# THE AUTOMATIC RECOGNITION, LOCATION AND LABELLING OF TARGETS IN DIGITAL PHOTOGRAMMETRIC ENGINEERING MEASUREMENT

J. Chen & T.A. Clarke
Department of Electrical, Electronic, & Information Engineering,
City University, Northampton Square,
London, EC1V OHB,
ENGLAND.

**ABSTRACT:**

A major problem in digital photogrammetric analysis of engineering structures using multiple views and bundle adjustment techniques is the precise recognition, location, and labelling of targets. This task is complicated by: variable surface reflectivity, non-ideal illumination, occlusion, and the variation of target size with distance from camera stations. In this paper the authors discuss the use of multiple images of a targeted section of a wind power generator blade to automatically calculate the coordinates of the targets. Three aspects are discussed: target recognition using a binary image, subpixel location using a grey scale image, and unique labelling.

(i) The target recognition process uses a binary matching technique where prior knowledge of the target characteristics are used to reject spurious targets.
(ii) The local region around each target image is independently analysed and an adaptive location algorithm used to give a subpixel estimation of the position of the target.
(ii) A small number of control targets are used to perform a transformation of the camera images to enable reliable automatic labelling of each unique target image.

Results from applying these methods to provide the coordinates of a targeted section of a wind power turbine blades are presented, and the accuracy and efficiency analysed.

**KEY WORDS:** Targets, Recognition, Image-processing, Subpixel, 3-D.

## 1. INTRODUCTION

In digital photogrammetry, using multiple camera views, the 3-D spatial relationship of targets which are placed on an object can be determined using the bundle adjustment technique from the 2-D spatial locations of the targets in each camera view. An important consideration, if the full potential of this system is to be exploited, is unique target identification which must take place reliably if the resulting 3-D spatial coordinates of the targets are to have any meaning. It is common for this process to take place by moving a cursor to the target and labelling it according to some predetermined scheme. However, this method requires a skilled operator and is time consuming with the possibility of poor reliability. Another aspect is the precision with which the 2-D coordinates of the targets can be obtained.

With the development of computer vision and pattern recognition techniques the use of digital image processing of CCD camera produced data is being more widely used (Gruen, 1989). Bethel, 1990, discusses the fact that photogrammetrists might benefit by looking towards computer vision techniques where some methods have become established over many years for dealing with digital imagery (Bhanu, 1986). The methods applied here rely heavily on a machine vision approach to the problems of target recognition such as described by West, 1988. The machine vision community primarily works in image space to perform task such as: matching, pattern recognition, target location, edge detection, etc. Photogrammetrists are more used to mapping image coordinate information back into the object space. Some cross fertilisation may be of mutual benefit. This paper represents the work of two machine vision researchers who are attempting to learn from the photogrammetric community.

This paper deals first with target recognition using a binary image and target location using grey scale images. The aim is the detection of the coordinates of the target object with respect to the sensor. The second area which is considered is the consistent labelling of the targets with respect to a reference image and subsequent images which contain views of the same target. The method proposed uses the best features of each method for an optimal solution. Research into binary and grey scale target location have been performed elsewhere (Mikhail, 1984; Rosenfield, 1988). However, in this paper both are considered along with the labelling problems for multiple views.

A novel feature of this investigation is the use of the Bundle Adjustment technique as a means of comparing the accuracy of target location methods. The aim of the investigation is the automatic 3-D measurement of a section of a wind turbine blade. The research results indicate that these methods can efficiently: recognise the targets, locate them to subpixel accuracy, and properly label the them in the multiple images to produce a 3-D range map of these targets.

The study uses a PC-AT computer, a Pcvision Framestore Board, and some general purpose image processing software. The image size and resolution is 512x512x8 Bits. All the software is written in the C language except the interface software, which is written in Assembly language.

