Automatic Calibration of a Video Camera Lens System

A thesis presented in partial fulfillment
of the requirements for the degree of
Master of Technology
at Massey University

Christopher Louis Hunt
1997

Abstract

An automated camera calibration procedure was successfully developed to allow accurate measurements to be made from images obtained from camera-lens systems exhibiting geometric and lens distortion. The calibration procedure is based upon image analysis methods programmed in script using Mathworks MATLAB software.

The process initially involved capturing an 512 x 512 pixel image of a calibration chart consisting of a regularly spaced two dimensional grid of circle shaped fiducial marks. Background leveling was used to correct intensity gradients due to non uniform illumination. A region of interest was determined for each fiducial mark in the image, by thresholding followed by an identification process. A grayscale Centroid Calculation Method was then used to accurately determine the center position of each fiducial mark. Two polynomial equations were fitted in the least squares sense to describe an inverse spatial transformation function that mapped fiducial mark position (x^T,y^T) in the image to the actual positions (x,y) in real world coordinates on the calibration chart. A second image of a calibration chart was obtained so that the overall calibration error could be assessed for the procedure.

The effect of altering the degree of the fitted polynomial equation on the error was investigated. No significant reduction in error was achieved by increasing the order of the fitted polynomial equations above order 4.

The effect of altering fiducial mark size on error was investigated. For the Fujinon TV lens being tested at a focal distance of 32 cm and a field size of approximately 18 cm square, an optimal fiducial mark diameter was determined to be 3 mm (8.5 pixels). Increased calibration error was obtained for fiducial mark diameters both greater and less than this figure.

The effects on calibration error of varying aperture were investigated. Greater calibration error observed at low aperture settings was attributed to primary lens aberration and also the possibility of systematic error in the Centroid Calculation Method due to spatial undersampling in the image. Increased diffraction resulting in loss of definition at the reduced aperture probably explains the increase in error that was observed at aperture setting, f22. An optimum aperture setting of f11 was determined, for this particular lens.

The use of this camera calibration procedure has resulted in a large increase in accuracy for position determination or measurement from an image. When the non-linear effects of geometric or lens distortion are ignored, the maximum observed error was seen to be as high as 5.10 pixels compared to the maximum error in a 4th order calibrated image of 0.77 pixels. Mean error was observed to decrease from 1.25 to 0.33 pixels in the calibrated case.

The mean error obtained as a result of the calibration closely approached the estimated uncertainty present in the physical calibration chart.

The computation time required for the calibration of an image of a control chart having 320 control points, including the calculation of a verification image, was found to be 6.5 minutes on a Pentium 100MHz computer. The advantage of the automated procedure is that it is accurate and fast, unlike manual methods that are tedious, time consuming and prone to error.

Acknowledgments

I am indebted to the people who have made this thesis possible. I would especially like to thank my supervisor Dr. Roger Browne who provided guidance and constructive criticism throughout this study.

I would also like to thank Dr. Bruce Campbell, AgResearch, for making the facilities and time available to conduct this work, and AgResearch for providing the necessary funding.

Table of Contents

Chapter 1 -Introduction.

Machine vision systems have become important components of a large range of manufacturing industries. These applications include part recognition, sorting, welding, robotic guidance and assembly operations, (SME, 1984).

Another field where computer assisted vision is becoming more prominent is in the area of photogrammetry. Photogrammetry is a dimensional measurement technique which is achieved through the analysis of two-dimensional photographic images. The technique has been employed for many decades in the marine construction industry as well as in architecture, and civil engineering roles. Precise optical instruments are used to extract positional information associated with the points of interest and, in the past, this has formed the basis of aerial surveying. Traditionally this was accomplished by photographic plates but the increasing application of video cameras in these areas has increased the requirement to obtain more accurate measurements from video camera and machine vision systems.

More recently image analysis systems have become of fundamental importance in research institutions. The ability of image analysis systems to obtain greater volumes of experimental data means that the value and the statistical significance of the results can be improved. Some examples of applications which are to be found in the agricultural research sector are seed shape characterization, leaf area, shape, size and color classification, root length measurement, stomata density and length measurement, white clover root nodule size and root architectural classification.

A typical image acquisition system that could be used for some of these applications is shown in Figure 1-1.

Figure 1-1

A video camera is placed in an appropriate position to obtain an image of an object under suitable lighting conditions. An image of the object is formed through the camera optics and electronics and is converted into a series of electrical impulses which is stored as an array of intensity values in an image matrix or memory of the frame grabber device. This image is accessed by a host computer and can be processed using appropriate software resident in the host computer, in order to extract the required information from the image.

In the majority of cases it is desirable to preserve the linear dimensions of the object in the image so that measurements obtained from the image accurately represent those dimensions of the object that need to be measured. In reality, this is not the case, as during the image capture process, the image is modified by the non linear characteristics of the lighting conditions, or illumination, the camera optical system (the lens) and the camera and frame grabber electronics. A diagrammatic representation of the image processing system is shown in Figure 1-2 with these main components illustrated.

During the image capture process a number of illumination regimes are possible depending on the type of the information that is required from the image. Back lighting may be used for object dimension determination in the case of root length, seed shape or enumeration applications. Direct illumination may be preferential to maximize object color or spectral contrast, and diffuse side lighting may be desirable if textural information is required from the object.

The correct aperture and focus settings are important if the amount of information in the image about the object is to be maximized. The correct aperture will ensure that the contrast in the image will be optimized and the detail of object features preserved. For the same reason it is important that the illumination profile is as uniform as possible across the image (Russ, 1995) as subsequent image leveling operations will result in a reduction in contrast.

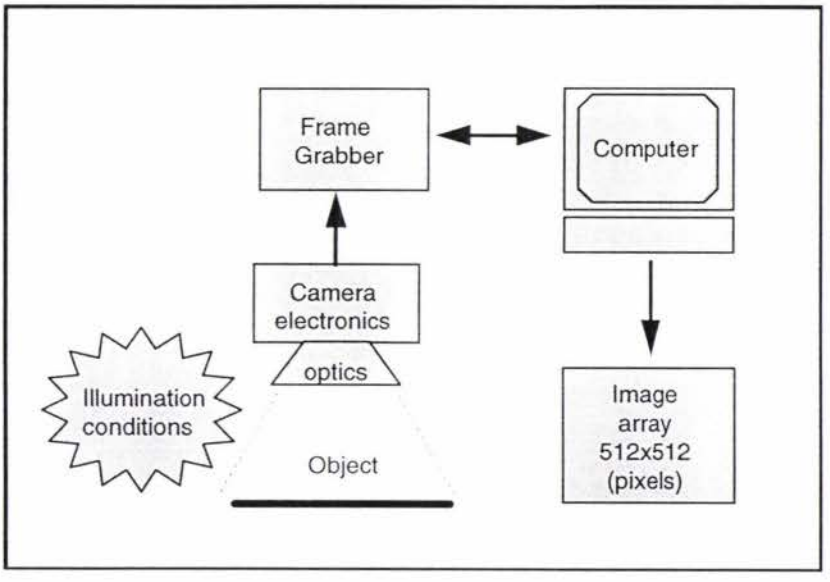


Figure 1-2

Unlike the high quality metric cameras that are used in photogrammetry, the solid-state video cameras used in these computer vision applications tend to have a much lower spatial resolution, due to the use of charge coupled device (CCD) sensors. The resolution is determined by the geometric construction of the CCD which consists of a fixed array of optical sites that detect light reflected from the object. An example resolution for such a device is given for the Sony ICX087AK which has an effective pixel resolution of 500x582 when used in PAL video format.

In addition to this restricted resolution, video cameras typically incorporate the use of asymmetric lens' design which exhibit a high degree of non-linearity due to distortion. It is because of this that the majority of commercially available video cameras are affected by radial and decentering distortions. Variation of these distortions are greater for zoom lens systems as the focal length and focus are adjusted and the magnitude of these distortions can result in non-linear displacement of the object well in excess of a pixel on the image sensor plane.

The use of these cameras in the mensuration of an image, that is the extraction of length, distance or volume information, has resulted in the development of image processing techniques that allow super resolution or subpixel accuracy to be achieved despite the limited resolution of the CCD sensor that is used in the camera. The use of these subpixel techniques are futile if larger errors attributable to the various lens distortions or camera errors are not taken into consideration and the greatest precision will not be achieved unless these errors are incorporated into camera lens calibration procedures.

Camera calibration is the process of determining the internal camera geometric and optical characteristics and the 3D position and orientation of the camera frame relative to a specified real world coordinate system which is defined by a spatial coordinate system present on a calibration chart or control grid. It is therefore necessary to adopt a calibration procedure that will take into account the geometric and optical characteristics of the video camera system and to obtain a spatial transformation function that maps the image pixel positions into 3D real world coordinates. Calibration procedures can also take into account sources of error in the camera electronics, such as line jitter, voltage fluctuations and temperature induced drift which may be significant at the subpixel level.

The magnitude of distortion present in a video camera system varies considerably depending on the focus setting of the lens. Because of this variation, it is necessary to recalibrate a video camera at the specific focus setting at which it will be operating.

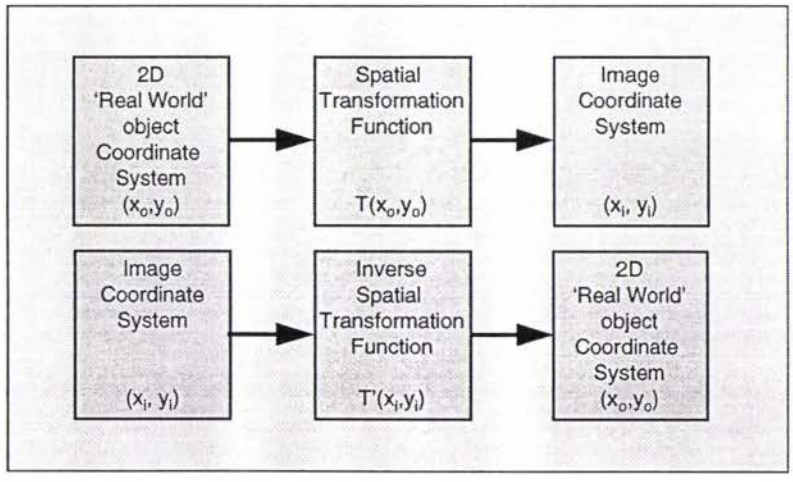


Figure 1-3

In our particular case the model can be simplified by reducing the problem to that of a two dimensional case due to the fact that the objects being investigated appear in coplanar form with measurement information to be extracted from photographic plates. The obtained transformation takes into account camera geometric and optical distortion as well as the position and orientation of the camera frame. This enables the accurate mensuration of subsequently acquired images, that is the extraction of increased accuracy measurements made from images captured from photographic plates for a particular camera lens arrangement (Shih, 1993). The result of this calibration is the determination of a spatial transform function T and its inverse T', as shown diagramatically in Figure 1-3. Having obtained the inverse spatial transform function it is possible, using subpixel techniques, to obtain more highly accurate information about an object's position from its image.

Traditional camera calibration methods include the 'Plumb line' technique developed by Brown (1971) for photogrammetric studies but have proved very expensive both in terms of labor, time and equipment. An automated method of calibration can offer the benefits of a reduction in time and effort as well as reducing potential errors (Fryer & Mason, 1987).

