Multi-sensor system self-calibration

Article in Proceedings of SPIE - The International Society for Optical Engineering

DOI: 10.1117/12.220890

CITATIONS
62

READS
655

3 authors, including:

Clive Fraser
University of Melbourne
240 PUBLICATIONS 7,085 CITATIONS

SEE PROFILE

Mark R Shortis
RMIT University
183 PUBLICATIONS 4,763 CITATIONS

SEE PROFILE

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Estimation of the Koala distribution on French Island using RPAS and LandSat images. View project

Video based automatic fish species classification View project

Multi-sensor system self-calibration

C. S. Fraser, M. R. Shortis and G. Ganci

Department of Geomatics, University of Melbourne Parkville 3052, AUSTRALIA

Telephone: +61 3 9344 6806 Facsimile: +61 3 9347 2916 Email: Clive Fraser@mac.unimelb.edu.au

ABSTRACT

This paper will address the topic of multi-sensor self-calibration in vision metrology systems employing still-video imagery. Under this approach greater fidelity is afforded in the recovery of sensor interior orientation and distortion parameters. Moreover, a combination of wide- and narrow-angle imaging sensors affords a better insight into the problems of in-plane image distortion and out-of-plane sensor surface deformation, which are often overlooked sources of error in metric applications of CCD cameras. This on-site calibration approach is demonstrated in the measurement, to better than 1:100,000 accuracy, of a 5m x 2.5m bond tool used in aircraft manufacturing.

Keywords: Calibration, CCD cameras, Vision Metrology, Accuracy, 3D Measurements

1. INTRODUCTION

In a modern single-sensor videometrics or vision metrology system for industrial measurement, object space positional accuracies surpassing 1:100,000 of the principal dimension of the object are now routinely attainable with large-area CCD cameras and photogrammetric data processing³,13,19,20. To achieve such optical triangulation precision it is critical that the geometric relationship between the image space of the CCD sensor and the object space be comprehensively understood. Central to this understanding is the aspect of system calibration and, more specifically, system self-calibration.

Self-calibration is well known to photogrammetrists, since it has been accepted for over two decades as the most versatile and rigorous means to perform a general perspective transformation from 2D image to 3D object space coordinates. In controlled tests, self-calibration has yielded up to 10-fold improvements over conventional calibration approaches for close-range CCD-based camera systems^{2,14}. All parameters determined by the photogrammetric triangulation, namely the sensor interior and exterior orientation, the XYZ object point coordinates and the essential calibration data, are recovered without the need for reference to external geometric constraints such as targeted test ranges of known dimensions. The accuracy of photogrammetric triangulation is fundamentally a function of, firstly, the directional measurement resolution of the CCD camera and, secondly, the geometry and number of the intersecting bundles of rays (one per image) forming the optical triangulation network^{8,18}. The angular measurement resolution is in turn a function of three variables: the accuracy of 2D image coordinate mensuration (1%-4% of a pixel for well defined targets), the focal length of the camera lens, and the fidelity of the mathematical model of calibration which describes the deviations of the physical imaging process from a geometrically ideal perspective projection.

Optimal self calibration relies on the provision of both geometric strength and a high level of observational redundancy in the imaging network. In industrial photogrammetric networks employing a single large-format film camera, triangulation accuracies exceeding 1:500,000 have been realised in highly convergent networks of 20 or more images, with an associated angular measurement accuracy of about one second of arc^{9,17}. Such resolution requires that image coordinate mensuration be carried out to 1 µm precision. In order to yield triangulation accuracies of 1:100,000 or better with large area CCD cameras such as the Kodak Megaplus and DCS420 series, multi-station convergent networks and angular resolution to 1-2 arc seconds are again required. Yet to yield such accuracy, xy image coordinate data needs to be determined to a precision of decimals of a micrometre in view of the shorter focal lengths involved. For example, to produce 1 arc second angular accuracy in a Kodak DCS420 still video camera with 9 µm square pixels and a lens focal length of 28 mm, an image mensuration accuracy of 0.15 µm, or better than 0.02 pixel is required. Since virtually all calibration errors manifest themselves as perturbations in image location on the image plane, accurate geometric calibration is of greater relative significance in high-precision vision metrology than in film-based photogrammetry.

Thus it is worthwhile to examine closely some of the traditional calibration models which have long served film-based photogrammetry well, but which have to withstand closer scrutiny in the presence of image coordinate mensuration to $0.1~\mu m$ precision.

As has been mentioned, both triangulation precision and the accuracy of recovery of system calibration parameters are enhanced through the provision of greater observational redundancy, obtained simply from more bundles of imaging rays. This additional data need not come from a single CCD camera, but instead from additional sensors. The concept of multi-sensor self-calibration is well established in film-based photogrammetry^{6,7,11} even though it is not frequently called upon in practise. In this paper, the authors investigate the calibration problem for large area CCD cameras through the approach of multi-sensor self-calibration whereby image data from six photogrammetric networks comprising a common object point field, but effectively six different cameras, is utilised. The six "different" CCD cameras actually comprise two bodies, a DCS200 and a DCS420, and three different lenses of 28 mm, 20mm and 15mm focal length. Thus, an examination can be made of the behaviour of parameters common to both individual lenses and particular CCD sensors.

In the following sections we first review the self-calibration process for digital close-range cameras. The algorithms are given in sufficient detail to clearly define the concepts for the reader who is not a calibration specialist. The test measurement phase is then overviewed and the general triangulation results summarised. The accuracy levels attained assist interpretation of the calibration results produced, which illustrate some hitherto unreported characteristics and so provide an interesting insight in to the mathematical fidelity of the "standard" photogrammetric calibration model, at least as far as the two DCS still video cameras and three lenses are concerned.

2. SELF-CALIBRATION

2.1 Mathematical background

In any discussion of selfcalibration it is important to keep in mind a common thread that binds all photogrammetric calibration approaches, but is not necessarily adhered to in the machine and computer vision communities. Put simply, this could be said to be the comprehensive geometric (and radiometric where appropriate) modelling of the image space of the camera system, independent of the target scene. The resulting sensor system calibration is then without regard to the object space and is largely independent of the geometric configuration of the camera stations employed for the particular 3D measurement task at hand. This accounts for the calibration models employed in close-range photogrammetry being expressed solely in terms of image space coordinates.