## 2. TARGET IDENTIFICATION IN A BINARY IMAGE.

The detection of legitimate targets can be described as an object shape recognition problem - the targets are round. Segmentation of the image into 0 = black or 1 = white ( black for parts of the image below a specific intensity level and white for all those on or above) has often been successfully used to solve these problems in machine vision applications (Haralick, 1985). The advantages are: compact images, ease of processing, and potential for hardware solutions. The disadvantages are that in some circumstances this method is likely to prove unreliable because of: variable surface reflectivity, non-ideal illumination, the possibility of occlusion, the variation of target size, and many false targets; i.e. good contrast is needed across the whole image. Furthermore, distortion and deformation of targets by the imaging process and the subsequent digital processing can have an influence on the measurement accuracy. The targets chosen are required to be distinguishable from the background of the object under investigation. Consequently, either black or white diffusely reflecting targets are commonly used. In this case because the background was of a light colour, black targets were chosen.

### 2.1 Local Image Normalisation.

The grey levels of the background of the object are seldom constant over the whole image, hence, segmentation of the image will often give non-ideal results. There are many possible solutions to this problem such as: building a mathematical model of the background image, high pass filtering, dividing the image into sections so that each subimage is processed separately, or performing a Fourier Transform of the image, removing the low frequency components and doing an inverse Fourier Transform. Each method has its own advantages and disadvantages.

In practice, using prior knowledge of both the targets and the structure being measured can assist in the choice of detection algorithms. In the case of this application there are many equal sized targets and the background intensity changes slowly and provides good contrast between the dark targets and the light background. To reduce the problem of uneven illumination, as shown in Figure 2, affecting the binary segmentation; a local area of the image is considered for normalisation of the background intensity level.

The targets were found to occupy a range of pixel sizes in the image, from 3X3 to 5X5, therefore the 512X512 image was divided into 32X32 subimages. The mean of all the intensity values was calculated for the whole image and the subimage. The subimage was then normalised by using Equation 1.

$$IM[i,j] = IM[i,j] + (i\_mean - l\_mean) + C \qquad (1)$$

Where **C** is a constant, **i_mean**, and **l_mean** are the mean intensity values for the whole and subimage respectively, and IM[i,j] is the subimage array.

The advantage of this method is to provide reliable thresholding in spite of background intensity variations. The value of **C** can be altered to adjust the image background to any desired level. Figure 1. shows the original image and Figure 2, shows the inverse intensity profile of the section of the image at the position of the line at the top of the image.

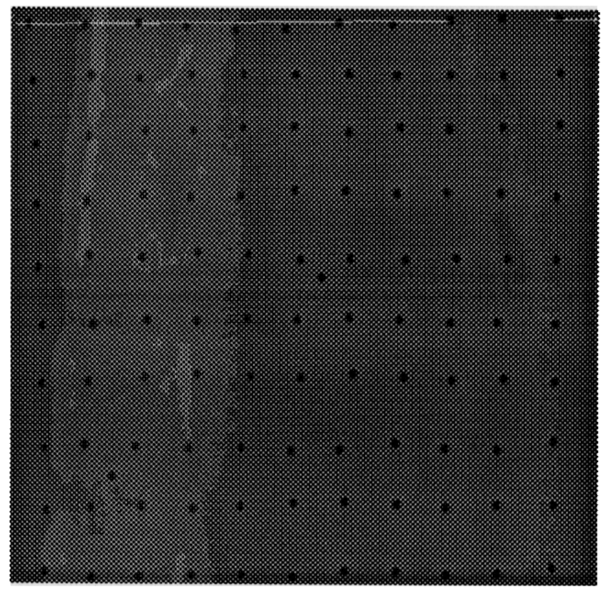

Figure 1. Original image.

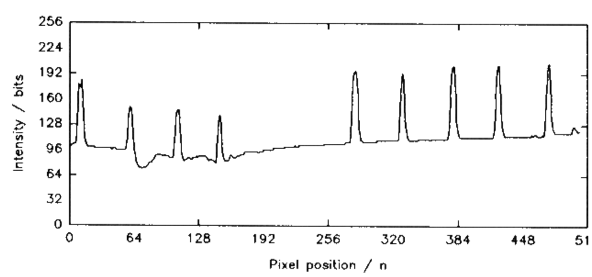


Figure 2. Intensity profile of marked section.

If the whole image were segmented above the highest background level, then some of the targets would be of reduced size, or in the worse cases, not visible at all. Figure 3, shows the result of using the image normalisation, where it can be seen that segmentation will give good definition of the targets.