There are many benefits to be obtained by the development of an automated technique, using image analysis. The properties of such a process have been investigated by Tsai (1987), who describes a versatile camera calibration procedure and specifies the following attributes. It must be autonomous in that it requires no operator intervention. It must be accurate in that resolution is enhanced by the use of subpixel or super-resolution techniques. The efficiency of the process must compare favorably with manual methods in terms of time and cost, and the calibration technique must be versatile in that it will operate successfully for a range of different camera and lens systems, optical configurations and distances. It must work on cameras that are commonly available.

Chapter 2 -Camera Lens Systems.

Two simplified camera lens configurations are shown in Figure 2-1 and Figure 2-2.

In the first case, Figure 2-1, illustrates a dual lens system in which the effective focal length lies between the two lenses. The fixed focal length system is focused by fine adjustment of the objective lens, which alters the distance of the lens system from the image plane allowing objects to be in focus from infinity down to a short distance. In this particular case, an object at infinity is shown in black to be in focus. However the image of an object at a short distance from the lens, shown in red, falls behind the image plane. In this instance it is necessary to increase the distance of the lens system from the image plane to ensure that the image of the object at the specified distance is focused sharply onto the image plane.

In the mechanical construction of a camera, adjusting the position of the lens is accomplished by rotation of the lens system in a track or thread, the pitch of which determines the extent of the lens travel in response to the rotation angle.

In the case of a solid state video camera, the image plane is occupied by the CCD sensor which performs the role of converting the image into a series of electrical impulses.

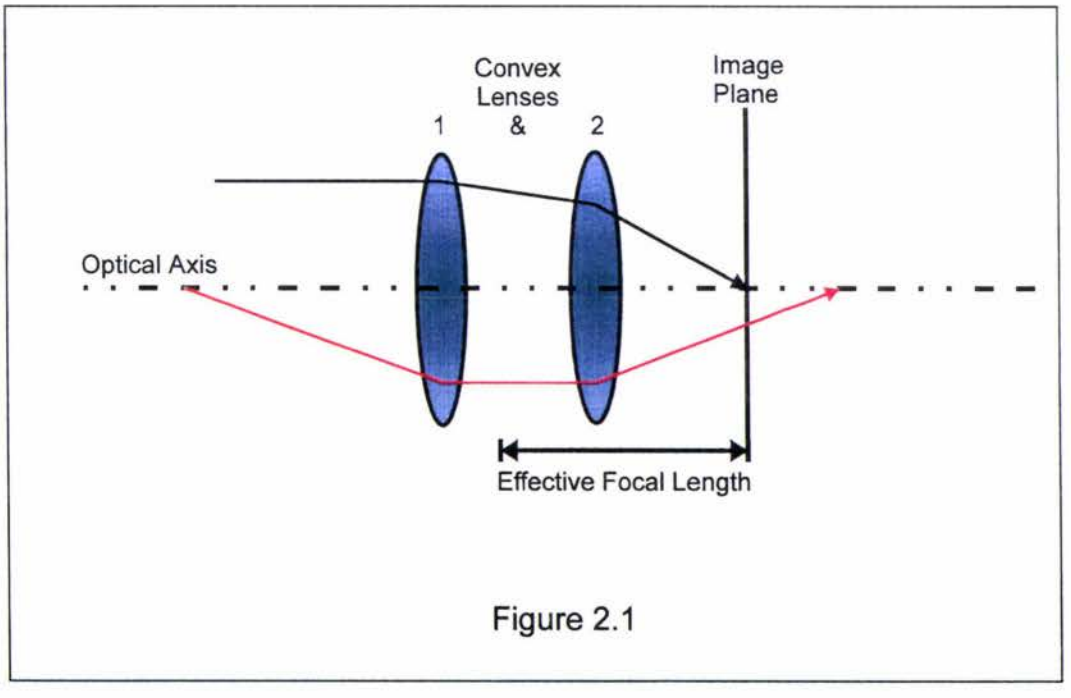


Figure 2.1

In the second configuration, a negative (concave) lens is used in the optical path, to alter the focal length of the system and thereby achieving a variable magnification or zoom. As the two lenses are shifted along the lens barrel towards each other, there is a increase in magnification corresponding to the increased focal length of the system. Focus in this case is also achieved by fine adjustment of the objective lens.

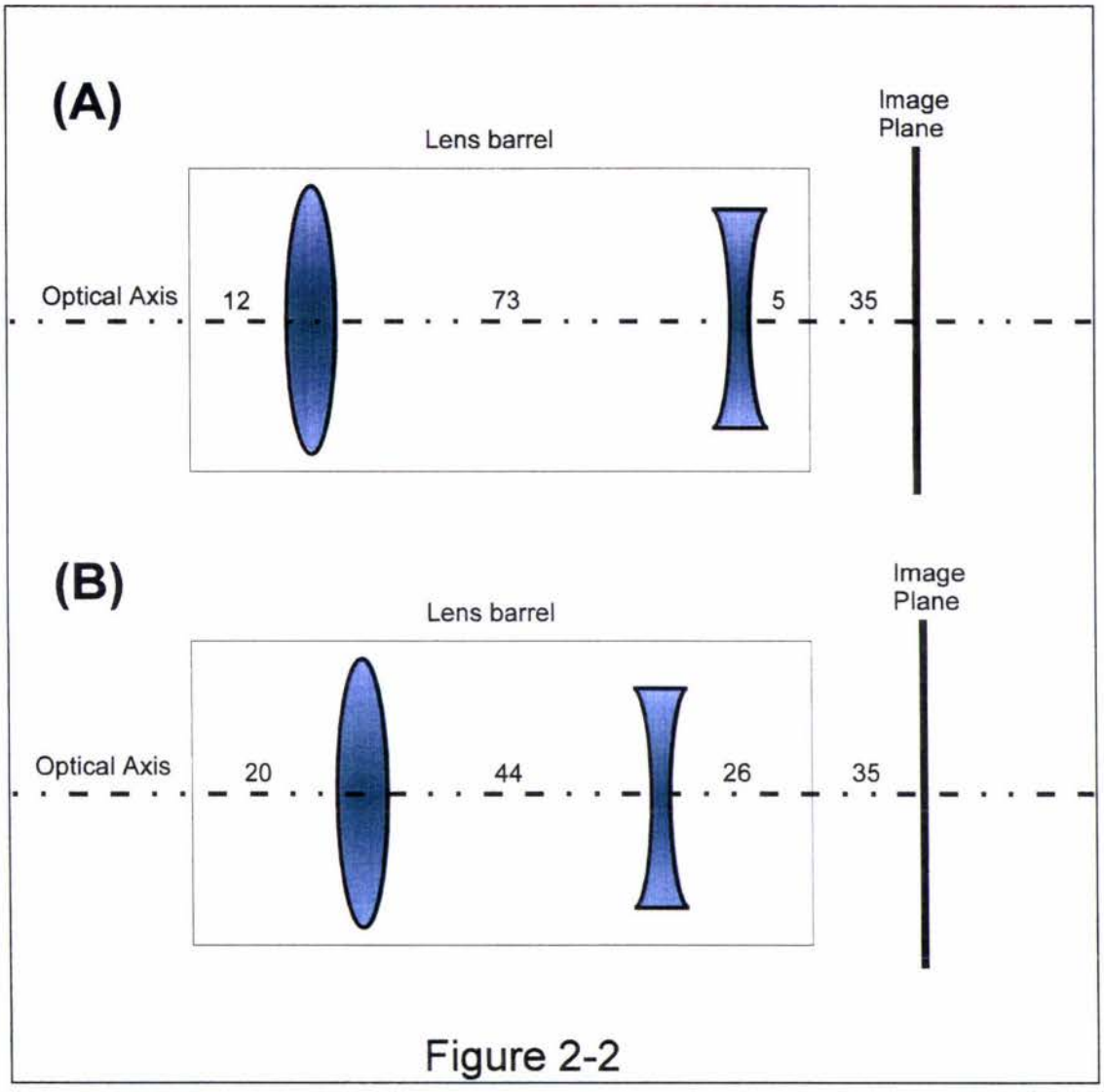


Figure 2-2

A zoom lens system is shown in Figure 2-2, and illustrates how the positioning of lens combinations are adjusted to achieve a variation in magnification due to a corresponding change in focal length. The distances stated represent an approximate measurement in millimeters, for a Canon 35-70 mm FD zoom lens (Fryer, 1986). Figure 2.2(A) represents the zoom lens configuration for the case when the combined equivalent focal length of the system is equivalent to 35 mm. In Figure 2.2(B), the negative lens is moved towards the objective lens (that lens closer to the object) and the effective focal length of the system is increased to 70 mm.

2.1 Lens Aberration.

The purpose of a lens in a camera system is to receive light from an object and to focus that light onto a corresponding position in an image of that object. The ability of a lens to accomplish this successfully, and to form an ideal image of the object, is prevented by a number of errors, termed primary lens aberrations. There are seven primary lens aberrations described by Walker (1994). These aberrations are:

- Spherical aberration, which is the lack of a common focus for on-axis (rays parallel to the optical axis) light rays incident to the lens at different distances from the primary optical axis, causing a point in the object to appear as a spot in the image.

- Coma, where off axis rays (rays not parallel to the optical axis) from an object are incident to the lens in different zones and are focused into the image in different positions and spot sizes, resulting in a triangular shaped blurring.

- Field curvature, is the error resulting when an essentially curved image from the lens system is formed on a flat surface.

- Astigmatism, occurs when the lens exhibits a different effective focal length for a tangential fan of light rays from an object than that for a sagitally orientated fan of rays. This occurs when the spherical lens is obloid and results in an elliptical image of a point source.

- Axial Color Aberration, (chromatic aberration) occurs because refraction by the lens is a function of wavelength. When light rays that are parallel to the optical axis are refracted by the lens, rays of varying wavelength will be focused at slightly different positions resulting in a blurred spot consisting of a green central region surrounded by a purple outer ring.

- Lateral color aberration, (chromatic aberration) occurs when off-axis light rays are focused by the lens at slightly different heights on the image plane resulting in a blurred image.

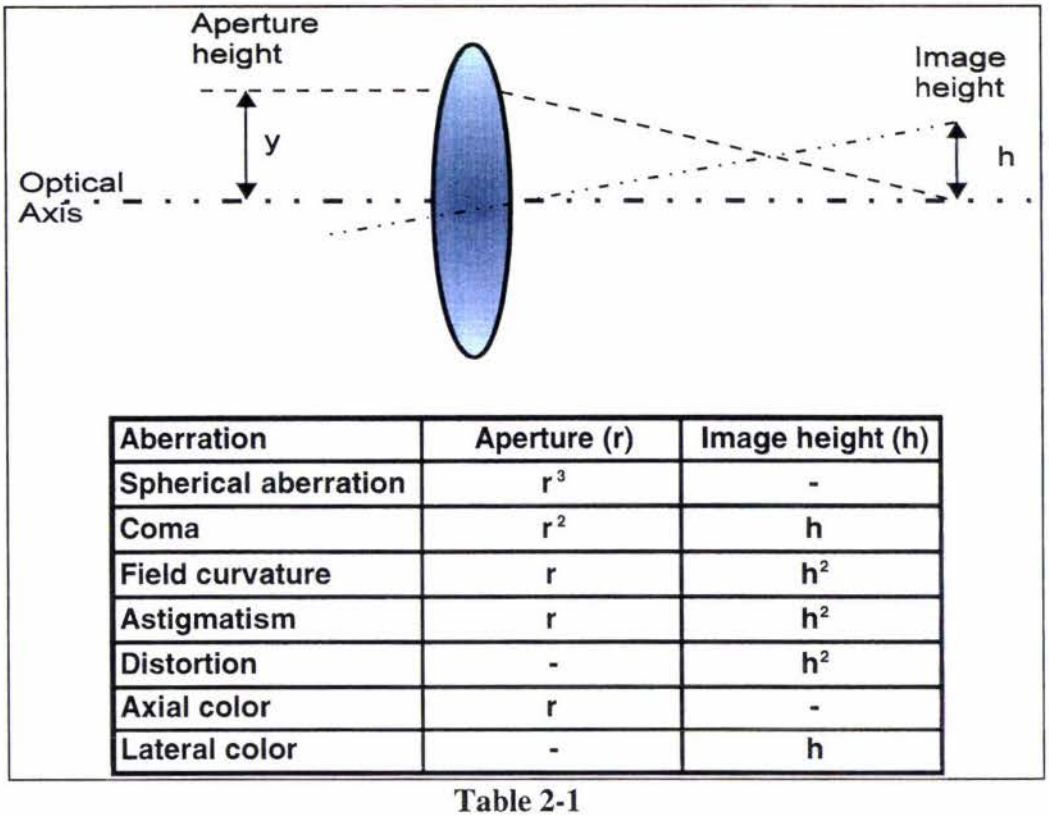
- Distortion of the image is present when the focal length of the lens varies as a function of field size or image height.

