An estimator for the random error in subpixel target location and its use in the bundle adjustment.

T.A. Clarke, M.A.R. Cooper, & J.G. Fryer.

Abstract.

A CCD camera and frame store provide image data that are affected by a number of imperfect processes. However, most of these can be quantified or estimated. For instance, the quantisation process of allocating integer values from 0 to 255 for the intensity level of each pixel has well known statistical properties. This paper discusses how these can lead to an estimate of the precision of location of target image centroids. Two centroid algorithms are analysed. The theory is tested against experimental and simulated data.

1. Introduction.

A typical photogrammetric method of measuring the dimensions of an object is to place targets at sites of interest and to use their resulting 2-D target image co-ordinates from multiple camera views to calculate the 3-D co-ordinates. The targets may be retro-reflective, black on a white background, white on a black background, or any distinctive colour which stands out against the background. The shape of targets varies from crosses to circular patches. This paper concentrates on the use of circular targets which are easy to place on the subject, may be imaged on as few as nine pixels, and whose estimated location will remain the same provided the viewpoint is not too oblique and given a uniform background and illumination.

In close range photogrammetric 3-D measurement using digital cameras, the subpixel location of images of targets is highly important. Typically, a target image will occupy a region of 5x5 pixels to 20x20 pixels, and will have an approximately Gaussian shaped intensity distribution. This intensity distribution will, in practice, be influenced by: the background, the distance of the target from the camera, and the angle of inclination of the target to the camera. Such variations will influence the accuracy with which the target image can be located. A number of algorithms can be used to compute the locations of these target images to subpixel precision, for example, centroid methods and least squares template matching (West, and Clarke, 1990). The least squares method, which is computationally expensive, is able to indicate the "goodness of fit" between the target and an ideal template. Centroid methods, which are simple to calculate, do not provide such information. In this paper a method of estimating the precision of the positions of centroids of target images computations is proposed, based on information derived from the target image itself.

The 2-D locations of target images, acquired from multiple views, are used in a bundle adjustment to calculate the 3-D co-ordinates of the targets on the object. The subpixel precision of the location algorithm has a direct influence on the overall precision that can be achieved by the measuring system. In general, all the derived co-ordinates of the target images are weighted equally in a bundle adjustment. However, it is sometimes necessary to change the weight of some of these co-ordinates relative to others. Because the least squares minimisation process distributes errors amongst all the observables, any intuitive procedure

which attempts to reduce the weight of specific observations is liable to error and can be highly time consuming if physical reasons are sought to explain errors. Although in a well designed configuration, changes in a priori weights of the target image co-ordinates do not seriously affect the estimated co-ordinates of the targets on the object, correct a priori weights are necessary if proper statistical properties are to be associated with those object target co-ordinates. In this paper a random error estimator is used to assign weights to the co-ordinates of the centroid of each target image in the bundle adjustment. An analysis is performed of the results with and without the use of the error estimator.

2. Definitions of the centroid of a target image.

One of the most common, and intuitively acceptable definitions of the centre of a target image is the position of its centroid. The position can be defined by co-ordinates relative to an arbitrary origin and arbitrary axes, but for convenience an origin and orthogonal axes are taken to coincide with a corner and the edges of the two dimensional array of pixels. To define the position of the centroid of a target image within the array, it is convenient to take a rectangular shaped subset of pixels (or window) which encloses the target image (Trinder, 1989). Then a local co-ordinate system with origin at a corner of the window and orthogonal axes parallel to its sides is convenient as a basis for calculating the position of the centroid of the target image. Figure 1 illustrates the arrangement. The window has m pixels in the x direction and n in the y direction.