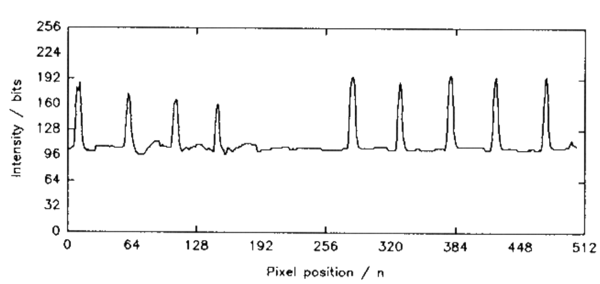


Figure 3. Results of local image normalisation.

Having performed the image normalisation the segmentation of the image is possible using a single threshold value, as shown in Figure 4.

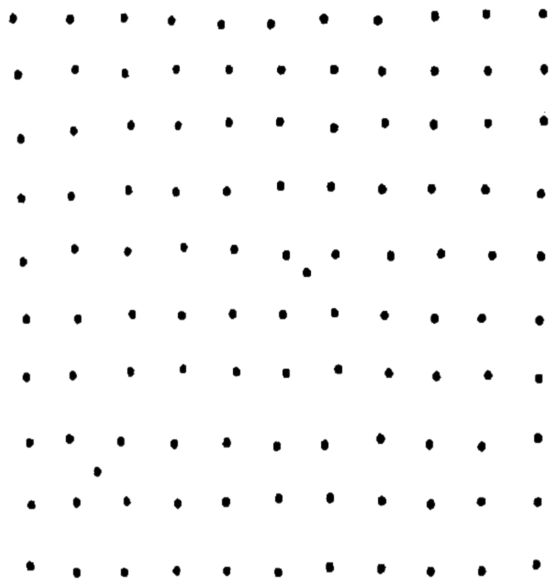

Figure 4. Segmentation of original image.

The improvements gained by using this method are only of benefit for initial target recognition as the possibility exists of a target appearing partially in more than one subimage and hence, being distorted.

**2.2 Detection and recognition of targets.**

The image normalisation discussed in the previous section compensates for the non-uniform background illumination and reflectivity. After image segmentation the process of searching for and recognising targets is performed. A single threshold value is chosen to create the binary image. In the case of image detection and recognition the prior knowledge of the size, shape and possible orientation of the targets is used. The process divides into two steps: (i) contour tracing of object, and (ii) extraction of a structure parameter, which includes area, perimeter and circle factor, to decide on the validity of targets.

2.2.1 Contour tracing of targets. It is necessary to trace the contour of all objects which appear in the binary image. There are many methods can be used for this purpose such as the chain code method (Pavlidis, 1982). These are not discussed here as they are well understood and documented techniques. In the case of this study, the target appears as a black circular blob. The X,Y coordinates of the traced contour are extracted for use in analysis of the shape of the object.

2.2.2 Extraction of a structure parameter and recognition of targets. When an image is segmented there may be many objects other than the legitimate targets, so that it is necessary to find a suitable method to distinguish between targets and non-targets. Typical features which can be used are (i) perimeter length, (ii) size, and (iii) shape.

(i) Perimeter. The perimeter length of the subject can be calculated using the traced contour **X,Y** coordinates.

(ii) Area. The area can be calculated by counting all of the pixels inside and on the perimeter of the subject.

(iii) Shape. A shape factor is used to express the differences between circular subjects and non-circular. The definition of the shape factor is given in Equation 2.

$$Q = A/[\pi .( L / 2 )^2 ] \qquad (2)$$

Where **A** is the area of the object, **L** is the longest distance across the object. The equation gives the ratio of the area of the subject to the area of a circle which circumscribes the subject. The nearer to a circle the object is, the closer to 1 the ratio is.The variation in the distance between pixels which are connected in the x,y and the pixels connected in diagonal directions has to be compensated for. Commercial cameras have differing scale factors in vertical and horizontal (a typical ratio for an industrial camera is 4:3 ). So there must also be an adjustment for this factor.