The main effect of the primary lens aberrations, described in this section, is to cause light from a specific point on the object to be focused onto an image not as a point but as a blurred circle or a spot. As the aberration becomes worse the diameter of this spot increases resulting in a reduction of the modulation or contrast in the image.

Aberration has a varying effect on image quality depending on both the lens aperture radius (r) and also image height (h), or displacement from the optical axis. The manner in which the blurred circle diameter varies with each of these parameters, for each type of primary lens aberration, is described by Walker (1994) and shown in Table 2-1.

The effect that spherical aberration has in an image causes the blur spot diameter to increase in proportion to the cube of the physical aperture size, but it is independent of image height. Blurring in an image attributable to coma increases in proportion to the square of physical aperture and also in proportion to image height. Astigmatism is seen to vary with the square of image height.

In general it can be seen that loss of definition in an image due to lens aberration increases with both increasing aperture radius and image height.



Aberration	Aperture (r)	Image height (h)
Spherical aberration	r^3	-
Coma	r^2	h
Field curvature	r	h^2
Astigmatism	r	h^2
Distortion	-	h^2
Axial color	r	-
Lateral color	-	h

Table 2-1

Another phenomenon that can affect the definition in an image is diffraction. Diffraction occurs when light from an object passes through an aperture (Ray, 1992). Some of the of light around the edge of the aperture is deviated from its normal path. As the aperture size is reduced, a greater proportion of the total light is deviated resulting in an image that is not a spot but an 'airy disk'.

Although not classed as an aberration, diffraction has the effect of a reduction in image contrast and a loss of definition.

2.2 Lens Distortion

Of all the lens aberrations described, lens distortion is the only one that does not alter the image quality in terms of sharpness or focus. The result of distortion when present in a lens is to warp the overall shape of the image and to introduce non linearity into the image. Hence the image does not represent a scaled reproduction of the original object.

From Table 2-1, it can be seen that distortion is independent of aperture radius (r), but increases with the square of image height (h) or displacement from the optical axis.

For visual systems, image distortion of up to 10% may be acceptable and for camera and projection systems distortion values of up to 2% are common, (Walker, 1994). For photogrammetric applications where accurate measurements are required, 2% represents a significant unacceptable error, and distortion is a problem that must be eliminated.

The distortion present in the lens systems is primarily radial distortion and decentering distortion, (Fryer,1987).

2.2.1 Radial Distortion.

Radial distortion results in an apparent displacement of the radial distance from the center of the image. If this shift is an increase in distance then the result is observed as 'barrel' or positive distortion. If the shift is a reduction in the radial distance then the observed result is a negative distortion or 'pin cushion distortion'.

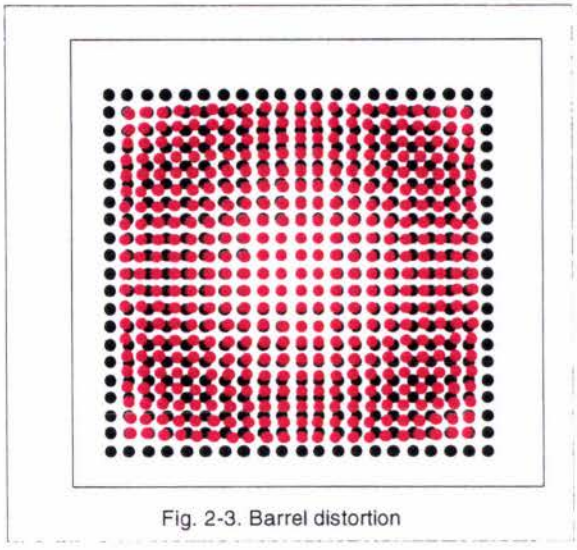


Fig. 2-3. Barrel distortion

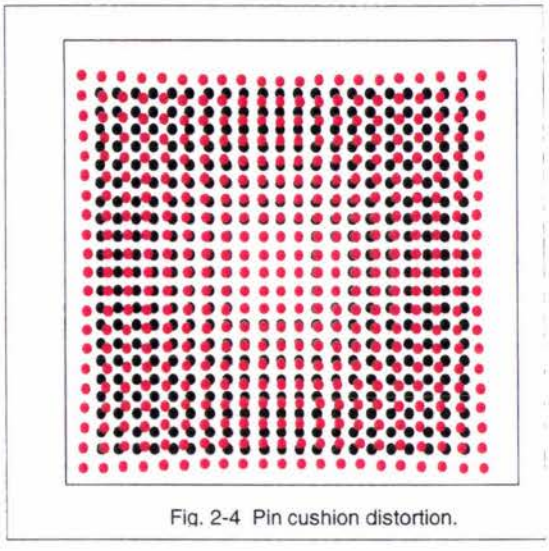


Fig. 2-4 Pin cushion distortion.

Figure 2-3 illustrates the effect of barrel distortion. The displaced or transformed coordinates in the image are shown in red with the original point positions shown in black. It can be seen that as the radial distance of each point increases from the center, there is a proportional reduction in the displacement of that point from the center of the image.

If the radial displacement increases with increasing distance from the center, pin cushion distortion is observed. The effects of pin cushion distortion are shown in Figure 2-4.

Although it is stated by Weng (1992) that radial distortion is the result of flawed radial curvature of the lens elements, radial distortion is actually an inherent characteristic of lenses due to their spherical surfaces. Radial distortion is strictly symmetric about the central optical axis with distortion being a minimum at the center and progressively increasing with an increase in distance from the center.

Radial distortion can be described in the following mathematical relationship,(Weng,1992), (Fryer 1986).

$$\delta_{\rho r} = d(\rho) = k_1\rho^3 + k_2\rho^5 + k_3\rho^7 + \ldots \qquad \text{Equation 2.1}$$

where:
δ= the radial distortion as a change in radial displacement, d(ρ)
ρ= the radial distance from the principal point of the image plane
k_n=coefficients of radial distortion.

The image points may also be expressed in terms of a Cartesian coordinate system in the following relationship:

$$\begin{bmatrix} \sin\phi \\ \cos\phi \end{bmatrix} = \frac{1}{\rho}\begin{bmatrix} y \\ x \end{bmatrix} \qquad \text{Equation 2.2}$$

Using this relationship it is possible to derive the component of radial distortion corresponding in the x and y-axis directions:

$$\delta_x = k_1 x(x^2 + y^2) + O[(x,y)^5] \qquad \text{Equation 2.3}$$

$$\delta_y = k_1 y(x^2 + y^2) + O[(x,y)^5] \qquad \text{Equation 2.4}$$

Where:

δ_x	=	x-axis component of radial distortion.
δ_y	=	y-axis component of radial distortion.
x	=	x-axis displacement.
y	=	y-axis displacement.
O(x,y)	=	the principal point on the image axis.

It must be noted that the principal point on the image plane corresponds to where the optical axis of the lens system intercepts the image plane. This does not necessarily correspond to the geometric center of the image plane or in the case of solid state video cameras, the CCD array.

A study performed by Fryer (1986), on Canon 35-70 mm zoom lenses, found that the coefficients of radial distortion, kn, in Equation (2.1), where n>=3, proved to be insignificant. Radial distortion is known to change in a predictable manner when the lens is focused at different distances less than infinity, (Magill 1955).

2.2.2 Decentering Distortion

Decentering distortion occurs in a lens system when the optical center of each element in the system is not precisely aligned, (Brown, 1966). In this event, the actual center of an image will not coincide with the center of distortion that is associated with the lens system.

This problem can become increasingly significant in complex multi-element lens systems and results in distortion that has both a radial and tangential components.

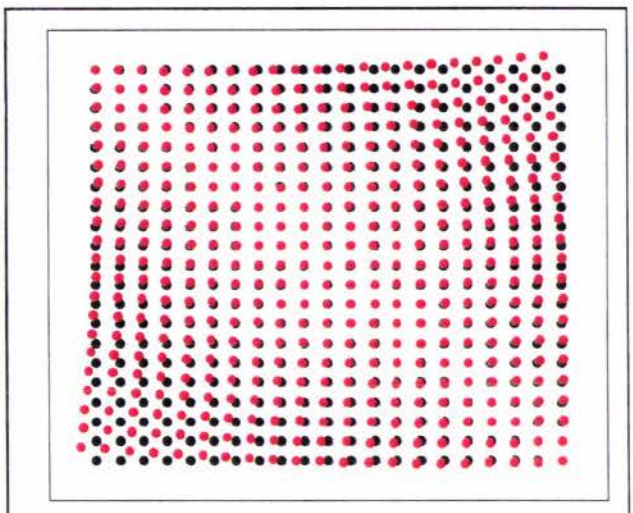


Fig. 2-5. Tangential distortion

Figure 2-5 demonstrates the effect of tangential distortion. In the image, points experience a shift in angular displacement around the center of the image or optical axis. The angular displacement is intensified with increasing radial displacement from the center. There is an axis of maximum tangential distortion and correspondingly an axis where this distortion is a minimum.

A mathematical relationship (Weng, 1992) that describes the effects of decentering distortion is given in the following expression:

$$\delta_{\rho d} = 3\,(\,j_1\rho^2 + j_2\rho^4 + j_3\rho^6 + \ldots)\,\sin(\varphi - \varphi_0) \qquad \text{Equation 2.5}$$

$$\delta_{td} = (\,j_1\rho^2 + j_2\rho^4 + j_3\rho^6 + \ldots)\,\cos(\varphi - \varphi_0) \qquad \text{Equation 2.6}$$

where:

$\delta_{\rho d}$	=	radial component of decentering distortion
δ_{td}	=	tangential component of decentering distortion
φ_0	=	axis of maximum tangential distortion
φ	=	position angle of object element
j_n	=	coefficients of decentering distortion

Equations 2.5 & 2.6 can be expressed in terms of x and y-axis decentering distortion:

$$\delta_{xd} = p_1(3x^2 + y^2) + 2p_2xy + O[(x,y)^4] \qquad \text{Equation 2.7}$$

$$\delta_{yd} = 2p_1xy + p_2(x^2 + 3y^2) + O[(x,y)^4] \qquad \text{Equation 2.8}$$

2.2.3 Thin Prism Distortion:

Another type of distortion present in optical systems is thin prism distortion, (Weng, 1992). This results from defective design or construction as well as incorrect lens assembly and like decentering distortion is especially significant in multiple lens systems. The nature of the distortion that results is similar and includes components of radial and tangential distortion. The term 'thin prism distortion' arises from the fact that this aberration can be modeled by the addition of a thin prism into the optical system.