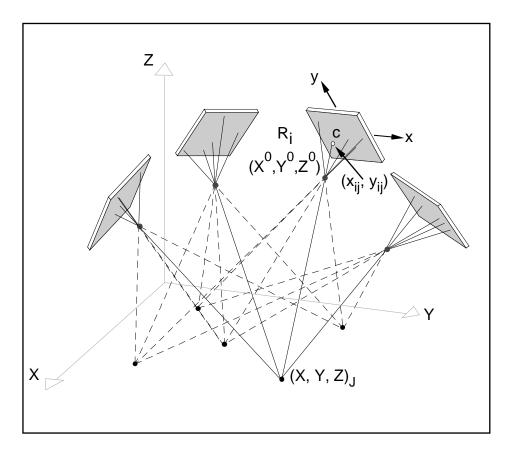


Figure 1. The concept of bundle triangulation

The essence of the photogrammetric bundle adjustment is illustrated in Figure 1, where an array of targets is imaged in a convergent network of four sensor stations. Through the process of optical triangulation, XYZ object space coordinates are obtained from the spatial intersection of homologous rays, the individual directions of which are determined by the position (X^0, Y^0, Z^0) and attitude (rotation matrix R) of the camera perspective centre and the xy image coordinates of the point at which the ray intersects the focal plane. To solve for the relative orientation of the individual bundles of rays only photogrammetric image coordinate data is needed. Although exterior orientation parameters relate to an object space XYZ coordinate system, this reference system can be arbitrarily assigned. The fact that the bundle triangulation is concerned with the reconstruction of 3D shape, independent (if necessary) from scale, position of orientation of the XYZ datum, makes it more imperative that calibration parameters relate only to the image space.

The mathematical model of the self-calibrating bundle adjustment is based on the well-known collinearity condition which is implicit in the perspective transformation between image and object space:

$$x - x_0 + \Delta x = -c \frac{R_1}{R_3}$$

$$y - y_0 + \Delta y = -c \frac{R_2}{R_3}$$
(1)

where

$$\begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix} = R \begin{pmatrix} X - X^0 \\ Y - Y^0 \\ Z - Z^0 \end{pmatrix}$$

The 'calibration terms' in the equation are represented by the principal point offsets x_0 , y_0 , the principal distance, c (the interior orientation parameters) and the perturbation terms Δx and Δy which account for the departures from collinearity due to lens distortion and in-plane and out-of-plane focal plane distortion. The purpose of the present paper is to concentrate on the recovery of calibration parameters through the simultaneous solution of the collinearity equation, Eq. 1. Although reference will be made to the precision of recovery of triangulated XYZ object point coordinates, no further discussion will be made concerning the mechanics of the bundle adjustment other than to refer to specific network requirements concerning multi-sensor self-calibration. Readers requiring further information on the topic of digital camera self-calibration are referred to the wealth of literature on the subject 1,10,14,25 .

In seeking appropriate parameters for the functions Δx and Δy it is necessary to consider the four principal sources of departures from collinearity which are physical in nature. These are symmetric radial distortion, decentring distortion, image plane unflatness and in-plane image distortion. The net image displacement at any point will amount to the cumulative influence of each of these perturbations. Thus,

$$\Delta x = \Delta x_r + \Delta x_d + \Delta x_u + \Delta x_f$$

$$\Delta y = \Delta y_r + \Delta y_d + \Delta y_u + \Delta y_f$$
(2)

where the subscript r is for radial distortion, d for decentring distortion effects, u for out-of-plane unflatness influences and f for inplane image distortion. The relative magnitude of each of the four image coordinate perturbations depends very much on the nature of the camera system being employed. The presence of radial lens distortion is usually seen in the form of barrel distortion, decentring distortion is typically small in magnitude, unflatness effects in CCD cameras would arise through chip bowing or the "crinkling" of thin wafers, and in-plane distortion can be introduced through electronic influences such as clock synchronisation and rate errors, and long period effects from line jitter^{1,2} or time base correction²².

2.2 Lens distortion

Symmetric radial distortion in analytical photogrammetry is universally represented as an odd-ordered polynomial series, as a consequence of the nature of Seidel aberrations:

$$\Delta r = K_1 r^3 + K_2 r^5 + K_3 r^7 \tag{3}$$

where K_i are termed the coefficients of radial distortion and r is the radial distance from the principal point:

$$r^{2} = \overline{x}^{2} + \overline{y}^{2} = (x - x_{0})^{2} + (y - y_{0})^{2}$$
(4)

The necessary corrections to the xy image coordinates follow as $\Delta x_r = x \Delta r/r$ and $\Delta y_r = y \Delta r/r$. The K_1 term alone will usually suffice in medium accuracy CCD applications to account for the most commonly encountered third-order barrel distortion. Inclusion of the K_2 and K_3 terms is often warranted for higher accuracy applications and wide-angle lenses. In either case the decision as to the order of the correction function can be made on the basis of statistical tests of significance of the individual and generally highly correlated radial distortion coefficients. For the self-calibration computations carried out for this research, all three terms were carried in the distortion model for reasons of consistency and maximum model fidelity.

The distortion profile Δr associated with a particular principal distance value c is termed the Gaussian distortion profile. It is well known the radial lens distortion varies both with focussing and within the field of view. Variation with focus due to changing principal distance is of limited consequence in self-calibration so long as the camera is used at a fixed focus. The amount of variation of distortion with object distance is a function of the distortion gradient and this effect can be of metric significance for lenses exhibiting large radial distortion 12 . There is a projective coupling between the linear component of radial lens distortion and the principal distance which gives rise to an interesting and fortuitously beneficial feature in the self-calibration of selected CCD cameras. It is not uncommon for these cameras to utilise only a modest portion of the available field of view of the lens, as exemplified by the Kodak DCS series of still video cameras. Hence, any variation within the essentially paraxial, linear section of the distortion curve will be largely compensated by the projective coupling.

A lack of centring of lens elements along the optical axis gives rise to a second category of lens distortion which has metric consequences in analytical restitution, namely decentring distortion. The misalignment of lens components causes both radial and tangential image displacements which can be modelled by correction equations due to Brown⁴:

$$\Delta x_d = P_1(r^2 + 2\bar{x}^2) + 2P_2\bar{x}\bar{y} \Delta y_d = 2P_1\bar{x}\bar{y} + P_2(r^2 + 2\bar{y}^2)$$
 (5)

For a typical "quality" lens the magnitude of the decentring distortion rarely exceeds a few tens of micrometers at the extremities of the image format, and for a CCD camera of reduced image format is often much less. Decentring distortion also varies with focusing, but the resulting image coordinate perturbations are typically very small and the distortion variation is universally ignored in analytical photogrammetry. There is a strong projective coupling between the decentring distortion parameters P_I and P_2 and the principal point offsets x_0 and y_0 . This correlation has practical consequences in self-calibration for it means that to a significant extent decentring distortion effects can be compensated for by a shift in the principal point (and an effective tilting of the optical axis). The projective compensation can usually be anticipated with CCD cameras and hence a self-calibration may indicate that the lens be treated as if it were largely free of decentring distortion.