These three factors, described by Equations 3,4, & 5, allow the building of a decision function which is able to establish the likelihood of a given subject being a target. All parameters used are expressed in relative values for convenience.

$$\text{Area\_factor} = a1.ABS(( A - AA ) / AA ) \qquad (3)$$

$$\text{Perimeter\_factor} = a2.ABS(( P - PP ) / PP ) \qquad (4)$$

$$\text{Circle\_factor} = a3.(1 - Q) \qquad (5)$$

Where:

a1,a2, and a3 are risk weight coefficients
A is the estimated ideal target area
P is the estimated ideal target perimeter
Q is the circle factor
AA is the actual area of the object
PP is the perimeter of the object

When all of these factors are within some predetermined bounds there is a high probability of this object being a target so the object coordinates are stored, otherwise the object is rejected. This process is repeated for all of the image until all of the objects are recognised as targets or rejected. Figure 5, shows the final results of the recognition process.

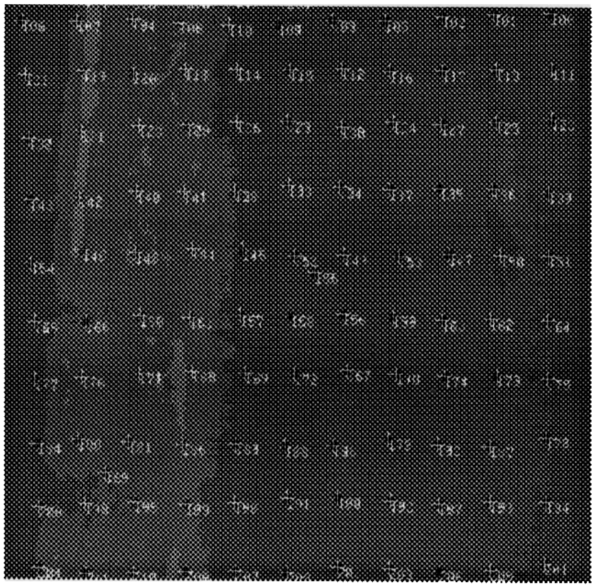

Figure 5. The final results of target recognition.

A number is given to each of the targets in the order in which it is recognised and a cross placed at that position. The use of a normalised images and binary segmentation allows for efficient recognition of the targets. While other methods using a grey scale image may be contemplated most will be more computationally expensive.

## 3. TARGET LOCATION IN A GREY SCALE IMAGE.

Grey scale processing uses the image intensity values at each pixel site directly. In this case the image blurring that inevitably occurs in the imaging process can be used to match not only shape but expected intensity distributions. Cross sections of a small round target will have a Gaussian shaped intensity distribution. It is possible to perform image matching in grey scale but this method will be computationally more expensive. Hence, the grey scale image is only used to provide the precision location of the target to subpixel accuracy.

Once the legitimate targets have been identified it is then necessary to refine the precision with which they are located in the image. A survey (West & Clarke, 1990) of subpixel techniques using grey scale images indicates that high location accuracies have been reported by a number of authors for objects such as edges, Gaussian blobs, patterns, etc. A number have demonstrated that the centre of the targets used in photogrammetry are no exceptions (Trinder, 1989; Wong, 1986). The subpixel accuracy with which a target can be located using these methods has been reported as high as 0.01 of a pixel. Accuracy is limited to a large extent by noise (Deng,1987) if linejitter (Beyer, 1990) is not a problem.

Many methods have been used for subpixel location of targets, among them are interpolation, correlation, centroiding, differential, and shape fitting. In this study the centroiding method was chosen for investigation because the computation will give consistent results on small circular shaped targets even when viewed from differing angles.

The initial positions of the targets given by the binary image are used to place a small rectangular window around the target. A suitable threshold was used to eliminate the background from the target. A 15x15 pixel window was found to be a suitable size. From this window the precise centre of the target was calculated using the well known Equation 6.

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

$$y = 1/M \sum_{i=1}^{n} \sum_{j=1}^{m} i . g_{ij} \qquad (6)$$

Where

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

$g_{ij}$ is the grey scale value of each pixel and $n = m = 15$.

The results of the target coordinates for each image were then stored in a file for subsequent labelling with respect to the reference image. Following the labelling, described in the next section, the bundle adjustment procedure (Granshaw, 1980) was used to calculate the coordinates of the targets. The results gave a global variance factor for the procedure of 1.578 which could be equated to a subpixel accuracy of 0.13 of a pixel.