Thin prism distortion can be expressed analytically in the following relationship.

$$\delta_{\rho p} = (i_1\rho^2 + i_2\rho^4 + i_3\rho^6 + \ldots) \sin(\varphi - \varphi_1) \qquad \text{Equation 2.9}$$

$$\delta_{tp} = (i_1\rho^2 + i_2\rho^4 + i_3\rho^6 + \ldots) \cos(\varphi - \varphi_1) \qquad \text{Equation 2.10}$$

where:

$\delta_{\rho p}$	=	radial component of thin prism distortion
δ_{tp}	=	tangential component of thin prism distortion
φ_1	=	axis of maximum tangential distortion
φ	=	position angle of object element
i_n	=	coefficients of thin prism distortion

The x and y-axis distortion components can be expressed as:

$$\delta_{xp} = s_1 x(x^2 + y^2) + O[(x, y)^4] \qquad \text{Equation 2.11}$$

$$\delta_{yp} = s_2 (x^2 + y^2) + O[(x, y)^4] \qquad \text{Equation 2.12}$$

where: $s_n = -i_n \sin \varphi_n$

It must be noted that although both thin prism distortion and decentering distortion both give rise to similar relationships and coefficients, that they model two different types of distortion with the tangential components more than likely having a different axis of maximum tangential distortion.

(e.g.) $\varphi_0 <> \varphi_1$

2.2.4 Total distortion.

The total distortion that is associated with any particular lens system is an addition of these three main types of distortion.

This can be estimated from the sum of equations for radial, decentering and thin prism distortion. (2.3,2.4),(2.7,2.8) & (2.11,2.12) respectively.

Assuming that the terms in the equation of order 3 and higher are negligible, the combined equation describing the total lens distortion present can be given.

$$d_x(x,y) = (s_1 + 3p_1)x^2 + 2p_2xy + (s_1 + p_1)y^2 + k_1x(x^2 + y^2) \quad \text{Equation 2.13}$$

$$d_y(x,y) = (s_2 + p_2)x^2 + 2p_1xy + (s_2 + 3p_2)y^2 + k_1y(x^2 + y^2) \quad \text{Equation 2.14}$$

By grouping the coefficients for the powers of x & y, these equations can be expressed in the following simplified form:

$$d_x(x,y) = (g_1 + g_3)x^2 + g_4xy + g_1y^2 + k_1x(x^2 + y^2) \quad \text{Equation 2.15}$$

$$d_y(x,y) = g_2x^2 + g_3xy + (g_3 + g_4)y^2 + k_1y(x^2 + y^2) \quad \text{Equation 2.16}$$

2.2.5 Lens distortion in video cameras.

Distortion is inherent in the design of video camera lens system due to the asymmetrical lens arrangement around the position of the aperture stop (Ray, 1992). As described in Section 2.0, the nature of the construction of either a standard fixed focus lens or a zoom lens will mean that distortion will vary with focal length variation in a zoom lens or with focus in the standard lens. This is because the rotation of the lens during focus will change the position of the center of any thin prism distortion that is present (due to the non alignment or orthogonality of multiple lens elements in the system), and will also change the axis of any tangential component that may also be present due to decentering distortion. For a zoom lens, the variation in the positioning of an inner lens can significantly affect the coefficients relating to decentering distortion, as it tracks along the optical axis (Fryer, 1987).

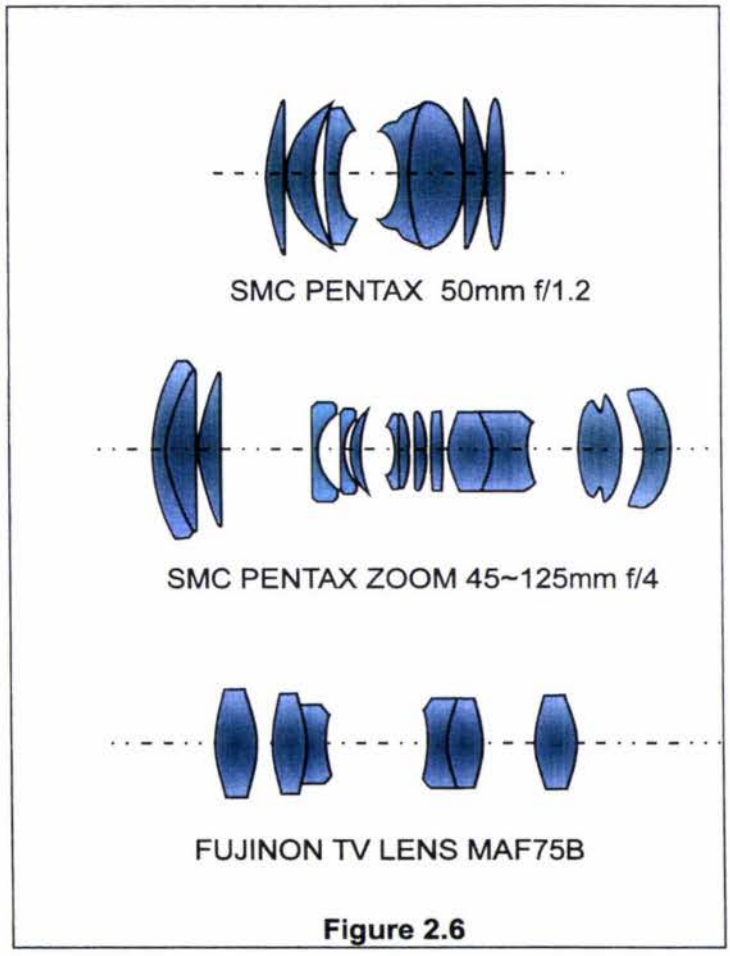


Figure 2.6

Figure 2.6 shows several camera lens configurations including a fixed focal length Fujinon TV lens. In general off the shelf video camera lenses tend to be of a simpler design when compared to the more elaborate lenses that are available for Single Lens Reflex (SLR) cameras, Fryer (1987). In terms of radial distortion, Fryer & Mason (1989), notes that most video camera lenses are designed with minimal distortion occurring at the principal focus point, or when the camera is focused at infinity. For industrial robotic applications as well as close range photogrammetric mensuration, the magnitude of radial distortion will be far greater than that experienced for distant focus applications. Fryer concludes that it is therefore very important to have a fast and accurate method for video lens calibration so that this may be accomplished at whatever focus setting is being used.

Chapter 3 -Camera Alignment Geometry

The alignment of the camera on its stand can influence the spatial transformation function that maps the positions of a bilinear calibration grid onto the image plane. The effects that camera perspective and alignment geometry can have on the captured image are illustrated in Figure 3-1.

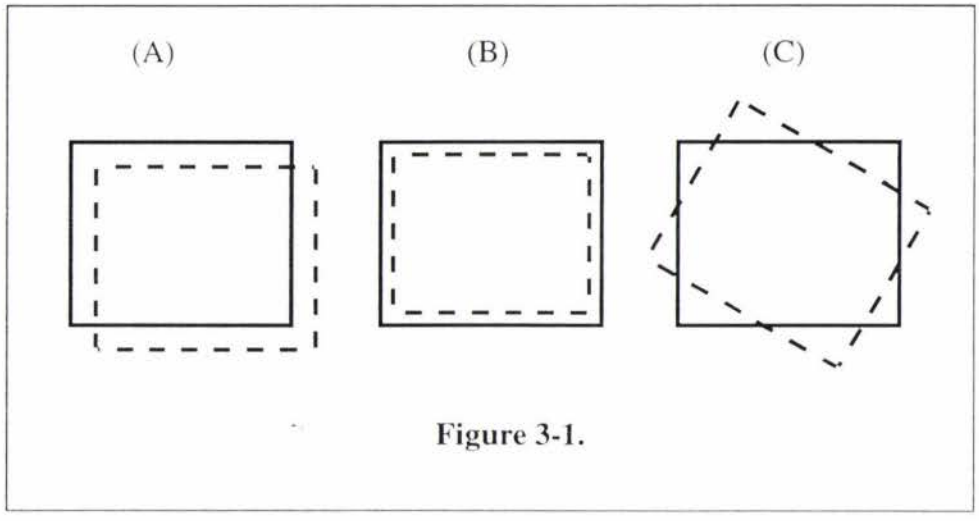


Figure 3-1.

Incorrect camera centering will result in the center of the camera's optical axis not aligning exactly with the center of the calibration grid. This is illustrated in Figure 3.1(A), and appears as a bilinear translation of the object in the image.

A variation in the camera's height will create an apparent distortion in the object's size, demonstrated in Figure 3.1(B) and a rotation of either the object within the field or the actual camera on the mount will cause a rotation of the spatial transformation function, illustrated in Figure 3.1(C).

A mathematical relationship that represents these transformations is described by Wolberg (1990) as an Affine warp and is represented by the following set of equations:

$$x^{T}(x, y) = a_{11}x + a_{21}y + a_{31} \qquad \text{Equation 3.1}$$

$$y^{T}(x, y) = a_{12}x + a_{22}y + a_{32} \qquad \text{Equation 3.2}$$

In this case, $x^{T}(x, y)$ and $y^{T}(x, y)$ are the transformed coordinates in the image plane which are a function of the untransformed coordinates, x and y, which are coordinate positions in the object real world plane. The coefficients, a_{nm} , are determined by the magnitude of the transformation.

These equations take into account image transformations involving position translation, object or camera rotation and scale effects due to camera height or object distance which are shown in Figure 3-1.

Aspect angle distortion will occur if the horizontal or vertical axis of the camera is not perpendicular to the object plane and this is illustrated in Figure 3.2(A) and 3.2(B). This type of distortion results if, due to the angle of the camera, one end of the object plane is actually closer to the camera than the opposite end. Consequently that part of the object that is closer to the camera will appear to be larger in size when compared to that part of the object that is at a greater distance.

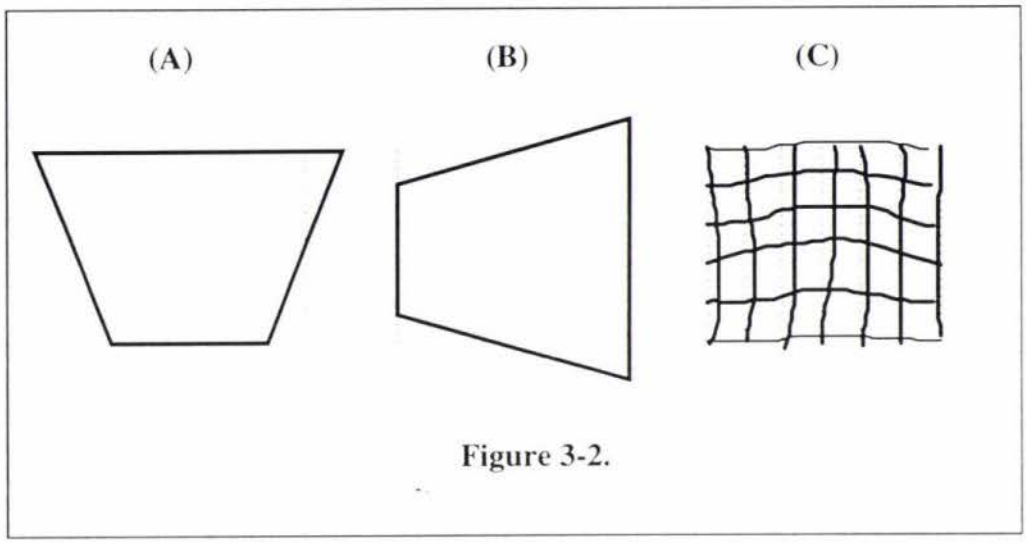


Figure 3-2.