2.3 In-plane and out-of-plane distortion

Systematic image coordinate errors due to focal plane unflatness constitute a major factor limiting the accuracy of the photogrammetric triangulation process. The induced radial image displacement Δr_{tt} is a function of the incidence angle of the imaging ray. Thus, long focal length, narrow-angle lenses are much less influenced by out-of-plane image deformation than short focal length, wide-angle lenses. Unfortunately, many machine vision systems employ wide-angle lenses to achieve a workable field of view in cameras with CCD arrays of small format.

In high-precision industrial photogrammetry, the image plane topography of a metric film camera can be measured directly⁵. From a series of "spot heights" a polynomial surface model is computed, and the associated image coordinate corrections are then made according to the equations

$$\begin{cases} \Delta x_u \\ \Delta y_u \end{cases} = \begin{cases} \overline{x} / r \\ \overline{y} / r \end{cases} \sum_{i=0}^n \sum_{j=0}^i a_{ij} \overline{x}^{(i-j)} \overline{y}^{(j)}$$
 (6)

where a_{ij} are the coefficients of the polynomial surface model and the order is typically restricted to n=3 or 4, depending on the number and distribution of data points. The applicability of this approach to CCD matrix arrays is uncertain, for after installation within the camera, the CCD chip surface most often does not lend itself to direct surface contour measurement. Moreover, information regarding chip topography seems very difficult to come by, especially from manufacturers for whom 'flat' implies a much freer tolerance than the micrometre level sought by photogrammetrists. It is even conceivable that the CCD array may exhibit a degree of planarity that does not warrant any unflatness correction. However, focal plane unflatness should not be ignored; at an incidence angle of 45° a departure from planarity of 10 micrometers will give rise to an image displacement of the same magnitude. Moreover, the influence of unflatness is most insidious in that it invariably leads to significant accuracy degradation in the object space, without aggravating the magnitude of triangulation misclosures.

It has long been recognised that attempts to recover an image plane surface model from self-calibration are fraught with danger due to the problem of projective absorption of the induced error. Nevertheless, in instances of very strong network configurations there is the potential to model the error signal to some degree. In the present multi-sensor case this possibility is afforded through the combination of three lenses with each of two CCD cameras thus providing a means to better isolate the chip specific out-of-plane deformation effects.

In-plane distortion is distinguished from unflatness by the fact it takes place within the plane of the image. In film cameras, inplane distortion can arise from the processing and subsequent storage of film. The measured image coordinates can be analytically corrected for the effect by a model of the form

where, once again, n is typically 3 or 4. In the self-calibration of CCD cameras the statistically significant coefficients b_{ij} and c_{ij} are recovered and others are suppressed to alleviate the problems of overparamerization, which can lead to ill-conditioning and singularities in the least-squares bundle adjustment process. One encouraging feature of CCD arrays is the high positional integrity of the pixel elements. This would normally indicate that distortion is not a problem with CCD cameras, and indeed it appears not to be for digital CCDs in which A/D conversion occurs in the camera, and image data is output digitally (for example the Kodak Megaplus and DCS series, and the ProgRes 3000 from Kontron). The distortion which arises in an analog CCD camera is chiefly attributable to A/D conversion (especially to pixel clock non-synchronisation) and video signal transmission. Common symptoms are line jitter and errors in the scale of the horizontal scan, or x-axis, of the image. Recent research^{1,2} shows that while these distortions can be large enough to significantly influence measurement accuracies with analog CCD cameras, they can be rendered metrically insignificant through pixel synchronous A/D conversion and due attention to aspects such as camera warm-up and power supply fluctuations.

2.4 The additional parameter model

It should be clear from the terms comprising the correction Equations 6 and 7, for out-of-plane image deformation and inplane distortion, respectively, that their combination to form a single model for unflatness and distortion effects will lead to a duplication of terms, with resulting linear dependencies. If the correction coefficients are determined by independent means, this issue is of limited consequence. If, on the other hand, it is desired to recover the coefficients a_{ij} , b_{ij} and c_{ij} via self calibration, a number of problems can be anticipated. First and foremost, a successful modelling of unflatness and distortion effects through polynomial functions requires a well distributed array of points throughout the image. The denser the image point pattern the better the expected fidelity of the empirical correction functions. For machine vision networks, which typically involve a modest number of target points, experience suggests that the determination of the coefficients of both out-of-plane deformation and in-plane distortion via a direct solution of the collinearity equations should not be contemplated.

For this investigation, a variation of Eq. 6 has been employed to model focal plane unflatness, whereas the correction model for in-plane distortion has been reduced to two terms in the x-coordinate only, one to account for differential scaling between the horizontal and vertical pixel spacings, and one to model non-orthogonality between the x and y axes:

$$\Delta x_f = b_1 \overline{x} + b_2 \overline{y} \tag{8}$$

Upon collection of the terms comprising the interior orientation and the models for lens distortion and in-plane image distortion, the following additional parameter set is obtained for the self-calibrating bundle adjustment:

$$\Delta x = -x_0 - \frac{\overline{x}}{c} \Delta c + \overline{x} r^2 K_1 + \overline{x} r^4 K_2 + \overline{x} r^6 K_3 + (2\overline{x}^2 + r^2) P_1 + 2 P_2 \overline{x} \overline{y} + b_1 \overline{x} + b_2 \overline{y} + \Delta x_u$$

$$\Delta y = -y_0 - \frac{\overline{y}}{c} \Delta c + \overline{y} r^2 K_1 + \overline{y} r^4 K_2 + \overline{y} r^6 K_3 + 2 P_1 \overline{x} \overline{y} + (2\overline{y}^2 + r^2) P_2 + \Delta y_u$$
(9)

in which the terms Δx_u and Δy_u comprise, for any of the six combinations of lenses and camera bodies, only the statistically significant parameters a_{ij} forming the correction model, Eq. 6. The implication here is that all image distortion effects beyond axis affinity and non-orthogonality are attributable to unflatness of the CCD chip sensor surface. It should be noted that alternative functional models are available for image distortion, yet research has shown that most of the popular options yield similar results 16 .

In a multi-sensor self-calibration the calibration parameters x_0 , y_0 , c, K_1 , K_2 , K_3 , P_1 , P_2 , b_1 , b_2 and a_{ij} are carried for each of the six camera/lens combinations. Although the correction terms for in-plane distortion and out-of-plane deformation could be included as 'sensor-invariant' parameters they will be tied to specific lens/sensor combinations to enable an evaluation of the repeatability of their recovery. Under ideal circumstances, the radial lens distortion parameters should be very repeatable for the same lens and the in-plane and out-of-plane distortion parameters should also be repeatable for the same CCD matrix array. Interior orientation parameter values can be expected to change in accordance with variations in the lens mounting between different camera bodies. Decentring distortion parameters can also be expected to exhibit small variations due to the projective coupling with interior orientation elements and small changes in orientation of the focal plane with respect to the nominally perpendicular optical axis of the lens.