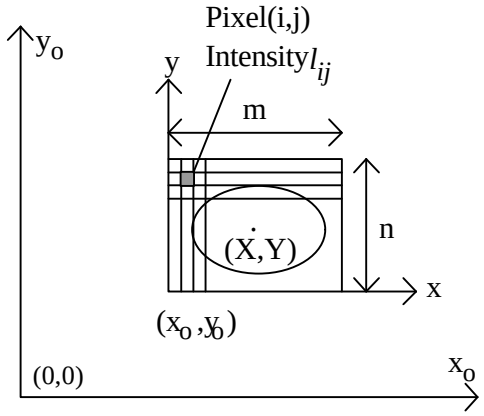

Figure 1. Co-ordinate axes.

If, for pixel (i,j) in the window the intensity level is l_{ij}, then the co-ordinates of the centroid are, if the factor at each pixel is taken to be proportional to the intensity level l_{ij}:

$$X = \frac{\sum_{j=1}^{n}\sum_{i=1}^{m} j.l_{ij}}{\sum_{j=1}^{n}\sum_{i=1}^{m} l_{ij}} \quad \text{and} \quad Y = \frac{\sum_{j=1}^{n}\sum_{i=1}^{m} i.l_{ij}}{\sum_{j=1}^{n}\sum_{i=1}^{m} l_{ij}} \qquad (1)$$

The coordinates (X,Y) in (1) are relative to the local axes; the co-ordinates of the centroid in the image co-ordinate system are (x_o+X, y_o+Y). This method for defining the centroid position is often used because it has low computational cost, it produces acceptable results for target images having different sizes and intensity levels, and it is simple to implement. It does however, have some disadvantages in practice. The extent of the target image may be difficult to define and enclose by a rectangular window. If the intensity level of the background is significant, a threshold value can be chosen and all values less than that are made zero in the computation (Wong & Ho, 1986). If a second target image lies close to the first, the window could contain unwanted intensity values which can not be readily excluded without setting a

background threshold so high that it reduces the number of intensity levels to the point where they become inadequate for precise centroid location. Target images situated with their centres along a diagonal of the rectangular window are most likely to cause mutual interference. This is illustrated using a section of a real CCD image where two targets are close together (Table 1). A square window of 13x13 pixels centred about the centre of the lower image is shown surrounded by the bold lines. Figure 2 is a bar chart of the italicised intensity levels in Table 1 which shows two alternative positions of a threshold, each of which will give an incorrect centroid position for the target image in the middle of the window.

255	255	240	95	7	5	3	3	2	1	1	0	0	0	0
232	255	255	197	36	8	5	3	2	1	1	0	0	0	0
91	255	255	190	37	7	5	4	2	1	1	0	0	0	0
14	232	170	76	18	5	3	3	2	2	1	1	1	1	0
5	***91***	25	13	6	4	2	1	1	1	0	0	0	0	0
3	14	***6***	4	3	3	3	2	1	0	0	0	0	0	0
3	5	3	***3***	3	2	2	3	1	1	0	0	0	0	0
4	3	3	2	***3***	3	7	11	7	2	0	0	0	0	0
3	3	2	2	2	***6***	26	61	48	14	3	1	0	0	0
2	4	3	3	4	15	***81***	142	122	46	6	0	0	0	0
2	3	2	2	4	25	125	***181***	163	72	13	1	0	0	0
2	2	3	1	4	25	118	176	***165***	73	13	0	0	0	0
2	2	1	2	1	12	73	135	126	***46***	8	1	0	0	0
1	2	1	1	1	4	27	60	49	17	***3***	1	0	0	0
2	2	1	1	0	1	1	3	8	6	3	***0***	0	0	0
1	1	0	1	1	1	2	2	2	2	0	0	***1***	0	0
1	2	1	0	1	0	2	0	0	1	0	0	0	***0***	0

Table 1. Intensity values taken from a real CCD image.

An alternative method for defining the centroid of a target image has been developed which avoids the above problem of an adjacent target image in a rectangular window. An arbitrary pixel within the target image is chosen and a recursive fill algorithm (Pavlidis, 1982) is used to visit all pixels within the image which have intensity levels above the threshold value. A series of vectors at 90° to each other are defined and, depending on the identified starting location within a specific target image, a search sequence is performed (Figure 3).