A problem has been reported using the centroid method (Trinder, 1989) because of the influence of the background. This will not affect targets which have a uniform background illumination, but will cause some asymmetry in the image of targets with a nonuniform background. This effect is illustrated by viewing a small section of the profile of the image shown in Figure 1. Although this is just a single slice of the target image it demonstrates the problem.

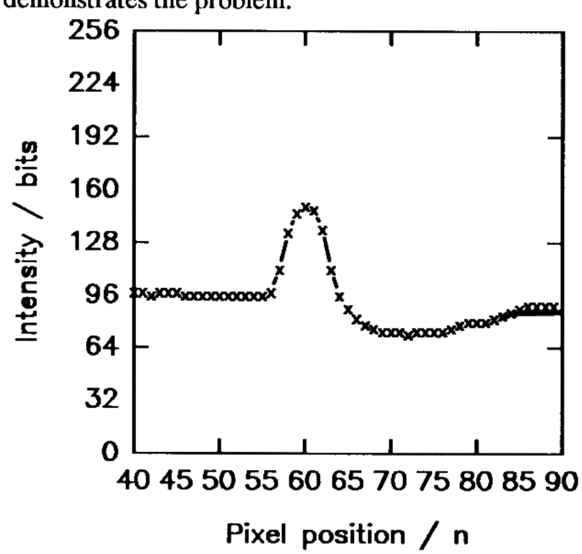


Figure 6. Section of target image shown in Figure 1.

A weighting factor was introduced as suggested by Trinder, 1989, changing Equation 6 into Equation 7.

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

$$y = 1/M \sum_{i=1}^{n} \sum_{j=1}^{m} i \cdot g_{ij} \cdot w_{ij} \qquad (7)$$

Where

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

$g_{ij}$ is the grey scale value of each pixel and $n = m = 15$, and $w_{ij} = g_{ij}$.

In this case the higher intensity values of the target are given a greater weight in the calculation so that the influence of the background is decreased. Further tests were performed using this equation instead of Equation 6, with unaltered: data sets, network, camera calibration, method of target location, labelling, and bundle adjustment procedure. This resulted in an improvement over the simple centroid method of 25% giving an overall subpixel accuracy of 0.11 of a pixel.

The benefit of the approach adopted is that an independent measure of accuracy achieved is used to assess the two methods. Furthermore, this is possible in a measurement situation with real problems of variable illumination, target orientation, and target distance. The results of these tests show that the location of the targets using these methods produces reasonable accuracy. A few targets exhibited residuals of up to one pixel in the bundle adjustment. These larger residuals occurred at the same image positions for each method but were not examined further as the primary purpose was to compare the two target location methods.

## 4. UNIQUE LABELLING OF TARGETS.

It is has been shown that the coordinates of legitimate targets can be extracted with high reliability and accuracy. These coordinates provide the basic information required for the bundle adjustment program to calculate the 3-D coordinates of the targets. However, although it is possible to identify and locate the targets from each camera station, the differing camera orientations mean that: the subject may be distorted, some of the targets possibly occluded, or targets may be out of the field of view. Therefore, it is necessary for the targets from each view to be uniquely identified with respect to each other.

Ideally, the locations of some or all of the labels of the targets could be mathematically modelled in 3-D space and a transformation performed for each varying camera location. However, this presupposes just the information which is the end product of the whole measurement process. Unfortunately many objects are complex to model accurately and so an approximation may be a better approach. Another method may be to use uniquely shaped targets. However, this has serious implications for the imaging of these targets because for unique identification it is likely that the targets would need to be larger than the small circular targets used and would also be non symmetric under translation. Fortunately the wind vane under consideration approximates to a flat surface and an affine transformation can be performed. See Figure 7.