3. THE EXPERIMENTAL PROJECT

3.1 Network design and photogrammetric measurement

The experimental setting for this investigation was a surface contour measurement of the 5m x 2.5m bond tool for an aircraft undercarriage door shown in Figure 2. The network design adopted for the measurement was consistent with the requirements for a single-sensor vision metrology measurement employing a DCS420 still video camera fitted with a 28mm lens. The adopted network geometry is shown in Figure 3. Ten camera stations were employed in a convergent imaging configuration, with a mean camera-to-object distance of 7.5m, leading to an image scale of 1:270. Covariance analysis at the planning stage indicated that such a configuration could be expected to yield an object point triangulation accuracy (RMS 1-sigma) of about 50 µm (approximately 1:100,000) if one exposure per station was observed and an image coordinate mensuration accuracy of 0.3 µm or 0.03 pixels could be maintained for the array of 120 retroreflective targets. Two exposures per station would then yield 35 µm accuracy (1:145,000) and three exposures 30 µm accuracy (1:170,000) in the self-calibrating bundle adjustment. The influence of image plane unflatness diminishes as the field of view of the sensor decreases. Thus, of the three lenses considered the longest lens could be anticipated to be least effected by out-of-plane deformation effects. For this reason the 28mm networks were expected to have the greatest fidelity and comprised in each case 3 images per station or 30 images in total. The networks for the 15mm and 20mm lenses had the same relative geometry as the 28mm networks, with the camera set-back distance being reduced in each case so as to maintain the 1:270 imaging scale. These networks comprised only two exposures per station, primarily due to the finite limit of on-board storage of images for the cameras.

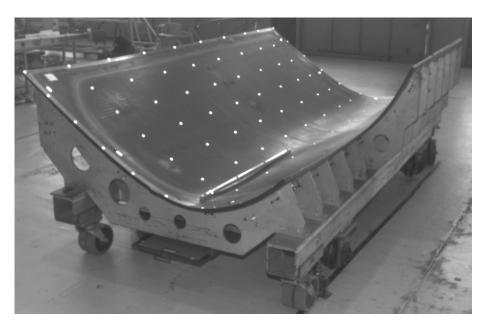


Figure 2. Bond tool prepared for surface contour measurement.

The image capture phase consumed about 2 hours, with each of the six networks being observed in about 15 minutes. The environment was unchanged prior to and during the photography, so it can be safely assumed that the temperature of the tool was stable and there were no dimensional changes. A total of 140 images were taken, with the cameras being hand held and illumination being provided by a flash unit of 200 watt-seconds power. The strobe light intensity was optimised by trial and error to minimise image blooming and read-out noise which has been experienced for the Kodak KAF sensors elsewhere²³. The targets spanned diameters of 22 mm and 6-10 pixels in the object and image spaces respectively.

Subsequent to both the semi-automatic measurement of the first half-dozen images and a preliminary bundle adjustment to determine refined approximations for the target point XYZ coordinates, all remaining image coordinate data was measured automatically via resection-driveback using a PC-based digital monocomparator employing intensity weighted centroiding²⁴.

To fully investigate the potential of multi-sensor self-calibration, both as a means to enhance triangulation precision and as an avenue to provide improved recovery of camera calibration parameters, a number of bundle adjustments were carried out. Of these, eleven are referred to in this analysis, namely the single-sensor self-calibrations for each network and five multi-sensor self-calibration adjustments, one for each lens and one each combining the 28mm and 20mm networks, and the 28mm and 15mm networks. Single-sensor self-calibration adjustments comprised 20 or 30 images, whereas 2-camera combined adjustments comprised 60 images for the 28mm data and 40 images for the 15mm and 20mm networks. The combined 4-camera self-calibrations comprised 100 images. To appreciate the geometric strength and redundancy involved in the multi-sensor self-calibrations it need only be recalled that spatial intersections involved close to 100 rays for some targets in two of the combined networks.

In recognition of the variation in influence of out-of-plane distortion between the lenses, different *a priori* precision was assigned to the observed image xy coordinate data for each lens. Prior standard error estimates of 0.3 µm, 0.4 µm and 0.5 µm were given to the 28mm, 20mm and 15mm lens data, respectively. While such differential weighting is of no consequence in both the single-sensor and 2-camera adjustments for each lens, it does mean that in the 4-camera self-calibrations involving two lenses, the 28 mm data carried nearly twice the weight of the 20mm and close to three times the weight of the 15mm data. Subsequent

variance component analysis indicated that the relative weighting was reasonably correct for the 28mm and 15mm lenses, but a little pessimistic for data from the 20mm lens.

3.2 Triangulation accuracy

Table 1 lists the measures of accuracy and precision obtained in both the six individual self-calibrating bundle adjustments and in the five combined adjustments. A number of interesting features are apparent from these results.

In the case of the 28mm lens, the two cameras yield very similar results, namely an RMS value of image coordinate misclosures of close to 0.03 and object pixels, point coordinate precision (RMS, 1sigma) of 0.033mm in XY and 0.026mm in Z. The standard error values are separated in this way for, as will be discussed marked later, a accuracy difference was found between planimetry and height in all Moreover, networks. the discrepancy was more anticipated pronounced than given the respective standard error values, a feature which is not uncommon, but still a little perplexing in vision metrology networks^{13,15}.

In the presence of satisfactory functional model fidelity, the triangulated object point coordinates should be free of significant systematic error. This can be simply assessed by comparing the level of XYZ coordinate discrepancies between the two 30-station, 28mm networks against the anticipated standard errors of the coordinate differences.

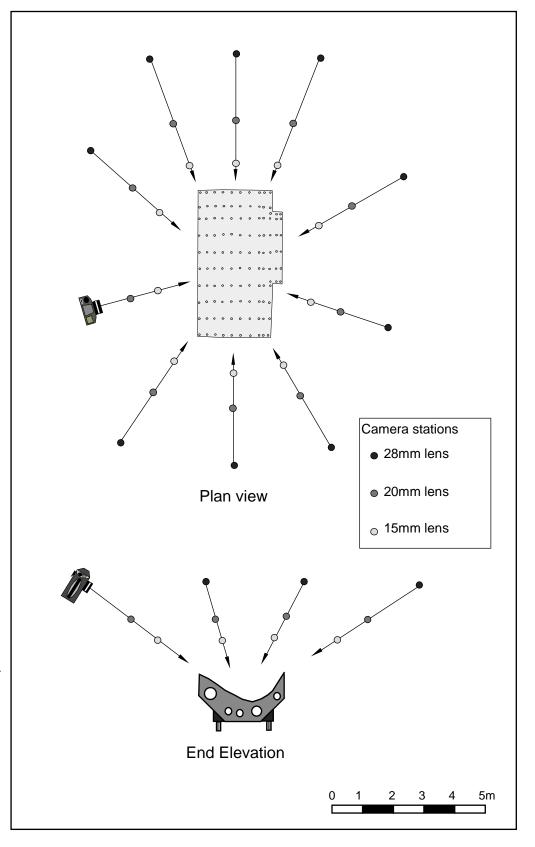


Figure 3. Network geometries for multi-sensor self-calibrations.