The mathematical relationship that describes this type of a transformation of the object coordinates onto the image plane is given by the following set of equations: (Wolberg, 1990).

$$x^{T}(x,y) = (a_{11}x + a_{21}y + a_{31})/(a_{13}x + a_{23}y + a_{33}) \quad \text{Equation 3.3}$$

$$y^{T}(x,y) = (a_{13}x + a_{23}y + a_{33})/(a_{13}x + a_{23}y + a_{33}) \quad \text{Equation 3.4}$$

The perspective transformation that results cause equally spaced lines or elements in the object to appear as non linearly spaced elements in the image.

Although not directly associated with camera alignment geometry, Figure 3.2(C) shows the effect that terrain relief distortion would have if the object or calibration grid is not, as in the ideal case, only occupying a single plane. In the application of digital photogrammetry, or using a video camera to extract information from a diagram or a photograph, this effect can be successfully eliminated by covering the object with a sheet of glass and hence ensuring that it is as flat as possible, and making terrain relief distortion negligible.

Chapter 4 -Modeling

Optical lens distortion and geometric distortion associated with camera alignment have been described in the preceding sections. It is these factors that determine the components of the spatial transformation function that map points in the object plane onto corresponding points in the image plane. Each of these distortions has been characterized by gradual variations in position across the image and not localized rapid changes. Because of this property, it is possible to model the overall spatial transformation function with low order (N) polynomial equations. Common values that have been used in practice range up to N=4, Wolburg (1990).

The standard formula for the polynomial equation is given:

$$x^T=f(x,y)=a_1+a_2y+a_3x+a_4y^2+a_5xy+a_6x^2+ a_7y^3+ a_8xy^2+a_9x^2y+a_{10}x^3+..$$ Equation 4.2

$$y^T=f(x,y)=b_1+b_2y+b_3x+b_4y^2+b_5xy+b_6x^2+ b_7y^3+ b_8xy^2+b_9x^2y+b_{10}x^3+..$$ Equation 4.3

where:

(x,y)	are the positions in the object plane
(x^T,y^T)	are the spatially transformed coordinates of the object in the image
a_n & b_n	are the polynomial coefficients.

An inverse transform function can be determined from this equation enabling measurements made in the image to be transformed to accurate real world measurements in the object plane. This is possible providing the polynomial coefficients are determined accurately.

In order to accomplish this for a known camera-lens and system geometry, a control grid consisting of a series of fiducial marks, whose positions are known to a high degree of accuracy, is used to calculate the spatial transform function from the positions of the fiducial marks determined in the acquired image.

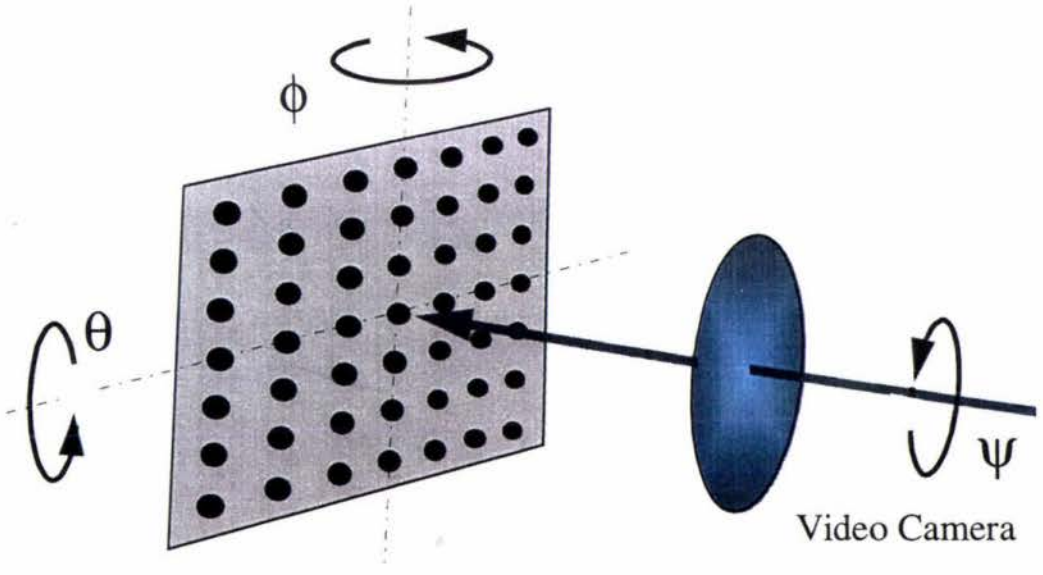


Figure 4-1

In this manner it is convenient to 'lump together' the parameters associated with the system geometry, or camera positioning, with those parameters associated with lens distortion in the system. The reason for this is that it is not possible to position the camera to subpixel accuracy using manual methods. The variables associated with camera positioning, (e.g.) camera roll angle, ψ, camera yaw angle, θ, and camera pitch angle, ϕ, as described by Nomura (1992), as well as camera height, and position of the principal point or optical center are included in the model. These parameters are illustrated in the diagram Figure 4-1. The spatial transform function will then take into account all variations associated with geometric distortions in the camera sensor, optics and positioning.

In order to calculate the spatial transformation function for any system, spatial information must first be supplied in an object as control points which have precisely defined positions. This can be achieved by the use of a control grid of fiducial marks whose positions are accurately known. The position of these marks represent absolute real world coordinates.

An image of this control grid is recorded using a camera-lens system in a fixed geometric position.

To then calculate the spatial transformation function, the control points must be identified in the image and their transformed positions accurately determined. The order of the polynomial to be fitted for the function determines the number of control points required to model the system. In general, accuracy increases with increasing number of control points (Shih, 1993) and if a high number of points are used then it is desirable to incorporate an automated image analysis algorithm that can be used to extract the position of the fiducial marks and then give transformed positions of the marks in the image plane, so that mathematical derivation of the polynomial coefficients is then possible.

Having obtained the coefficients for the forward transformation, calculation of the inverse transformation is straight forward. This will give the operator the ability to accurately measure the location and distances of objects in subsequent images taken for the particular camera lens system in the fixed position in which the camera was setup for the calibration image.

Chapter 5
-Calculation of the polynomial coefficients

Wolberg (1990) discusses several practical mathematical methods that can be used to derive the coefficients of the polynomial equations for the spatial transformation function that best fits the model that has been described previously in Chapters 3 & 4.

It is generally accepted that the equations are bivariate polynomial transformations described in the following set of equations. (Bivariate means that the transformed coordinate for either x or y, that is x^T or y^T, is a function of both x and y, the untransformed coordinates).

$$x^T = \sum_{i=0}^{N} \sum_{j=0}^{N-i} a_{ij} x^i y^j \qquad \text{Equation 5.1}$$

$$y^T = \sum_{i=0}^{N} \sum_{j=0}^{N-i} b_{ij} x^i y^j \qquad \text{Equation 5.2}$$

Where N is the order of the polynomial equation.

There are various methods that can be used to calculate the polynomial coefficients.

5.1 Pseudoinverse solution.

A correspondence is achieved between M points in the untransformed (object plane) and transformed coordinate systems. The bivariate polynomial that is chosen to best represent this transformation function is decided to be of order, N.

As an example, for a second order polynomial, the number of coefficients to be generated from equation (5.1) would be six.

$$\begin{bmatrix} x_1^T \\ x_2^T \\ x_3^T \\ \\ \\ x_M^T \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 & x_1 y_1 & x_1^2 & y_1^2 \\ 1 & x_2 & y_2 & x_2 y_2 & x_2^2 & y_2^2 \\ 1 & x_3 & y_3 & x_3 y_3 & x_3^2 & y_3^2 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & x_M & y_M & x_M y_M & x_M^2 & y_M^2 \end{bmatrix} \begin{bmatrix} a_{00} \\ a_{10} \\ a_{01} \\ a_{11} \\ a_{20} \\ a_{02} \end{bmatrix} \qquad \text{Equation 5.3}$$

A similar equation can be stated for the transformed coordinates of y^T and both equations may be represented in matrix form by the following expressions.

$$x^T = WA \quad \& \quad y^T = WB \qquad \text{Equation 5.4}$$

Solving for the polynomial coefficients:

$$A = W^{-1} x^T \quad \& \quad B = W^{-1} y^T \qquad \text{Equation 5.5}$$

Given that it is possible to calculate the matrix, W in each case, and that the transformed coordinates for both x & y are available, all that remains to be done is to calculate the inverse of W and then to calculate the coefficients from Equation (5.5). However it is here that the situation becomes complicated. Because the matrix, W, is not a square matrix, the inverse for it does not exist, and the normal Gaussian elimination procedures can not be implemented.

A more complicated procedure must be adopted, and hence the name given to it, the Pseudoinverse Solution.

The preferred method for the solution of Equations 5.4 & 5.5 is by singular value decomposition, or QR decomposition, a method that is described in detail by Reichel (1991), and the method adopted by the Mathworks, MATLAB, matrix manipulation package, when used for the solution of non square matrix calculations, (i.e.) the matrix division of A into B, where A is an M-by-N matrix with M <> N and B is a column vector with M components.
$X = A^{-1}B$ is the solution in the least squares sense to the under or overdetermined system of equations $A^*X = B$. The effective rank, K, of A is determined from the QR decomposition with pivoting.

5.2 Least Squares with Ordinary Polynomials.

The solution to the Pseudoinverse method yields results that are identical to the traditional least squares procedure. Least squares is most commonly used in the solution of overdetermined sets of linear equations. By the term overdetermined it is meant that a greater number of equations are supplied than the number required to evaluate the polynomial coefficients exactly, i.e. M data points are used to determine K coefficients where M>K. As a consequence of this, an approximating mapping function is obtained as opposed to a function that yields exact results for the data that is submitted. The resulting polynomial equation gives the best approximation among the M control points that are obtained from the calibration grid. The best fit is obtained by minimizing the sum of the squared error values for the M number of control points that are used.

A slight variation to this procedure, which can be used in overdetermined systems, can be to identify control points that contribute more significantly to the total error and to exclude them from a repeated analysis. The iterative process could be repeated until a lower threshold for individual control point error is achieved. In this manner it may be possible to determine polynomial coefficients that more closely fit an increasingly consistent set of data. This approach would overcome singularly spurious control points that may arise due to poor thresholding, or noise present in the calibration image acquisition.

5.3 Least Squares with Orthogonal Polynomials.

As the degree of the fitted polynomial equation increases, K, the number of coefficients increases dramatically. As a consequence of this, the numerical solution of an increasingly large system of linear equations becomes unstable, and inaccurate (Wolberg, 1990), and is further restricted by the possibility of linear dependence between the equations. In this case, direct solution of the linear equations is not the most satisfactory method. The orthogonal polynomial method rewrites Equations(5.1&5.2) in the following manner:

$$x^T = \sum_{i=1}^{K} a_i P_i(x,y) \qquad \text{Equation 5.6}$$

$$x^T = \sum_{i=1}^{K} b_i P_i(x,y) \qquad \text{Equation 5.7}$$

Where: a_i and b_i are the coefficients for the orthogonal polynomials $P_i(x,y)$.