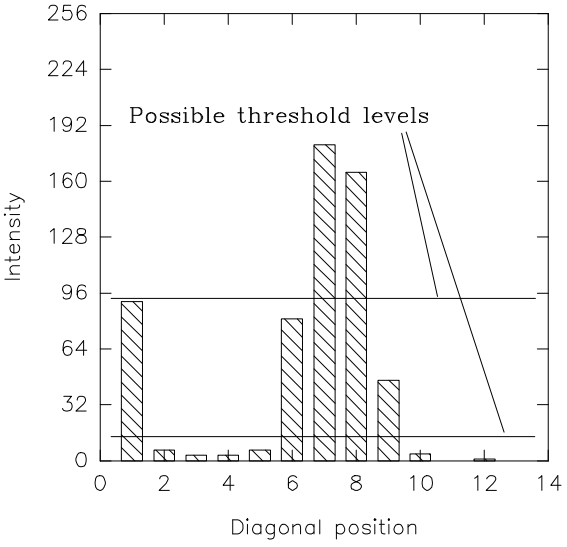

Figure 2. The choice of threshold will influence the centroid location.

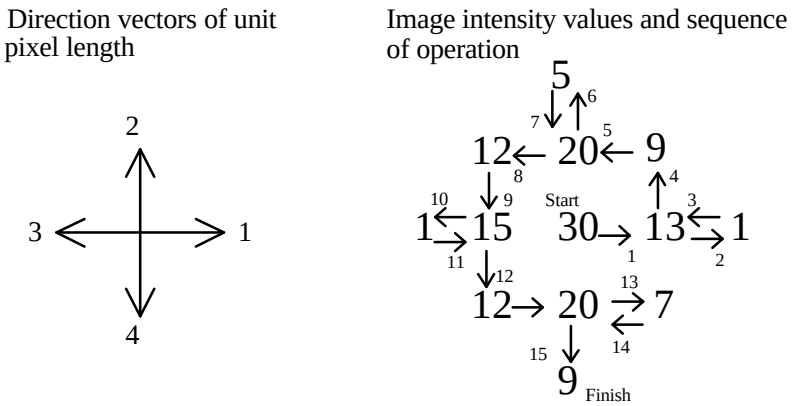

Figure 3. Direction vectors, example target image, and directions moved to visit all pixels.

The algorithm becomes less efficient with larger targets where alternative methods (Pavlidis, 1982) may be applied, but for the sizes of targets used here, the extra computation was not considered important. The pixel at which the fill algorithm starts is taken as the origin of the local co-ordinate system for defining the centroid position. For each pixel which is included in the target image, its co-ordinates (x_i,y_i) and intensity level l_i are recorded. The centroid of the target image is defined by local co-ordinates

$$X = \frac{\sum_{i=1}^{m} x_i . l_i}{\sum_{i=1}^{m} l_i} \quad \text{and} \quad Y = \frac{\sum_{i=1}^{m} y_i . l_i}{\sum_{i=1}^{m} l_i} \qquad (2)$$

where m is the total number of pixels in the target image. The co-ordinates in the image co-ordinate system (Figure 1) are (x_o+X,y_o+Y) where (x_o,y_o) are the image co-ordinates of the initial pixel used in the fill algorithm.

Another centroid defintion (West & Clarke, 1990) was also investigated in this study where the intensity values are squared. This has the effect of giving more weight to the higher intensity levels which, it may be argued, are more precisely measured and, in a real image, less affected by background intensity variations. This centroid could not be expected to be correct when a small image covering a few pixels is used, because under these circumstances the calculated centre will be incorrectly biassed toward the highest intensity level, rather than to the centre of a group of high intensity levels. The centroid co-ordinates in this case are given by (1), or (2) with l replaced by l^2.