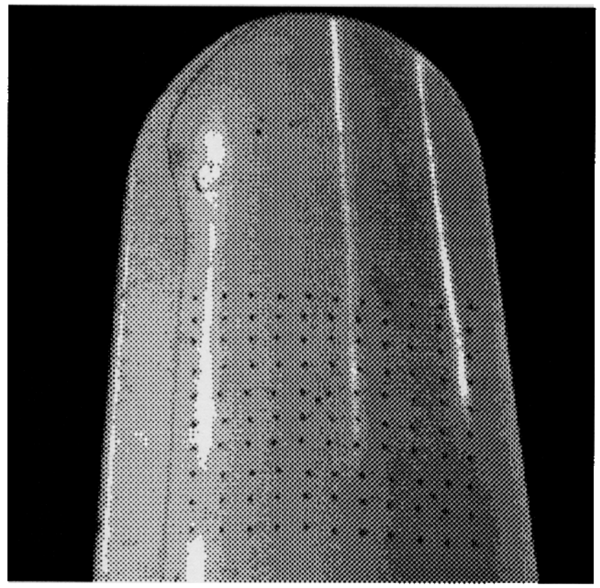

Figure 7. Image of the tip of the turbine blade.

### 4.1 Choice of control points.

The choice of parameters for the transformation of the image data to the same orientation is performed by choosing at least three, generally four, known control points which can be uniquely identified in each image, say, the corners. Then by performing the affine transformation the image is warped. These control points need to be positioned to minimise the distortion caused by the transformation of the 2-D image of a real 3-D object. It was found in practice the choice of targets at or near to the corners gave the best results.

### 4.2. Principle of the transformation.

The basic principle of the transformation is that one of the images is considered to be the reference image and the other images are transformed to match its orientation. If the number of control points identified on the master image is **m**, then the equations of a polynomial are:

$$X = T_x(X_n, Y_n)$$
$$Y = T_y(X_n, Y_n) \qquad (8)$$

$T_x()$, and $T_y()$ are single mapping functions. Because the position and orientation of the cameras are arbitrarily placed, the mapping relationship of $T_x()$ and $T_y()$ has to be approximated. In this case, where a linear affine transformation is sufficient, then:

$$X_i = a_0 + a_1 X_n^i + a_2 Y_n^i$$
$$Y_i = b_0 + b_1 X_n^i + b_2 Y_n^i \quad (9)$$

$(i = 1,2,....,m)$

$X_i$ and $Y_i$ represent the coordinates of the $i^{th}$ control point in the reference image. $X_n^i$ and $Y_n^i$ represent the coordinates of the $i^{th}$ control point in the image which are to be transformed. Since there are **m** control points, then m must be greater or equal to three. Hence, using Equation 9, and the usual matrix notation:

$$X = A.a \quad (10)$$

$$Y = A.b \quad (11)$$

Where:

$$X^T = [X_1, X_2,...X_m]$$

$$Y^T = [Y_1, Y_2,...Y_m]$$

$$a^T = [a_0, a_1, a_2]$$

$$b^T = [b_0, b_1, b_2]$$

$$A = \begin{bmatrix} 1 & X_n^1 & Y_n^1 \\ 1 & X_n^2 & Y_n^2 \\ \vdots & \vdots & \vdots \\ 1 & X_n^m & Y_n^m \end{bmatrix}$$

If the Equations 10 and 11 are solved separately then the coefficients $a_i,b_i$ are obtained. For $\mathbf{m} = 3$, the equations 10 and 11 are valid and:

$$a = A^{-1}X \quad (12)$$
$$b = A^{-1}Y \quad (13)$$

If $|A|$ is not equal to zero, then $|A^{-1}|$ exists.

When **m** is greater than 3 the solution to Equations 10, and 11 are expressed as the least squares estimate of parameters **a** and **b**.

$$a = (A^TA)^{-1}(A^TX) \quad (14)$$
$$b = (A^TA)^{-1}(A^TY) \quad (15)$$

Using the methods discussed all images from the differing camera stations were transformed into the same coordinate system. However, because of the approximations in the transformation to a new 512x512 image these images can only be used for target labelling. Targets which are very close together could cause problems to the raster scanning labelling system. In the case of conflict a test is used to measure which target in the warped image is closest to the reference image. When all of the targets in the warped image have been mapped to the reference image the reference image labelling system is applied. Figure 8, shows the location or the targets in from one of non-reference camera stations.