The set of polynomials, $P_i(x,y)$, are said to be orthogonal over a set of points (x,y) if they satisfy the following orthogonality principle:

$$\sum_{k=1}^{M} P_i(x_k, y_k) P_j(x_k, y_k) = 0 \qquad i \neq j \qquad \text{Equation 5.8}$$

The polynomials, $P_i(x,y)$ can be constructed using the Gram-Schmitt orthoganisation incremental procedure described by the relationship:

$$P_K(x,y)=\alpha_{K1}P_1(x,y) + \alpha_{K2}P_2(x,y)+.....+ \alpha_{KK}h_K(x,y) \qquad \text{Equation 5.9}$$

Where the first six linearly independent basis functions, $h_k(x,y)$,are:

$h_1(x,y) = 1$, $h_2(x,y) = x$, $h_3(x,y) = y$, $h_4(x,y) = x^2$, $h_5(x,y) = xy$, $h_6(x,y) = y^2$ ....

By applying the orthogonality principle of equation 5.8 to the polynomials, and by setting $\alpha_{11}=1$, it is possible to calculate the α_{ij} parameters in Equation 5.9 giving the orthogonal polynomials in Equation 5.9.

It can be shown (Wolburg, 1990) that the polynomial coefficients can then be calculated by:

$$a_i = \frac{\sum_{k=1}^{M} x_k^T P_i(x_k, y_k)}{\sum_{k=1}^{M} [P_i(x_k, y_k)]^2} \qquad b_i = \frac{\sum_{k=1}^{M} y_k^T P_i(x_k, y_k)}{\sum_{k=1}^{M} [P_i(x_k, y_k)]^2} \qquad \text{Equation 5.10}$$

In this method the orthogonal polynomials are used in combination with the data supplied from the control points to determine the required coefficients. The advantage of the orthogonal polynomial method is that by using Equation 5.10 it is not necessary to solve a system of linear equations. This method is known to be a faster and more accurate method of performing the least squares function fitting. A second advantage of the method is that from Equation 5.10

it can be seen that additional polynomial terms can be added to the calculation without having to totally re-evaluate the entire calculation. This means that the successive addition of polynomial terms to the equation can facilitate a means by which the error in the fitted model can be controlled, with terms being added until a minimum specified allowable error is achieved. This is not the case of other methods which must be totally recomputed.

5.4 Weighted Least Squares.

The weighted least squares approach is a method that overcomes the uniform error distribution that is a result of the ordinary least squares calculation. In the ordinary method the errors in the fitted function are effectively averaged over the entire image. In certain cases it may be desirable for the error contribution of the control points, that are more distant from the principal point in the image, to be less influential than those situated close to the center of the image. In photogrammetric applications that use the full image for measurement, it is almost certainly desirable to have the error averaged over the entire image. However, for close range robotic machine vision applications, it might be that only the center of the field is of interest. This would be similar to a human visual system that has a higher resolution located around the center or fovea. Therefore in this application, which involves taking accurate measurements from the center of the image, it would be desirable to minimize the error at the center and adopt the weighted least squares approach.

In our application, which involves taking measurements from the entire of the image, it is beneficial to average the error over the entire image and therefore weighted least squares will be mentioned only in passing. It is described in greater detail by Wolburg, 1990.

5.5 The Adopted Approach for fitting the polynomial.

The adopted approach for estimating the Spatial Transform function will be the pseudoinverse solution which is described in Section 5.1. There a two main reasons for this.

The first reason is that this procedure is the method that is incorporated into the MATLAB package for calculating the pseudoinverses of the non-square matrices. This must be evaluated in Equation (5.6) in order to obtain the coefficients for the polynomial.

The second reason is that this method minimizes the sum of squared error in the least squares sense, so that the error attributable to all control points in the calibration grid are averaged over the entire of the image, which is desirable in our application.

Chapter 6
-Fiducial mark position determination

The use of a control or calibration grid is necessary to identify fixed positions in the real world domain so that a correspondence can be established with their positions in the image domain. For this reason it is necessary to accurately determine the position of fiducial marks in the image or the calibration grid.

In the past, researchers have used various methods to provide this information. The nature of the fiducial mark, chiefly the shape, determines how the positional information is determined. Where it may be straight forward to establish the position of an object to the closest pixel by simply selecting the center pixel of a binary image, or choosing the pixel with a maximum or minimum intensity value in a gray scale image, it is possible to more precisely determine an object's position to an accuracy of rather less than a pixel. This concept is termed subpixel resolution or super-resolution.

Various methods have been used to accomplish this, and some of them are summarized by Valkenburg (1994). The method employed depends on the shape of the object used to determine position in the calibration image.

The following Sections 6.1-6.5 outline some of the techniques that have been used by other researchers to define and locate control points in calibration images.

6.1 Analytical Plumb Line Method

Originally developed by Brown(1971), and originally intended for manual measurement of photographic plates, the analytical plumb line method uses a control chart consisting of parallel lines to assess the amount of distortion present in an image. See also Fryer (1986), Fryer & Brown (1986) and Fryer & Mason, (1989).

The method uses the extent to which the lines present in the image of the control chart deviate, from the ideal linear case, to assess the magnitude of radial and decentering distortion present in the cameras optical system.

Ideally such a set of lines in an image will adhere to the relationship:

$$x^T \sin\theta + y^T \cos\theta = \rho \qquad \text{Equation 6.1}$$

Where θ = angle of orientation of the line, and ρ=radial distance from the origin.

Taking into account components of radial (Equations 2.3&2.4) and decentering distortion (Equations 2.7&2.8) the transformed coordinates become:

$$x^T = u + \delta_x(x,y) \;\;\&\;\; y^T = y + \delta_y(x,y) \qquad \text{Equation 6.2}$$

By direct measurement of the image it is possible to determine actual x and y axis components of distortion, $\delta_x(x,y)$ & $\delta_y(x,y)$ that correspond to different positions in the control chart, and it is possible to obtain an observational equation of the form:

$$f(x_{ij}, y_{ij}; x_p, y_p, k_1, k_2, k_3, p_1, p_2, p_3; \theta_i, \rho_i) = 0 \qquad \text{Equation 6.3}$$

In this equation x_p and y_p are the untransformed point positions on the plumbline and x_{ij} & y_{ij} are the plate coordinates in the image of the plumb line that correspond to the j'th measured point on the i'th plumbline. Each observed point generates a linear equation in the form of Equation (6.3) and providing a sufficient number of points are measured on each line, the number of equations will exceed the number of unknowns and a least squares determination of the coefficients can be made.

The method has been suggested by Fryer (1987) for the determination of components of radial and decentering distortion in digital systems in association with video cameras. In this case, the lens systems are very similar to those used in small format non-professional cameras.

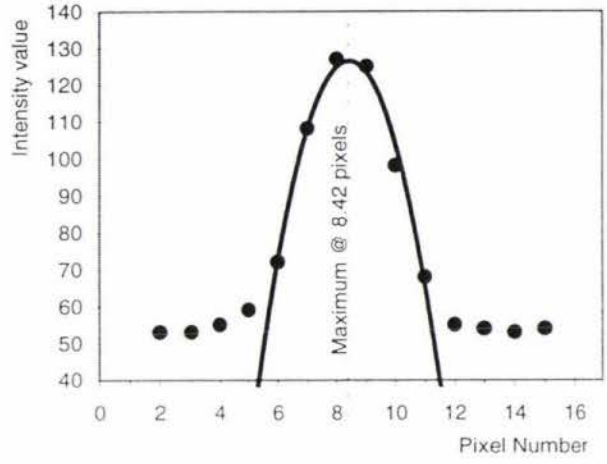


Figure 6-1

The plumbline method has been implemented by Fryer and Mason, 1989, where both line and edge detection routines were used to automatically extract plumbline data, allowing the calculation of the coefficients. In the line detection routine, illustrated in Figure 6-1, Fryer fitted a parabolic curve to the intensity level profile that occurred across a plumb line in order to determine the maximum intensity level and, therefore, the center position. This can be achieved to subpixel accuracy with the estimate of the central position of the line in this example located at 8.42 pixels. In the second method, the authors used an edge detection algorithm to locate the edge of the gray level step associated with the edge of the line. There was found to be no statistically significant difference between the two techniques.

In a similar manner, Seitz (1988), describes the use of parabolic and Gaussian interpolation, as well as center of mass of relative signal methods for determination of edge positions, and introduces a new method of continuous low-pass filtered function reconstruction for high accuracy edge detection and location determination.

6.2 Fourier descriptors and two dimensional moments.

Cross shaped objects have been used by Mikhail, Akey & Mitchell,(1984) as ground targets, or fiducial marks, to precisely determine the position in digitally acquired aerial photographic images.

In this method two principal steps are involved, detection and identification, and position determination.

Firstly, the automatic detection of high intensity cross shaped targets in a gray scale image, is achieved by initially thresholding the image to determine approximate object location. A pattern recognition technique, based upon normalized Fourier descriptors of the objects in the binary image, is then applied to recognize the cross shaped targets.

Secondly, two dimensional moments are applied to the identified cross shaped targets, which because of their two dimensional symmetry, yields the accurate position of the target locations.

The subpixel accuracy achieved by this method is claimed to be as high as 0.03 pixels.

This technique could be applied to the determination of distortion in an image by incorporating the use of a calibration image consisting of a grid of cross shape fiducial marks, providing that the position of these crosses could be accurately controlled on the grid.

6.3 Centroid Calculation Method (CCM)

The Centroid Calculation Method is one that has been widely used for the calculation of the central position of symmetrical objects in grayscale images (Valkenburg, 1994), (Alexander, 1991), (Fillard, 1993), including the center of fiducial marks in camera calibration, and for laser spot and stripe location.

Stanton (1987), has used the centroid method to accurately determine the position of stars for optical tracking using charged-couple devices. In this application, the centroid method was the preferred method due to its properties of simplicity, accuracy and ease of implementation, in particular to symmetrical point or disk shaped objects. The measured performance in this application was typically of the order of 0.005 pixels which was attributed at least in part to the high precision achieved in the manufacture of the CCD sensor.

Following acquisition of the star image, the algorithm performs two one-dimensional first moment calculations that provide the x and y centroid estimates of the stars position within the frame. For each star tracked, a 5x5 pixel window is selected and stored approximately centered around the displayed star. A background intensity value, consisting of the average of the lowest 10 pixel intensities around the edge of the window, is subtracted from each pixel in the windowed image. A minimum pixel value of zero is maintained. Centroids are then recalculated for each window and star position, to achieve an increased accuracy in position measurement.

Valkenburg (1994), notes three problems that can adversely affect the accuracy of the Centroid Calculation Method.

- For circular shaped items in an object, inaccuracy will arise in an image if the calibration disks are not coplanar with the CCD sensor in the camera, with a translation of the apparent centroid occurring due to the altered perspectives. As an example, the centroid shift will be 0.05 pixels if the object plane is angled at 45°, for a 20 mm diameter disk positioned 1m from a camera, focal length f=10 mm and square pixels of size 10μm. This shift can easily be minimized by ensuring that the object is coplanar with the imaging CCD, and is negligible if the object plane is angled to within 5°.

- Any non uniformity in the illumination of the background across the object in the image will cause a bias in the centroid estimate. It is therefore important to remove any intensity gradient that may occur across the image.

- Lastly, it can be shown from theory that the CCM method is only theoretically exact if the image signal is sampled at a rate greater than its bandwidth. This is more fully described in Section 7.1. It is therefore important to either bandlimit the sampling of the image or to design special fiducial marks that are essentially bandlimited in nature.

6.4 The Fourier Phase Shift Method (FPS)

Fillard (1993) formulated the Fourier Phase Shift (FPS) method to determine the center location of a disk. In this method, the phase component, $\phi(s)$ of the Fourier transform of the sampled image or signal is used to calculate $\bar{x}$, a mean value for the x coordinate, from the following relationship:

$$\phi(j) = -2\pi \frac{j}{NT} \bar{x} \qquad \text{Equation 6.4}$$

Where j is an integer representing the sampling frequency, N is the number of sample pixels and T is the pitch.