3. The derivation of error estimators.

Digital representation of an image consists of intensities at discrete pixels. Each intensity level contains errors due to processes of converting light to a voltage level (Chamberlain & Broughton, 1986) and this voltage to a digital value (Loriferne, 1982). Many errors can be reduced, for example in the design and implementation of the analogue to digital (A-D) converter in a framestore, but the error caused by rounding the intensity levels to integer values cannot be avoided. Hence, to investigate the precision of centroid algorithms only unavoidable quantisation errors are considered here.

3.1 A-D conversion.

It is assumed for the present that the output from the camera reaches the framestore error-free, and is sampled at the correct rate. There are four main sources of error in A-D converters. (a) Quantisation. Transitions in the digital output of the converter take place at voltage input values which are given by $(m + 1/2).(V_{ref}/2^n)$, where n is the number of bits of the A-D converter, and m is an integer in the range 0 to n-1. (b) Offset error. This is defined as the difference between the voltage input to give a '1' output and the theoretical 0.5 of the smallest quantisation level. (c) Gain error. When the digital output reaches a maximum before the input voltage has reached the reference voltage this is called a gain error. (d) Linearity error. Various descriptions are used for this term. It is the difference between the voltage level at which a digital transition should occur, and that at which it does occur. The last three errors (b,c, & d) are all temperature-dependent. Although compensation can be made, it is generally assumed by the chip manufacturer that an A-D converter will operate at a set temperature, and may exhibit varying degrees of these unavoidable errors at other temperatures.

3.2 Statistics of the quantisation process.

The A-D converter produces, for each pixel, an intensity level which is an integer formed by rounding the analogue signal value up or down. If this rounding, or quantisation process introduces a random error e, the probablilty density function (pdf) y = f(e) is shown in Figure 4.

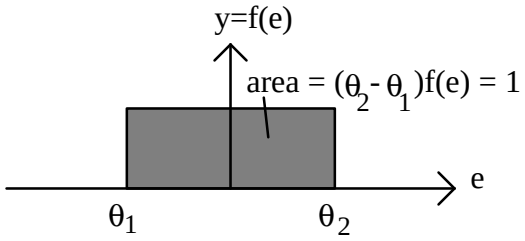

Figure 4. The uniform pdf.

The pdf is $f(e) = (\theta_2 - \theta_1)$ for $\theta_1 \le e \le \theta_2$ and $f(e) = 0$ for all other values of e.

Here, θ_1 and θ_2 are the minimum and maximum errors introduced by the rounding process, i.e. -0.5 and 0.5 respectively. The expection and variance of e are found by integration from $-\infty$ to $+\infty$ of $y.f(e)dy$ and of $y^2.f(e)dy - \mu^2$ respectively to be

$$\mu = (\theta_2 - \theta_1)/2 = 0 \text{ and } \sigma^2 = (\theta_2 - \theta_1)^2/12 = 1/12 \quad (3)$$

respectively. A quantised intensity level l is given by $l = l_o + e$, where l_o is the original, unquantised, value. Assuming that l_o is free from error (section 3, first paragraph) the expectation and variance of l are l_o and 1/12, respectively.

3.3 Covariance matrices of centroid co-ordinates.

The functions (1), (2), and (2) with l replaced by l^2, form the basis for deriving the (2x2) covariance matrix of centroid co-ordinates. For any vector $\mathbf{v}$ of random variables $v_1, v_2, ... v_k$ which is a function $\mathbf{f}(\mathbf{u})$ of random variables $u_1, u_2, ...u_l$ which have the (lxl) covariance matrix $\mathbf{C}_l$, the covariance matrix $\mathbf{C}_v$ is given by

$$\mathbf{C}_v = \mathbf{J}_{vu}\mathbf{C}_u\mathbf{J}^t_{vu} \quad (4)$$

(Cooper, 1987) where $\mathbf{J}_{vu}$ is the Jacobian matrix for $\mathbf{v} = \mathbf{f}(\mathbf{u})$. In the case of function (2) for example,