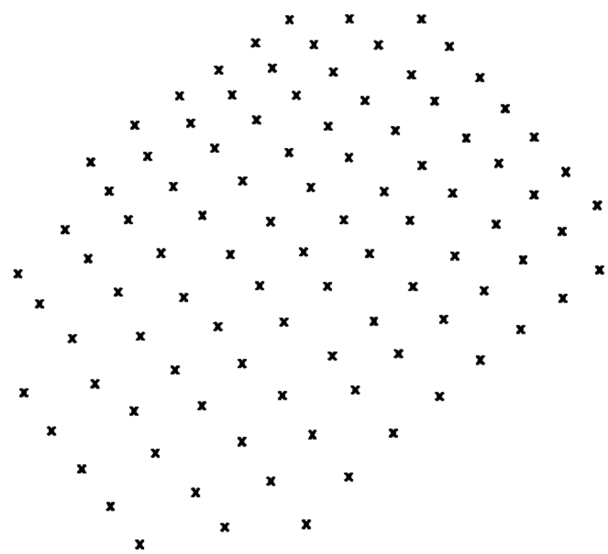

Figure 8. Target locations from camera station 2.

The results of the transform and are shown in Figure 9. by the circular symbols and the reference image shown by the triangular symbols.

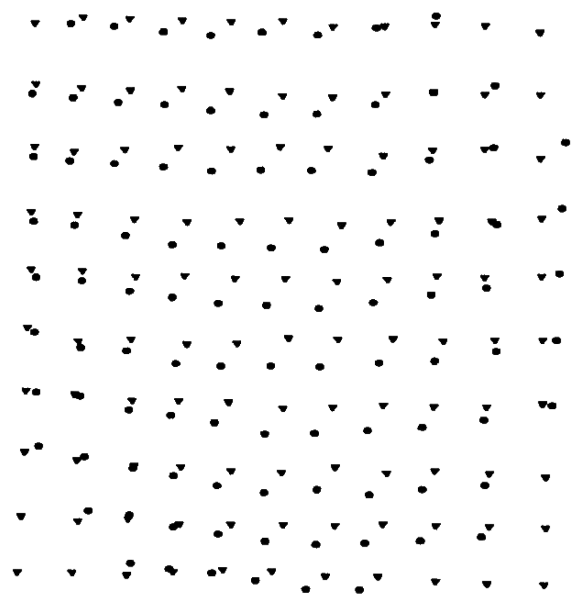

Figure 9. Results of target transformation.

It was found that in the case of the simple geometry of the turbine blade under test the targets could be uniquely identified in each of the images according to the numbering scheme of the reference image.

## 5. CONCLUSION

Using the methods described, several tests were performed in order to measure the 3-D coordinates of the targets placed on a turbine blade of a wind power generator. There were 10 x 11 targets placed on the blade. Five images were collected in each group as shown in Figure 10.

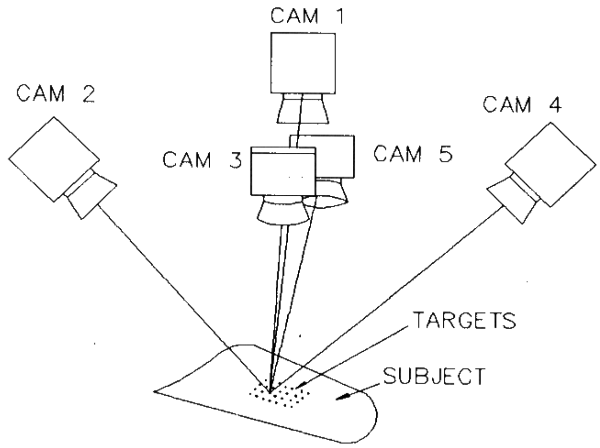


Figure 10. The position of the camera stations.

Following the recognition, location, and labelling process the coordinates of the targets from each camera station were passed to a software package called the General Adjustment Program (GAP) developed at City University, and the resulting 3-D data were used to construct a wire frame view of the turbine blade surface, see Figure 11.