The FPS method was found to be more accurate and less sensitive to noise than the centroid method although more computationally expensive. The requirements for the FPS method are that the sampling complies with the Nyquist criterion ($f_s > 2f_{bw}$) compared to the centroid method which only required the sampling frequency to be greater than the maximum frequency in the image, $f_s > f_{bw}$, to maintain accuracy.

6.5 Other Methods

In addition to these main methods there are several other variations worth noting.

Parallel lines have also used by Nomura (1992), and circles have been used by Shih (1993,1994). Rectangular grid lines were used by Grosky & Tamburino (1990). Weng (1992), has used the corners of squares followed by determination of intersecting edge locations.

Lyvers, Roberts & Mitchell (1989) have used an edge operator based on two dimensional moments to locate the position of edges in images to within 1/20th of a pixel, but this method is highly computationally expensive.

Tsai (1987) has used interpolation to determine the position of a 'true' edge for multiple points and then performed linear regression to fit a line to each edge of a diamond and then solved each equation to determine the intersections which determine the position of the control points which occur at the corners of the diamonds.

6.6 The adopted approach - CCM in detail.

The preferred method for fiducial mark location determination in this project was chosen to be the Centroid Calculation Method (CCM). There were three main factors that were responsible for this decision.

- In construction of the calibration chart, It is important to be able to construct fiducial marks easily and position them accurately in the calibration chart. In practice this was achieved by using the CorelDraw computer software package and the printing of the calibration charts accomplished on a Hewlett Packard LaserJet 5MP which has a nominal resolution of 600 dpi. Typically, with a field size of approximately 180 mm and an image resolution of 512x512 pixels, a positional precision of about 0.125 pixels was achieved.
- Use of the Centroid Calculation Method allows the use of filled circle shapes as fiducial marks. The use of these discrete objects, compared to plumblines or a hatch grid, minimizes the computation required to extract position information related to the location of each mark. Once the region of discrete support is identified, that is an area containing a single fiducial mark, by thresholding, implementation of the CCM is straight forward.
- Aliasing problems due to low image sampling rates is minimized with the CCM when compared to other techniques such as the FPS method that requires satisfaction of the Nyquist criterion, CCM can tolerate moderate levels of undersampling.
- The high frequency content or maximum bandwidth of a dot grid may be less than a line grid and hence any problem with undersampling may be further alleviated.

6.6.1 Description

Alexander and Ng (1991), have described in detail the conditions required to eliminate systematic error in subpixel determinations using the Centroid Calculation Method.

Consider first the one dimensional case involved when determining the center of a stripe or a line. The waveform f(x) is a representation of the line profile and is the result of the convolution of a pixel response or sensitivity function and the actual line intensity profile. This function is illustrated in Figure 6-2(A).

When the line is imaged by a CCD array, the waveform f(x) is effectively sampled at fixed points determined by the physical construction of the CCD sensor and the combination of factors concerned with the optics and camera geometry. This sampling function is illustrated in Figure 6-2(C).The sampling function s(x) is multiplied by the actual waveform f(x) in order to obtain the sampled waveform function g(x) which is illustrated in Figure 6-2(E).

Providing the pixel response profile is symmetrical, the centroid (or center) of the function f(x) will be the same as the centroid of the original line intensity profile.

Assuming that this is the case, then the centroid, $\bar{x}_f$, of the function f(x) can be determined analytically to be:

$$\bar{x}_f = \frac{\int_{-\infty}^{\infty} x f(x)\,dx}{\int_{-\infty}^{\infty} f(x)\,dx} \qquad \text{Equation 6.5}$$

The centroid estimate of the sampled function, $\bar{x}_g$, is given by the expression:

$$\bar{x}_g = \frac{\int_{-\infty}^{\infty} x g(x)\,dx}{\int_{-\infty}^{\infty} g(x)\,dx} \qquad \text{Equation 6.6}$$

It is necessary to determine the conditions for which the centroid of the sampled waveform, g(x), is equivalent to the centroid of the original waveform, f(x). This is best understood by examining the Fourier transforms of the two functions, which are defined in the following equations.

$$\bar{x}_f = \frac{\int_{-\infty}^{\infty} x f(x)\,dx}{\int_{-\infty}^{\infty} f(x)\,dx} = -\frac{F'(0)}{2\,pjF(0)} \qquad \text{Equation 6.7}$$

$$\bar{x}_g = \frac{\int_{-\infty}^{\infty} x g(x)\,dx}{\int_{-\infty}^{\infty} g(x)\,dx} = -\frac{G'(0)}{2\,pjG(0)} \qquad \text{Equation 6.8}$$

From Equations (6.7 & 6.8), it can be seen that the centroid of f(x) depends only on the value of the Fourier transform of the function, F(s) and F'(s), at the origin, F(0) and F'(0). It follows that if F(0)=G(0) and F'(0)=G'(0) then:

$$\bar{x}_f = \bar{x}_g . \qquad \text{Equation 6.9}$$

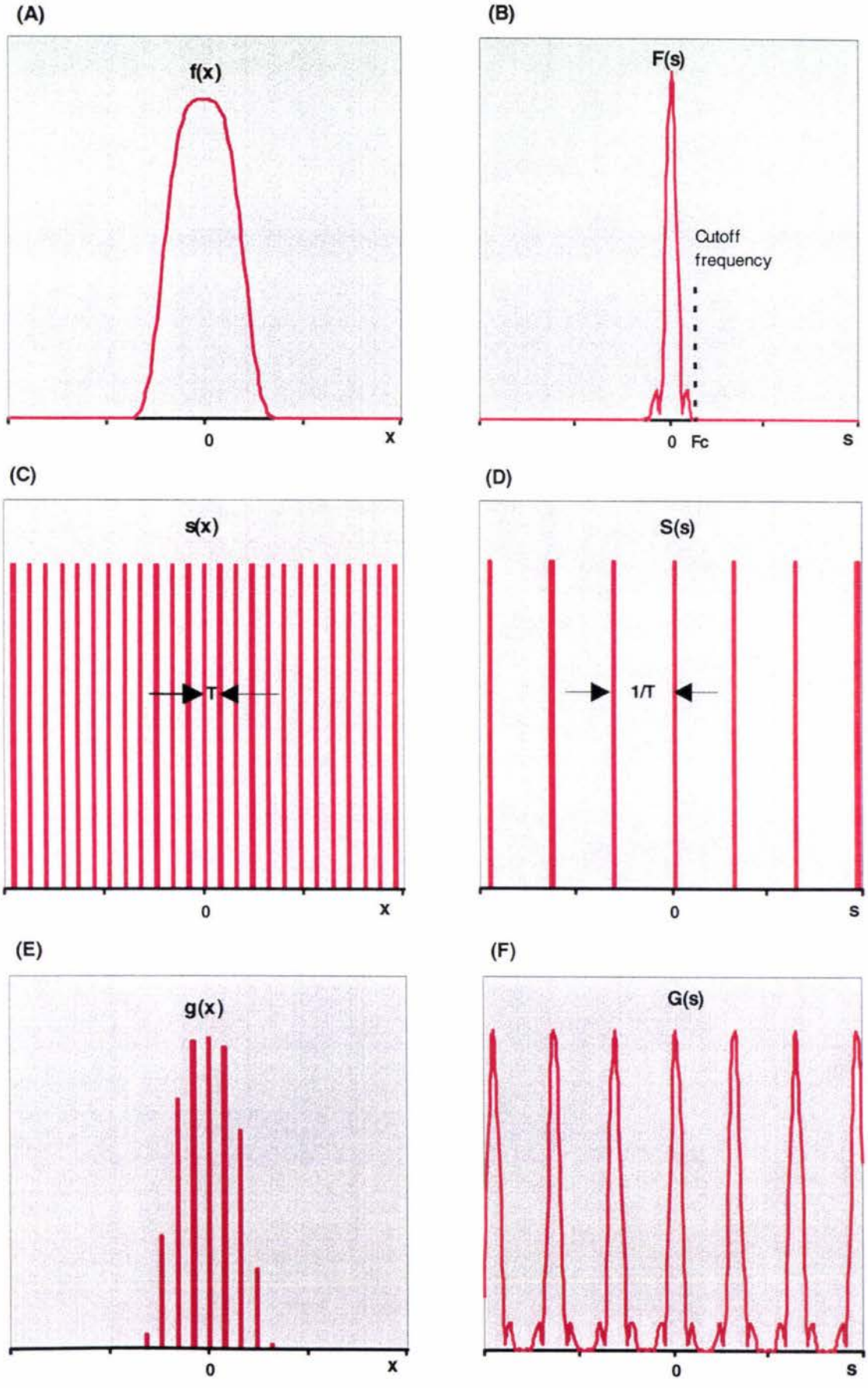

(A)
f(x)
0
x
(B)
F(s)
Cutoff
frequency
0
Fc
s
(C)
s(x)
T
0
x
(D)
S(s)
1/T
0
s
(E)
g(x)
0
x
(F)
G(s)
0
s