These figures are shown in parentheses in Table 1, where it can be seen that in planimetry the accuracy and precision measures are very consistent (ratio = 1.0), whereas in height the accuracy is significantly better than expected. It should be noted that the term "accuracy" is used instead of "repeatability", for although a common lens was employed, the two measurements are completely independent in all other respects. In the absence of "true" XYZ coordinate values, we will use comparisons against a "master" data set as being representative of the absolute accuracy. The master coordinate data was taken to be the 100-station combined adjustment of all images observed with the 28mm and 20mm lenses, from both cameras.

Network	Number	Image coordinate		Objec	Object point		Coordinate		Discrepancy/Associated	
	of	residuals		standard errors		discrepancies		standard error		
	Images	S_X	S_V	$\bar{\sigma}_{XY}$	$\bar{\sigma}_Z$	$\Delta_{ m XY}$	$\Delta_{ m Z}$	$\Delta_{XY}/\sigma_{\Lambda X}$	Δ_{Z}/σ_{AZ}	
			m)	1	m)		ım)			
<u>28mm</u>										
DCS200	30	0.29	0.28	0.033	0.026	0.030	0.018	0.8	0.6	
(200 vs 420)						(0.045)	(0.026)	(1.0)	(0.7)	
DCS420	30	0.29	0.27	0.033	0.026	0.031	0.018	0.9	0.6	
Combined	60	0.30	0.27	0.023	0.018	0.023	0.013	0.8	0.6	
20mm										
DCS200	20	0.27	0.26	0.035	0.030	0.080	0.043	2.4	1.3	
(200 vs 420)						(0.065)	(0.042)	(1.4)	(1.0)	
DCS420	20	0.24	0.23	0.032	0.028	0.050	0.044	1.4	1.4	
Combined	40	0.27	0.24	0.023	0.019	0.058	0.033	2.0	1.4	
<u>15mm</u>										
DCS200	20	0.48	0.40	0.053	0.054	0.077	0.058	1.4	1.0	
(200 vs 420)						(0.094)	(0.074)	(1.0)	(0.8)	
DCS420	20	0.48	0.61	0.076	0.079	0.109	0.079	1.4	1.0	
Combined	40	0.51	0.51	0.044	0.043	0.075	0.054	1.6	1.2	
28mm + 20mm										
28mm		0.29	0.27							
20mm		0.29	0.26							
Master set	100			0.016	0.013	-	-	-	-	
28mm + 15mm										
28mm		0.31	0.27							
15mm		0.50	0.51							
Combined	100			0.019	0.016	0.027	0.015	1.1	0.7	
	image coordinate misclosure			object space accuracy						
		$0.18 \mu m = 0.02 pixel$			0.025mm = $1:200,000$					
	$0.36 \mu \text{m} = 0.04 \text{pixel}$			0.050 mm = 1:100,000						
	$0.54 \mu \text{m} = 0.06 \text{pixel}$				0.075mm = 1: $67,000$					

Table 1. Summary of network adjustment results for the 28mm, 20 mm and 15mm lenses and for combinations. RMS values are shown, with the figures in parentheses being obtained from comparing the DCS200 and DCS420 networks for each lens.

The coordinate discrepancy results obtained for the data from the 28mm lens are quite consistent with the standard error values, which indicate an overall relative precision of 1:160,000 in the network for each camera and 1:230,000 for the combined network of 60 stations. Also shown in the table are the RMS XYZ coordinate discrepancies of each network compared to the master data set. Since the image coordinate observations for the 28mm networks carried the highest observational weight ($\sigma = 0.3$ µm), it is not surprising that there is good agreement with both the individual and combined 28mm networks.

In the two 20-station networks for the 20mm lens and in the combined network, similar levels of image coordinate misclosures were obtained, namely about 0.25 µm. The coordinate discrepancies between the networks for the DCS200 and DCS420 cameras was higher than in the 28mm case in terms of both absolute values and in relation to their associated standard errors. Thus, the 20mm lens yields smaller misclosures than the 28mm lens, but larger discrepancy values. It is noteworthy that the DCS420 produces significantly better results than the DCS200 for the 20mm lens, with the triangulation error field for the DCS200, illustrated in Figure 4a, showing a systematic shearing effect in XY. The corresponding plot for Z, Figure 4b shows only localised systematic trends. The results for the DCS420 showed a similar trend which was much less pronounced. In the combined network the discrepancies against the master data set were reduced compared to the DCS200, though the error vectors in XY continued to reflect the basic pattern shown in Figure 4a.

With the 15mm lens, the opposite accuracy situation to that for the 20mm lens was encountered. In this case, the DCS200 yielded significantly better absolute accuracy results than the DCS420, though the fall off was anticipated from the higher standard errors for the DCS420 which accompanied a significant rise in the misclosures of the y image coordinates. We attribute this to a number of instances of image blooming, smeared in the y direction due to the readout process, and radial image ghosting due to internal lens reflections in the DCS420/15mm network images. The XY coordinate discrepancies between the networks for the two cameras were in line with precision expectations, as reflected by the ratio values of 1.0 and 0.8, respectively. Once again, improved accuracy results were obtained in the combined adjustment compared to the DCS200 alone. Figure 5 shows the discrepancy vectors for the DCS420 with 15mm lens. Here, the main feature of the XY pattern is an apparent affinity in scale, with an expansion in X and a contraction in Y.

It is clear from Figures 4 and 5 that the largest discrepancies are at the periphery of the bond tool. It is widely understood that such a pattern of discrepancies is expected, given network geometries which consistently fill the field of view so that there is strong correlation between the periphery of the object and the extremes of the image. At the edges of the image there will be the greatest effect of image plane unflatness, the maximum gradient of the distortion profiles and the largest effects of in-plane distortions. Unflatness is a direct effect, whilst the latter two effects are indirect from uncertainties in the modelling caused by the predominance of data in the central portions of the images.