$\mathbf{v} = [X \quad Y]^t$ and $\mathbf{u} = [l_1 \quad l_2 \quad ....l_m]^t$

Assuming that the intensity levels l_i are independent and uncorrelated, each having variance σ^2, application of (4) to (1) gives, if the functions can be regarded as continuous,

$$s_X^2 = \frac{s^2}{(\sum l_i)^2}.(\sum x_i^2 - 2X\sum x_i + mX^2) \qquad (5.1)$$

$$s_Y^2 = \frac{s^2}{(\sum l_i)^2}.(\sum y_i^2 - 2Y\sum y_i + mY^2) \qquad (5.2)$$

$$s_{XY} = \frac{s^2}{(\sum l_i)^2}.(\sum x_i y_i - X\sum y_i - Y\sum x_i + mXY) \qquad (5.3)$$

with $\sigma^2 = 1/12$. The summations are over the m pixels which have been found to be within the threshold value for the target image whose centroid is required. Similar expressions can be found for the cases when the centroid is defined by using l^2 instead of l as the weighting function, or when it is defined by (1).

4. Trials with simulated data.

A program was written in 'C' and compiled on a Sun SPARC computer to generate synthetic images with a known Gaussian shape and maximum intensity

$$I = \text{round}(P.(\exp(-x^2/2.\omega^2).\exp(-y^2/2.\omega^2)) \qquad (6)$$

where I is the quantised intensity, P the peak intensity, (x,y) the position to calculate, and ω the standard deviation of the Gaussian shaped intensity function.

The centres of the simulated target images were created at a number of x,y, subpixel locations within a 1x1 pixel grid. At each location the discrepancy between the known centroid position and the centroid position computed from (2), and from (2) with l^2 replacing l, were stored with their variances and covariances. The RMS values of these discrepancies were finally computed with the mean of the standard deviations for each method. The simulation was repeated for a variety of image sizes and maximum intensity values shown in Table 2.

Intensity	4	8	16	32	64	128	256	512	1024	2048	4096
ω	0.5	0.6	0.7	0.8	0.9	1.0	1.2	1.4	1.6	1.8	2.0	2.5	3.0

Table 2. The sizes and maximum intensity values used in the trials.

In all, 13x11 trials of 10,000 locations were computed using both the l and l^2 as the weighting function for the centroid. The results for simulated target images with $\omega = 2$ (target images about 12 pixels in diameter) are shown in Table 3.

	With l as weighting function		With l^2 as weighting function	
Intensity peak	RMS centroid discrepancies in x	Mean σ_x	RMS centroid discrepancies in x	Mean σ_x
4	0.0566	0.0462	0.0497	0.0569
8	0.0368	0.0294	0.0253	0.0288
16	0.0186	0.0180	0.0139	0.0144
32	0.0119	0.0106	0.00646	0.0072
64	0.00651	0.00620	0.00354	0.00360
128	0.00368	0.00353	0.00175	0.00180
256	0.00252	0.00199	0.000897	0.000900
512	0.00135	0.00110	0.000450	0.000450
1024	0.000655	0.000606	0.000227	0.000225
2048	0.000380	0.000331	0.000109	0.000112
4096	0.000204	0.000179	0.000056	0.000056

Table 3. Results of the simulation trials of the error predictor for $\omega = 2$.

For a given maximum intensity level, little improvement in precision was obtained by increasing the target image size beyond about 5 pixels in diameter. The standard deviation of a centroid co-ordinate is approximately inversely proportional to the maximum intensity level. The simulated data compared with the computed data using (2) and (5) were in good agreement except for target images as small as 2 or 3 pixels in diameter. In these cases the relatively few discrete feasible positions for a centroid invalidated the use of (4) in deriving (5).