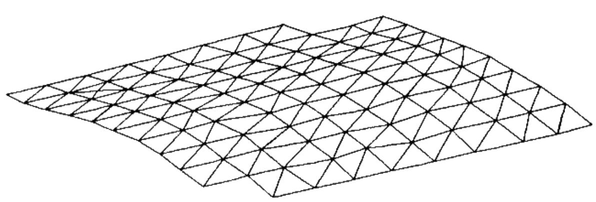

Figure 11. A 3-D view of a section of the turbine.

The research indicates that in some situations where there is prior knowledge of the subject to be measured automatic target recognition is possible. Furthermore, the use of subpixel techniques can enhance the accuracy of CCD cameras beyond the nominal accuracy suggested by sensor size and numbers of pixel. Hence, such image processing hardware and software methods may overcome some of the disadvantages of manual target identification and location which is both time consuming and tedious.

The investigations described in this paper used an essentially simple structure which allowed the use of straightforward strategies to achieve good results. However, in most cases, more complex structures with non-ideal imaging characteristics are required to be measured. This research has allowed the authors to become acquainted with some photogrammetric methodology and has given insight into directions for further work. Such work should address four main areas:

(i) The target recognition, location and labelling method could be seamlessly integrated into the bundle adjustment method.
(ii) The use of transformations based on central perspective projection e.g. epipolar, to extend the labelling process beyond 2-D object assumptions.
(iii) The target location algorithms may be further refined to give greater accuracy, reliability, or efficiency.
(iv) The binary method could be dispensed with if alternative methods could be devised for use in grey scale images.

Further research which will build upon the multi-disciplinary approach is underway. It is hoped that a greater understanding of photogrammetric methodologies will result in a useful contribution to machine vision problems.

## 6. ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of a number of colleagues: Prof. K.T.V. Grattan, Dr. S. Robson, Dr. T.J. Ellis, and the provision of the software package "GAP" which was developed in the Photogrammetry Unit in the Civil Engineering Dept, City University.

## 7. REFERENCES

Bethel, D.J., 1990. Digital image processing in photogrammetry, Photgrammetric record, B(76), October, pp. 493-505.

Beyer, H.A., 1990, Linejitter and geometric calibration of CCD cameras, ISPRS Journal of Photogrammetry and Remote sensing, Vol 45, pp. 17-32.

Bhanu, B., 1986, Automatic Target Recognition : state of the art survey, IEEE Transactions on Aerospace and Electronic Systems, Vol AES 22, No 4, pp. 364-379.

Deng, T., 1987, Accuracy of position estimation by centroid, Intelligent Robots and Computer Vision, SPIE Vol 848, pp. 141-150.

Granshaw, S.I., 1980, Bundle adjustment methods in Engineering Photogrammetry, Photogrammetric record, 10(56), pp. 181-207.

Gruen, A.W., 1989. Digital Photogrammetric Processing Systems: Current Status and Prospects, Photogrammetric Engineering and Remote Sensing, Vol.55, No.5.,pp.581-586.

Haralick, R.M. & Shapiro, L.G., 1985, Survey : Image segmentation techniques, Computer Vision, Graphics, and Image processing, Vol 29, pp. 100-132.

Mikhail, E.M., et al, 1984, Detection and subpixel location of Photogrammetric targets in digital images, Photogrammetria, Vol 39, pp. 63- 83.

Pavlidis, T., 1982, Algorithms for graphics and image processing, Published by Springer-Verlag, Berlin, 416 pages.

Rosenfield, A. & Cher, A.C., 1988, Detection and delineation of compact objects using intensity pyramids, Pattern Recognition, Vol 21, No 2, pp. 147-151.

Trinder, J.C., 1989. Precision of Digital Target Location, Photogrammetric Engineering and Remote Sensing, Vol.55, No.6, pp.883-886,

West, G.A.W., & Clarke. T.A., 1990, A survey and examination of subpixel measurement techniques, ISPRS Conf. on Close range photogrammetry and machine vision, SPIE Vol. 1395, pp. 456-463.

West, G.A.W. et al, 1988, Image processing and understanding for automatic inspection, Trans. Inst. M.C. Vol. 10, No. 5., pp. 265-272.

Wong, K.W. & Ho, W., 1986, Close range mapping with a solid state camera, Photogrammetric engineering and remote sensing, Vol 52, No 1, pp. 67-74.