If we speculate for the moment that the main source of error in object point XYZ coordinates is attributable to image plane unflatness, then two principal consequences can be anticipated. The first is that triangulation accuracy will improve as the lens focal length increases. This is indeed the case, in terms of both the repeatability between the two networks with the same lens, and in the comparisons against the master data. The second is that the image coordinate misclosures for the 15mm and 20mm images in the combined adjustments with the 28mm networks should undergo little change from the single-sensor and 2-camera solutions due to projective absorption of the out-of-plane image deformation effects. To a large extent, this also holds true in the two 100-station combined adjustments.

A few additional observations regarding the results in Table 1 are noteworthy. In relation to the ratio of the coordinate discrepancies over their associated precision, the least accurate results were obtained for the 20mm lens, and yet the 20mm networks produced the smallest image coordinate misclosures. Also, in all cases absolute accuracy was enhanced when individual networks for each lens were combined, though in the case of the 15mm and 20mm lenses the improvement fell short of the level anticipated from the *a posteriori* coordinate standard errors, as the discrepancy ratios were larger than expected. To a large degree, however, the accuracy results obtained are consistent with the estimated standard errors obtained from the covariance matrix of triangulated object point coordinates, especially in the 'depth' coordinate.

4. CALIBRATION RESULTS

4.1 Interior orientation parameters

The individual values for the interior orientation parameters x_0 , y_0 and c obtained by the two 100-station multi-sensor self-calibration adjustments are listed in Table 2. Because of the physical nature of the lens/sensor coupling it is not possible to infer any level of "repeatability" of the parameter value for different networks. It is, however, quite feasible to assess the level of repeatability obtained for the interior orientation elements in the combined 2-camera self-calibration adjustments as opposed to the single-sensor cases. Such an examination has been carried out for the eleven bundle adjustments conducted. In all cases the maximum variations were 0.003mm or less for the three parameters, which is quite consistent with the listed standard error values.

Thus, in spite of the high level of projective compensation of the principal point offset with decentring distortion, very precise and repeatable values for x_0 and y_0 were obtained.

Lens	Sensor	x_0 σ_{x_0}	$y_0 \qquad \sigma_{y_0}$	c σ_c
28mm	DCS200	0.180 0.004	0.062 0.002	28.728 0.001
	DCS420	0.149 0.004	0.051 0.002	28.693 0.001
20mm	DCS200	0.107 0.003	0.059 0.001	20.426 0.001
	DCS420	0.050 0.003	0.057 0.001	20.403 0.001
15mm	DCS200	0.164 0.003	-0.113 0.001	15.282 0.001
	DCS420	0.082 0.003	-0.162 0.001	15.272 0.001

Table 2. Interior orientation parameters and standard errors obtained in the multi-sensor self-calibration adjustments of 100 images. Units are mm.

From the standpoint of traditional photogrammetric cameras, the magnitudes of the principal point offsets x_0 and y_0 are very large, amounting to maximum shifts of 20 pixels in x and 18 pixels in y away from the nominal centre of the CCD matrix array. While such large offsets are of limited consequence for analytical restitution models which incorporate corrections for the principal point coordinates, they would introduce significant triangulation errors in models which assumed that the optical axis of the lens passed through the centre of the CCD array. Note also the variation in the interior orientation values which highlights the necessity to self-calibrate every lens/CCD sensor combination.

4.2 Radial lens distortion

Shown in Figure 6 are the Gaussian radial distortion profiles obtained for the three lenses in the self-calibration adjustments. Only one curve is plotted for each of the 28mm and 20mm lenses because the level of repeatability was so high as to render indiscernible differences in the plotted profiles. At the maximum working radial distance of about 7.5mm, the maximum variations were less than 0.5 μ m for the 28mm lens and 0.8 μ m for the 20mm lens. At mid-field, at a radial distance of 4.5mm or less, discrepancies were less than 0.3 μ m, which is consistent with the image mensuration standard error of 0.03 pixel for these two lenses. Although the 15mm curves for the two cameras differ by less than 1μ m throughout most of the image format, the discrepancy does reach 4μ m at the extremity of the field of view.

These results reinforce the self-calibration experience that the recovery of the radial lens distortion profile is very repeatable, irrespective of the lens/CCD sensor combination. The distortion coefficients K_1 , K_2 and K_3 are generally highly correlated, usually indicating that most of the error signal is accounted for by the cubic term $K_1 r^3$. This coupling is restricted to the radial distortion parameters and is of little importance as far as the overall numerical stability of the self-calibration adjustment is concerned. The form of the plots shown in Figure 6 clearly illustrates the overwhelming dominance of the K_1 parameter. Yet while omission of the K_2 and K_3 terms rarely made an impact on the level of triangulation misclosure in the bundle adjustments conducted, a small degradation in object space coordinate accuracy was introduced when the fifth and seventh order distortion terms were suppressed. For high-accuracy vision metrology applications involving relatively dense object point arrays the inclusion of all coefficients in the additional parameter model is recommended. In instances where the object target array contains only a few tens of points and the network comprises relatively few images, say four or less, it may prove judicious to include only the K_1 term in the distortion correction model.

The radial nature of image coordinate errors introduced by out-of-plane image deformation can mean that the radial lens distortion model provides a partial compensation for low-order unflatness such as bowing. The high degree of repeatability in the distortion profile obtained for each lens indicates that either little measurable compensation occurred or the error signal was largely common to each CCD sensor. It is nevertheless interesting that the 15mm lens, which is subject to the most pronounced influence from image plan unflatness, showed the largest discrepancy between associated radial distortion profiles.

4.3 Decentring distortion

A useful means of representing the decentring distortion is via the profile function P(r) and the phase angle which are obtained via Conrady's model⁴ from the parameters P_1 and P_2 as follows:

$$P(r) = (P_1^2 + P_2^2)^{\frac{1}{2}} r^2 \tag{10}$$

$$\tan \phi_0 = \frac{-P_1}{P_2} \tag{11}$$

The maximum magnitudes for the radial and tangential components of decentring distortion are then obtained as 3P(r) and P(r), respectively.

Decentring distortion appears least pronounced in narrow-angle lenses, though this may have more to do with the projective absorption of linear components of the distortion signal by the interior orientation parameters x_0 , y_0 than to small distortion per se. In the network adjustments conducted for this project, correlation coefficient values of up to 0.98 were encountered between the parameters P_I and x_0 .

Notwithstanding this high level of projective coupling, very repeatable results were obtained for the decentring profiles for each lens. At a radial distance of 7.5mm, profile function values of 4.5 μ m, 1.4 μ m and 1.2 μ m were obtained for the 15mm, 20mm and 28mm lenses, respectively. The maximum discrepancy between the values for the DCS200 and DCS420 cameras for a given lens at this radial distance was 0.2 μ m for the 20mm lens, with the discrepancy being 0.1 μ m or less for the other two lenses. A reasonable degree of repeatability in the phase angle determination was also obtained.

