) The precision of digital pointing to symmetrical targets will depend on size of pixel, level of quantization, and level of noise,but can approach 0.01 pixel under the best circumstances.

(ii) Systematic errors in pointing to digital data can exceed 0.2 pixel if the profile of the target is asymmetric. Appropriate choice of threshold can almost eliminate these systematic errors.

(iii) Image matching is being undertaken using edges extracted from the data to provide initial values for the matching of grid points. The paper describes the method of extracting the edges in the image data, and matching the two images based on these edges.

REFERENCES

Burns J.B., Allen R.H. and Riseman E.M (1986) "Extracting Straight Lines" IEEE Trans. PAMI-8 pp. 425-455

Forstner W. "On the Geometric Precision of Digital Correlation". International Archives of Photogrammetry. Vol 24-III pp 176-189.

Forstner W. (1986) "A Feature Based Correspondence Algorithm for Image Matching" Int. Arch. of Photogramm. and Rem. Sens. Vol. 26-3/3 pp. 150-166

Horn B. (1986) "Robot Vision" M.I.T. Press Cambridge USA

Lohman T. and Altroppe G (1986) "Interest Operators for Image Matching" Int. Arch. of Photogram. and Rem. Sens. Vol 26-3/2 pp. 459-474

Marr D. and Poggio T. "A Computational Theory of Human Stereo Vision" Proceedings of the Royal Society of London B. Vol 204 No. 1156 pp. 301-328

Mikhail E.M., Akey M.L. and Mitchell O.R. (1984) "Detection and Sub-pixel Location of Photogrammetric Targets in Digital Images" Photogrammetria Vol. 39 63-83

Piechel J. (1986) "Investigations of Different Interest Operators for DTM Generation". Int. Arch. of Photogram. and Rem. Sens. Vol. 26-3/2 pp. 564-572.

Trinder J.C. (1984) "Pointing Precisions on Aerial Photographs" Photogram. Eng. and Rem. Sens. Vol. 50 pp. 1449-1462.

Wong K.W. and Wei-Hsin H. (1986) "Close-Range Mapping with a Solid State Camera" Photogram Eng. and Rem.Sens. Vol. 52 pp. 67-74.