5. Use of the error estimator.

A test field was constructed from a flat panel 445x730mm in size which was spray painted black and targeted with 77 retro-reflective targets of three sizes: 2.0, 3.5, and 5mm. The targets were arranged to provide a range of images sizes when viewed from different directions. Five images of the test field were obtained using a Pulnix TM6CN CCD camera. The camera views and lighting were arranged to ensure a large variation of target peak intensities and target sizes across the field in each image. The camera axes varied from 90º to 20º relative to the plane of the test field giving a multistation convergent configuration. Hence, a wide range of centroid target location errors was expected.

For each image all of the target centroid locations were computed using a pre-processing stage to recognise the targets, followed by calculations of the target centroid co-ordinates using both l and l^2 as the weighting function. Four homologous points were manually identified in each of the five images to obtain good starting values for the exterior orientation parameters of the five cameras, followed by an iterative bundle adjustment matching procedure to correctly label the targets (Chen, Clarke, & Robson, 1993). The results of the bundle adjustment with minimum constraints using *a priori* weights based on (5) showed the presence of random errors in the image co-ordinates an order of magnitude greater than those from quantisation intensity levels. The evidence for this was an *a posteriori* variance factor of 169 and all measurement residuals of the order of ten times their *a priori* values. It is expected

that the causes of these larger errors can be identified and the errors reduced or eliminated by future work.

6. Conclusions.

A method of computing the centroid of a target which allows the separation of targets which are too close for reliable discrimination has been developed and tested. This method also minimises the unwanted effect of background illumination. A method of deriving random error estimators for target image centroid co-ordinates has been developed using information available from each image. For each of the centroid algorithms discussed, good agreement with simulated data tests was found provided the statistics of the method were applicable to the target image sizes. The characteristics of the errors due to quantisation were analysed and useful rules developed which should assist in the setting up of digital close range 3-D measuring systems. The problem of identifying and reducing other errors was not resolved in the course of the work for this paper, but it is hoped that this can be addressed in further work where a detailed assement of other aspects such as the effects of changing the threshold, or background illumination can take place. It is worth noting that ***if*** the random errors can be reduced to the order of magnitude of those arising from quantisation, precision of 3-D co-ordinates of targets on the object ***can*** be obtained to the order of 1 part in $2x10^5$.

7. References.

Chamberlain, S.G. & Broughton, J.H., 1986, "Technology progress and trends in solid-state silicon image sensors.", In: Proc. IEEE Conf. Custom Integrated Circuits, Rochester, New York, 9 pages.

Chen, J., Clarke, T.A., & Robson. S., "An alternative to the epipolar method for automatically finding corresponding target images in multiple viewpoint 3-D measurement". In: Proc. Optical 3-D Measurment Techniques, Zurich, 1993, in press.

Cooper, M.A.R., 1987, "Control Surveys in Civil Engineering.", BSP Professional Books, Oxford, 381 pages.

Loriferne, B., 1982, "Analog-Digital and Digital-Analog Conversion", Heyden & Son Ltd, London, 196 pages.

Pavlidis, T. 1982, "Algorithms for Graphics and Image Processing", pp. 167-193, Pub. Springer-Verlag, Berlin, 416 pages.

Trinder, J.C. 1989, "Precision of digital target location.", Photogrammetric Engineering & Remote Sensing, Vol. 55, No 6, June, pp. 883-886.

West, G.A.W, & Clarke, T.A., 1990, "A survey and examination of subpixel measurement techniques.", ISPRS Int. Conf. on Close Range Photogrammetry and Machine Vision, SPIE Vol. 1395, pp 456 - 463, Sept. 3-7.

Wong, K.W. & Ho, 1986, W, "Close range mapping with a solid state samera", Photogrammetric Engineering & Remote Sensing, Vol. 52, No 1, January, pp.67-74.

PAPER REFERENCE

Clarke, T.A. Cooper, M.A.R. & Fryer, J.G., 1993. An estimator for the random error in subpixel target location and its use in the bundle adjustment. Optical 3-D measurements techniques II, Pub. Wichmann, Karlsruhe:161-168.