Of the three lenses, only the 15mm lens displayed a decentring distortion profile that was clearly significant at the 95% confidence level and even with this lens the maximum value approached only half the pixel pitch. Nevertheless, in vision metrology applications which pursue the highest accuracy, decentring distortion cannot be ignored. The very repeatable results for P_1 and P_2 indicates that if any significant optical axis non-perpendicularity was present it was largely common to both cameras.

4.4 Affinity and non-orthogonality terms b₁ and b₂

The differential scaling and non-orthogonality terms b_1 and b_2 have proved beneficial both in the calibration of film cameras and in CCD camera self-calibration. In the CCD camera case, the affinity between the x and y axes most often arises through imprecise synchronisation between the sensor clock frequency and that of the frame grabber^{2,21}. With the digital signal transmission used in the DCS cameras, however, the source for the image shearing and differential scale variation between the x and y axes is in theory removed. The multi-sensor calibration technique affords an opportunity to investigate this issue further.

Lens	Network	Number	Affinity (b ₁)		Non-orthogonality (b ₂)		
		of images	DCS200	DCS420	DCS200	DCS420	
28mm	28	60	-0.18E-3	-0.15E-3	0.10E-3	0.50E-4	
	28+20	100	-0.18E-3	-0.14E-3	0.11E-3	0.44E-4	
	28+15	100	-0.17E-3	-0.15E-3	0.98E-4	0.53E-4	
20mm	20	40	-0.44E-4	-0.50E-5	0.33E-4	0.37E-4	
	28+20	100	-0.20E-4	-0.36E-4	0.88E-4	0.16E-4	
15mm	15	40	-0.24E-3	-0.24E-3	0.71E-4	0.11E-3	
	28+15	100	-0.23E-3	-0.25E-3	0.75E-4	0.10E-4	

Table 3. Values of affinity and non-orthogonality parameters obtained in the five multi-sensor self-calibrations

Shown in Table 3 are the values determined for b_1 and b_2 in the five combined self-calibration adjustments. Only in the 15mm and 28mm cases do the parameter estimates significantly depart from zero at the 95% confidence level. Lack of statistical

significance accounts for the poor repeatability between the values for the 20mm networks. The scale differential reaches 1.3 μ m and 1.8 μ m across the sensor for the 28mm and 15mm lenses, respectively, approaching a maximum of only 0.2 pixels. The shear effect is smaller in magnitude, giving rise to a maximum shift in the x-direction of 0.5 μ m at the edge of the format. The most noteworthy feature in the table is that for the 15mm and 28mm lenses there is a very strong correlation between the results for b_I between the two camera bodies, yet there is no visible correlation between the results for different lenses on the same camera body. This illustrates that the affinity term is modelling an error signal from the lens and not from the image plane. We must conclude therefore, firstly, that neither CCD chip displays any significant differential scaling between horizontal and vertical pixel spacings and, secondly, that the lens distortion terms fall short of providing complete functional model fidelity. A similar trend is seen for the non-orthogonality term, but here the correlation is not as strong. Once again, however, we see that there is no significant error signal common to each camera.

4.5 Focal plane unflatness

Within the polynomial correction model for out-of-plane distortion, only those terms of second to fourth order which were of statistical significance were retained in the final self-calibration adjustments conducted. First-order terms were omitted at the outset since these formed the model for in-plane distortion and in any event turned out to model an error signal generated primarily from the lens and not the CCD array. In each adjustment involving the 28mm and 20mm lenses, generally only two third-order terms were of significance, notably those in x^2y and xy^2 . For the 15mm lens, however, six terms were significant for both the DCS200 and DCS420, with the magnitude of the corrections Δx_u an Δy_u being substantially larger than those for the two longer focal length cases. This behaviour is consistent with the expectation that a more pronounced sensor surface unflatness influence will be seen for the wider angle lens.

In carrying image plane unflatness parameters in the multi-sensor self-calibration it was hoped that, in spite of the well known projective compensation effects of out-of-plane deformation, a modelling of the error would nevertheless be possible to some extent because of the very high geometric strength and high level of observational redundancy inherent in the combined network adjustments. Moreover, under ideal circumstances common trends would be seen in the derived error surfaces for each of the two CCD matrix arrays involved. To quantify the out-of-plane deformation, surface contours were determined for CCD sensor surfaces according to the following formula:

$$\Delta z = c \frac{\overline{\Delta r}}{(r + \overline{\Delta r})} \tag{12}$$

where Δz is the 'height' departure of the image surface from an ideal plane, and $\overline{\Delta r}$ is the radial component of the image coordinate correction obtained via Eq. 6.

The plots of focal plane topography determined in the two 100-station combined adjustments are shown in Figure 7, where it can be seen that the most conspicuous feature is the absence of any common trend for the DCS200 and DCS420 sensors. The results help to confirm the view that focal plane unflatness cannot be adequately modelled using analytical self-calibration techniques. When considered in isolation, any of the six plots in Figure 7 could be taken to plausibly represent 'true' CCD chip topography, but when all six are considered together no such conclusion can be drawn. Perhaps the only encouragement to be gained in this regard from the present investigation is that the level of out-of-plane image deformation of the two CCD sensors was not sufficient to invalidate the internal measures of triangulation precision, which are reasonably consistent with external accuracy indicators.

5. CONCLUDING REMARKS

From general considerations concerning the multi-sensor self-calibration of digital cameras, and from the results of the two-camera, three-lens experiment conducted, a number of observations can be made. Some offer further confirmation of findings from other reported investigations, while others provide a fresh insight into the powerful and versatile technique of the self-calibrating bundle adjustment applied to multiple CCD sensors.

The measures of internal precision and external accuracy were reasonably consistent. Combined multi-sensor networks always produced better accuracy, obviously through increased redundancy, but less obviously through an effective compensation

for remaining sensor-specific systematic errors. Accuracies well in excess of 1:100,000 were verified, with the precisions for the combined networks surpassing 1:200,000.

The combined networks illustrated that out-of-plane image deformation effects are absorbed into the parameter estimation without any significant influence on the triangulation misclosures. This result confirms that the signal is basically unrecoverable, a fact further indicated by the 20mm lens which produced smaller image misclosures but poorer accuracy than the 28mm lens. Some clear systematic trends remain in the patterns of discrepancy vectors in the object space, which indicates that the additional parameter model of self-calibration is still not providing perfect model fidelity. As was expected, in general accuracy increased with focal length, irrespective of the fact that the geometry and image scale was basically the same for all networks. Out-of-plane deformation of the images is the suspected source of these two phenomena.

Both the radial and decentring distortion profiles were very 'stable' and repeatable, in this case showing little influence from unflatness compensation in the radial case and little effect from correlation with the principal point location in the decentring case. The variability in distortion profiles obtained was very consistent with the level of observational error in the image coordinates, at least for the 20 and 28mm lenses. The affinity and non-orthogonality terms succeeded in modelling linear error signals, but these were common to the lenses and not the sensors. Thus, affine scaling and shear are not apparent for the CCD sensors in the DCS cameras. These linear distortion effects, which are small, may have arisen from variation of distortion within the photographic field.

The absence of common trends in the sensor surface topography mapping confirms the view that out-of-plane effects cannot be adequately recovered through self-calibration, even in very 'strong' networks obtained using camera station and object point fields with a wide variation in three dimensional geometry. This non-recoverability places the onus on users of vision metrology systems to evaluate or eliminate the expected accuracy degradation for large area CCD sensors. Either accuracy predictions must be deliberately conservative to account for the expected degradation, based on a model for sensor surface unflatness, or premeasurement of the sensor surface topography must be carried out prior to the assembly of the camera.

In spite of all the considerations cited above, the multi-sensor self-calibration produced a robust and accurate result for the surface measurement of the bond tool, and the calibrations of the different sensor/lens combinations. The multi-sensor experiment afforded a valuable insight into calibration which would not be given to the same extent by single-sensor networks. It is recommended that where multiple sensors are to be utilised in a vision metrology system, that their calibration be established using this approach.

6. REFERENCES

- 1. H. A. Beyer, "Accurate calibration of CCD-cameras". Proceedings. <u>Computer Vision and Pattern Recognition</u>, CVPR '92, pp. 97-101, Champaign, USA, 1992.
- 2. H. A. Beyer, "Geometric and radiometric analysis of a CCD-camera based photogrammetry close-range system". <u>Dissertation No. 9701</u>, Institute for Geodesy and Photogrammetry, ETH-Zürich, Switzerland, 186 p. 1992.
- 3. H. A. Beyer, "Digital photogrammetry in industrial applications". <u>International Archives of Photogrammetry and Remote Sensing</u>, 30(5WI), 373-378, 1995.
- 4. D. C. Brown, "Decentring Distortion of Lenses" Photogrammetric Engineering, 32(3), 444-462, 1966.
- 5. D. C. Brown, "A large format, microprocessor controlled film camera optimised for industrial photogrammetry". XV International Congress of Photogrammetry and Remote Sensing, Commission V, 29p., Rio de Janerio, Brazil, 1984.
- 6. J. Dold, "A strategy for photogrammetric multiple camera calibration without additional object information". <u>International Archives of Photogrammetry and Remote Sensing</u>, 30(5), 61-70, 1994.
- 7. C. S. Fraser, "Multiple focal setting self-calibration of close-range metric cameras". <u>Photogrammetric Engineering and Remote Sensing</u>, 46(9): 1161-1171, 1980.
- 8. C. S. Fraser, "Network design considerations for non-topographic photogrammetry". <u>Photogrammetric Engineering and Remote Sensing</u>, 50(8): 1115-1126, 1984.
- 9. C. S. Fraser, "Photogrammetric measurement to one part in a million". <u>Photogrammetric Engineering and Remote Sensing</u>, 58(3): 305-310, 1992.
- 10. C. S. Fraser, "Photogrammetric camera component calibration A review of analytical techniques". XVII ISPRS Congress, Workshop TU-1 on Calibration and Orientation of Cameras in Computer Vision, 36p, Washington, D.C., USA, 1992.

- 11. C. S. Fraser and S. A. Veress, "Self-calibration of a fixed-frame multiple camera system". <u>Photogrammetric Engineering and</u> Remote Sensing, 46(11): 1439-1445, 1980.
- 12. C. S. Fraser and M. R. Shortis, "Variation of distortion within the photographic field". <u>Photogrammetric Engineering and Remote Sensing</u>, 58(6): 851-855, 1992.
- 13. C. S. Fraser and M. R. Shortis, "Metric exploitation of still video imagery", Photogrammetric Record, 15(85): 107-122, 1995.
- 14. A. Gruen and H. A. Beyer, "System calibration through self-calibration", XVII ISPRS Congress, Workshop TU-1 on Calibration and Orientation of Cameras in Computer Vision, 33p, Washington, D.C., USA, 1992.
- 15. P. C. Gustafson, and H. B. Handley, "A video-based industrial measurement system". <u>International Archives of Photogrammetry and Remote Sensing</u>, 29(B5): 501-506, 1992.
- 16. E. Kilpela, "Compensation of systematic errors of image and model coordinates", Photogrammetria, 37: 15-44, 1980.
- 17. R. Kinzel, "Photogrammetric surface measurements of "CPTR" reflectors". <u>2nd Int. Symp. Environmental Testing for Space Programmes</u>, 553-556, ESTEC, The Netherlands, 1993.
- 18. S. Mason, "Expert system-based design of photogrammetric networks". <u>Institute of Geodesy and Photogrammetry</u>, 187 p., ETH-Zürich, Switzerland, 1994.
- 19. J. Peipe, "Photogrammetric investigation of a 3000 x 2000 pixel high resolution still video camera", <u>International Archives of Photogrammetry and Remote Sensing</u>, 30(5W1): 36-39, 1995.
- 20. J. Peipe and C.-T. Schneider, "High resolution still video camera for industrial photogrammetry", <u>Photogrammetric Record</u>, 15(85): 135-139, 1995.
- 21. M. R. Shortis, A. W. Burner, W. L. Snow, and W. K. Goad, "Calibration tests of industrial and scientific CCD cameras", <u>First Australian Photogrammetric Conference</u>, Paper 6, 11p, Sydney, Australia, 1991.
- 22. M. R. Shortis, W. L. Snow, B. A. Childers and W. K. Goad, "The influence of storage media on the accuracy and repeatability of photogrammetric measurements using CCD cameras", <u>Videometrics II</u>, SPIE Vol. 2067, pp 80-92, 1993.
- 23. M. R. Shortis, W. L. Snow and W. K. Goad, "Comparative geometric tests of industrial and scientific CCD cameras using plumb line and test range calibrations, <u>International Archives of Photogrammetry and Remote Sensing</u>, 30(5W1): 53-59, 1995.
- 24. M. R. Shortis, T. A. Clarke, and S. Robson, "Practical testing of the precision and accuracy of target image centring algorithms", Videometrics IV, SPIE Vol. 2598, Philadelphia, USA, 1995.
- 25. W. L. Snow, B. A. Childers and M. R. Shortis, "The calibration of video cameras for quantitative measurements", <u>39th International Instrumentation Symposium</u>, pp. 103-130, Albuquerque, USA, 1993.