Figure 6-2

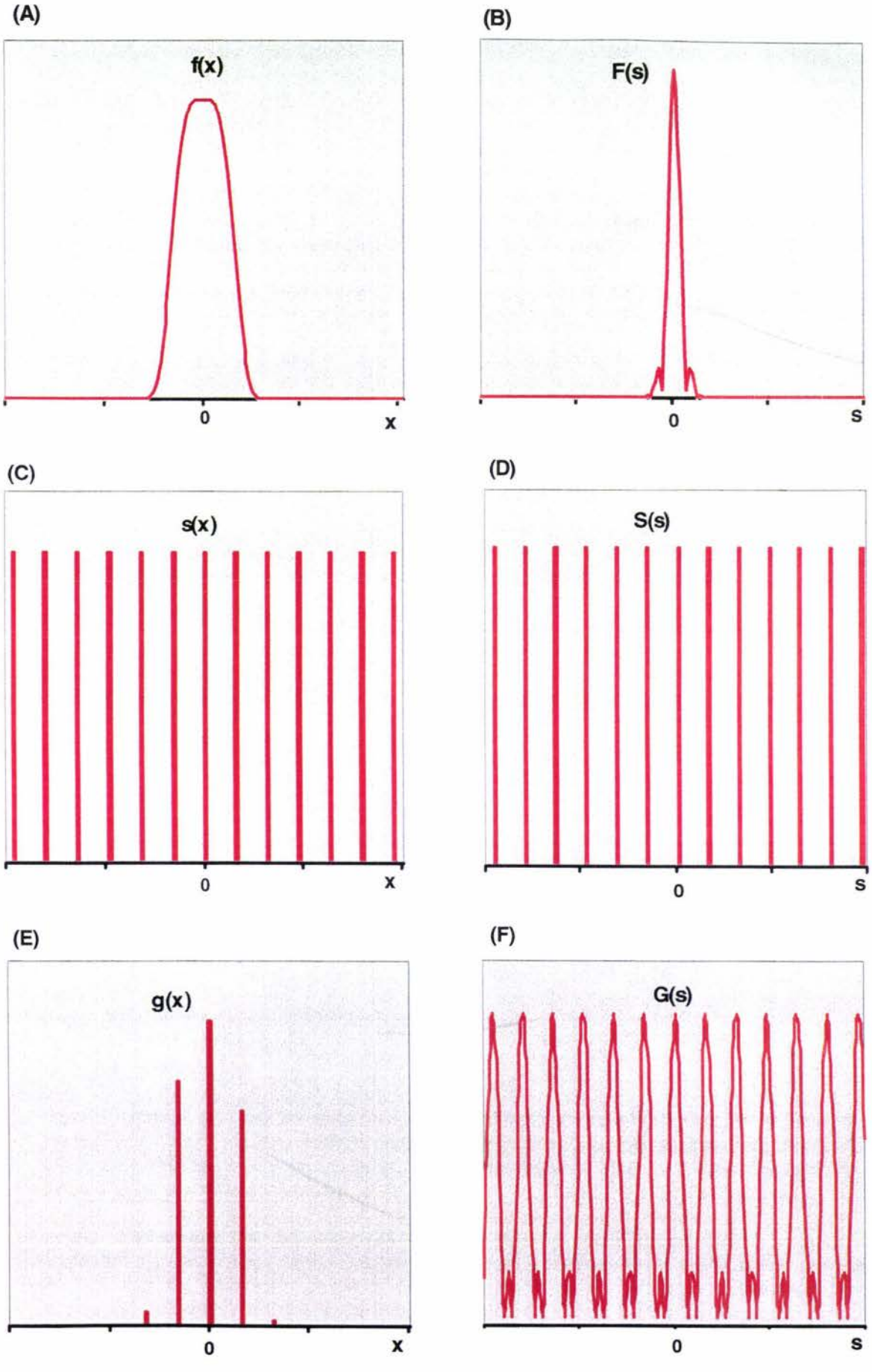


Figure 6-3

This observation is illustrated graphically in Figure 6-2. Figures 6-2(A,C,E) show the intensity profile f(x), the sampling regime s(x) and the sampled waveform g(x) plotted in the normal spatial domain. Fourier transformation of these functions allows them to be plotted in the frequency domain in Figures (B,D,F) respectively.

In Figure 6-2(B), the plot of the frequency spectrum of the intensity profile, a cutoff frequency, F_c, is defined as the maximum frequency component occurring in the intensity profile signal.

In Figure 6-2(C), a graph of the sampling function, f(x), the sampling interval, T, in pixels, is shown. The corresponding plot of this function in the frequency domain, Figure 6-2(D), shows spectral components (or harmonics) of the sampling function separated by the sampling frequency, 1/T.

It can be seen that providing the sampling frequency $1/T > F_c$, then higher frequency components of the waveform will not interfere with the value G(0) in Figure 6-2(F). In this instance, Equation 6-10, derived from Equations 6.7 & 6.8 is satisfied.

$$\frac{F'(0)}{F(0)} = \frac{G'(0)}{G(0)} \Rightarrow \overline{x}_f = \overline{x}_g \qquad \text{Equation 6.10}$$

Figure 6-3 shows an example of the case for when $1/T = F_c$. In this situation, it can be seen that the higher frequency components of the sampled waveform are starting to impact on the value of G(s) at the origin, in Figure 6-3(F). When this happens, systematic error is introduced into the value of the estimated centroid of the function f(x), and from Equation 6.10, $F(0) \neq G(0)$ and $F'(0) \neq G'(0)$ and $\overline{x}_f \neq \overline{x}_g$

Alexander and Ng (1991), have suggested a practical way in which systematic error can be avoided when implementing the Centroid Calculation Method. This involves changing the camera lens aperture in order to restrict the bandwidth of the image.

Choice of lens aperture can be used to ensure that the maximum frequency, F_c, present in the image is less than the sampling frequency, 1/T. For a diffraction limited system, such as is the case with a lens, it can be shown that for an ideal lens, the modulation or effective contrast in an image will become zero at a given cutoff frequency, Walker (1994).

$$F_C = \frac{A}{\lambda f} \qquad \text{Equation 6.11}$$

where λ = the wavelength of monochromatic light.
f = the focal length
A = aperture

From Equation 6.11 it can be seen that the cutoff frequency can be limited by reducing the physical aperture of the lens system. Using a high F-stop (e.g. f22) will bandlimit the spatial frequencies present in the image to a point where Equation 6.9 is satisfied allowing the centroid to be accurately calculated.

Another method that may be employed to ensure that the sampling condition is satisfied is suggested by Valkenburg (1994). The design of the fiducial marks in the calibration image can be such that the calibration image is essentially bandlimited. This could be achieved by varying the properties of the fiducial marks. Fiducial mark size, frequency (or number), and texture all contribute to the spectral energy present in the calibration chart. The bandwidth, or cutoff frequency, of the chart will increase will a reduction in size of fiducial marks, the increase in number of fiducial marks and an increased step size in intensity level at the margins of any object in an image.

Chapter 7
-Calibration of an image in practice

This chapter describes the practical procedure that has been developed to calibrate a video camera-lens system positioned on a stand for mensuration of images (shown in Figure 1-1). A summary of the procedure is shown in Figure 7-1.

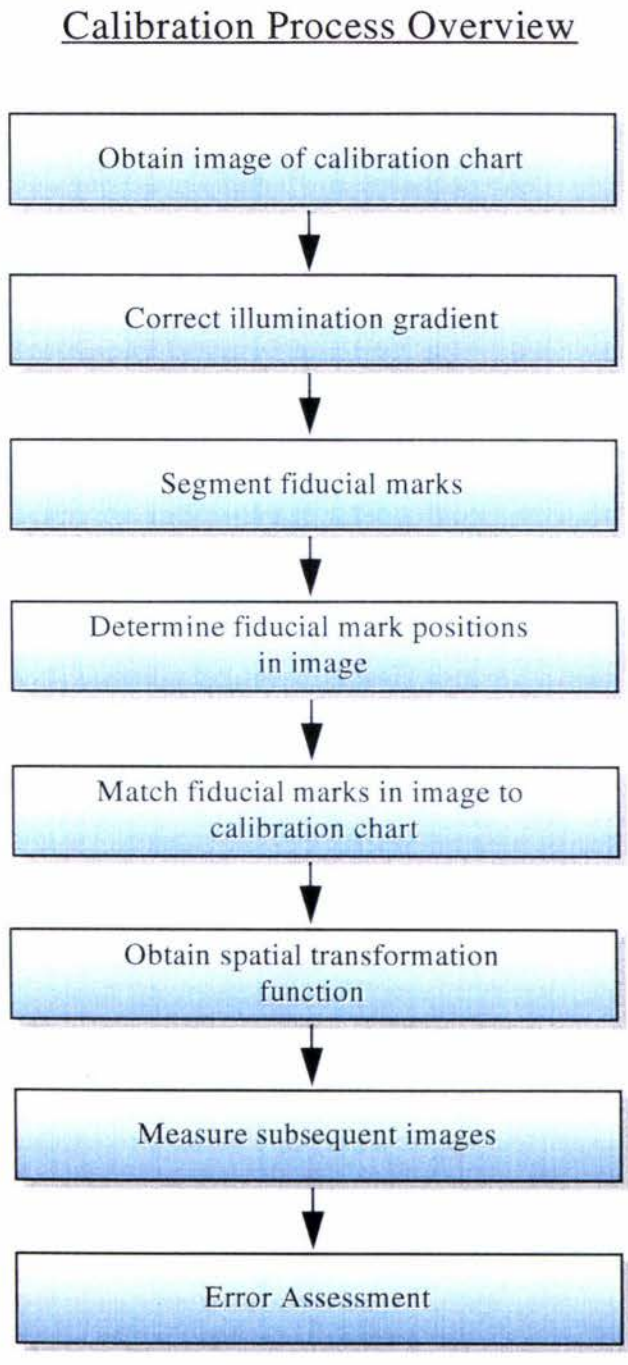


Figure 7-1

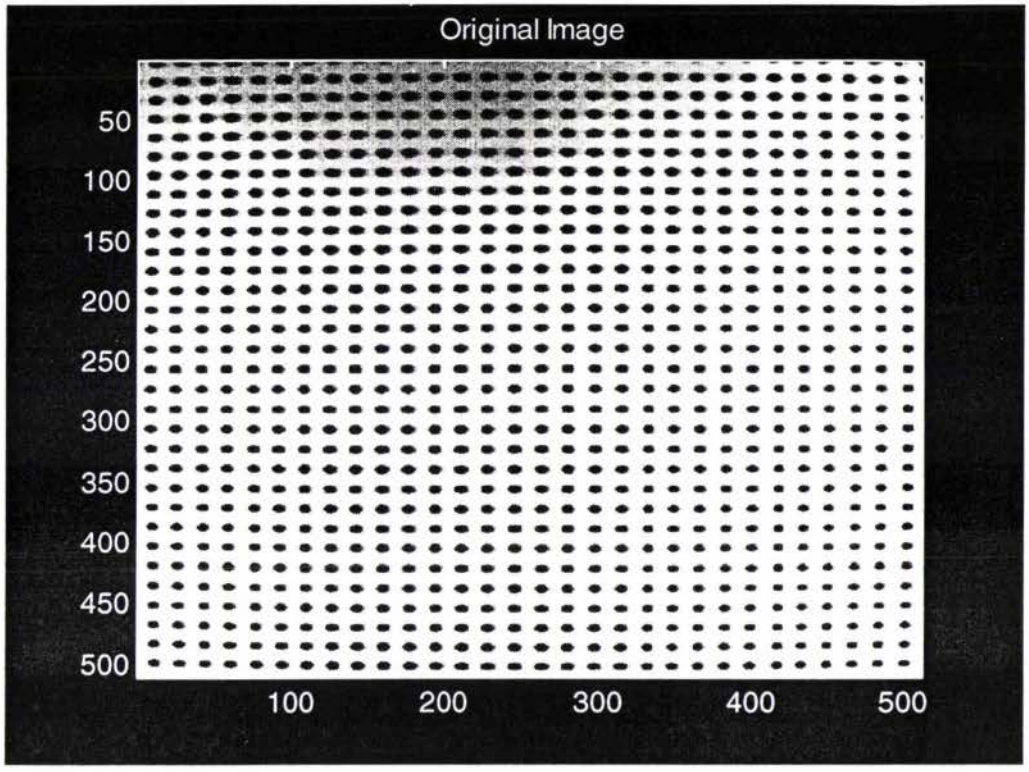


Figure 7.2

7.1 Image Acquisition.

In the initial step, the camera, a Sony b/w CCD video camera (Model XC-77CE), was adjusted on the stand in the fixed position in which the images were to be captured. The lamp angles were adjusted to give a uniform lighting distribution across the chart as well as a suitable exposure setting for the selected aperture setting. The object, in this case the calibration chart, was placed under a transparent sheet of glass to ensure flatness and the absence of the terrain distortion, mentioned in Chapter 3. Camera height was fixed to give a field size of approximately 18x18 cm. Grayscale images were captured from the video camera on a Kontron-Vidas image analysis system in a 512x512 pixel format, and stored for subsequent processing. The use of filled circles as fiducial marks in the calibration chart, as described in Chapter 6 section 6, means that in the image, fiducial marks will be represented by areas of pixels having low intensity values and background will be represented by pixels having high intensity values. To maximize the value of information in an eight bit intensity (grayscale) image it is desirable for the brightest pixel in image to have an intensity value of 255 and the darkest pixel to have a value of 0. This can be accomplished by careful manipulation of illumination conditions and control of aperture settings.

For subsequent processing of the image it is of benefit to have fiducial marks represented by pixels having high intensity values and the background, low intensity values. For this reason, captured images undergo a grayscale inversion operation as a preprocessing step.

Figure 7-2 shows a typical captured image of a calibration chart prior to intensity inversion. This image illustrates the phenomenon of both barrel distortion and non uniform illumination.

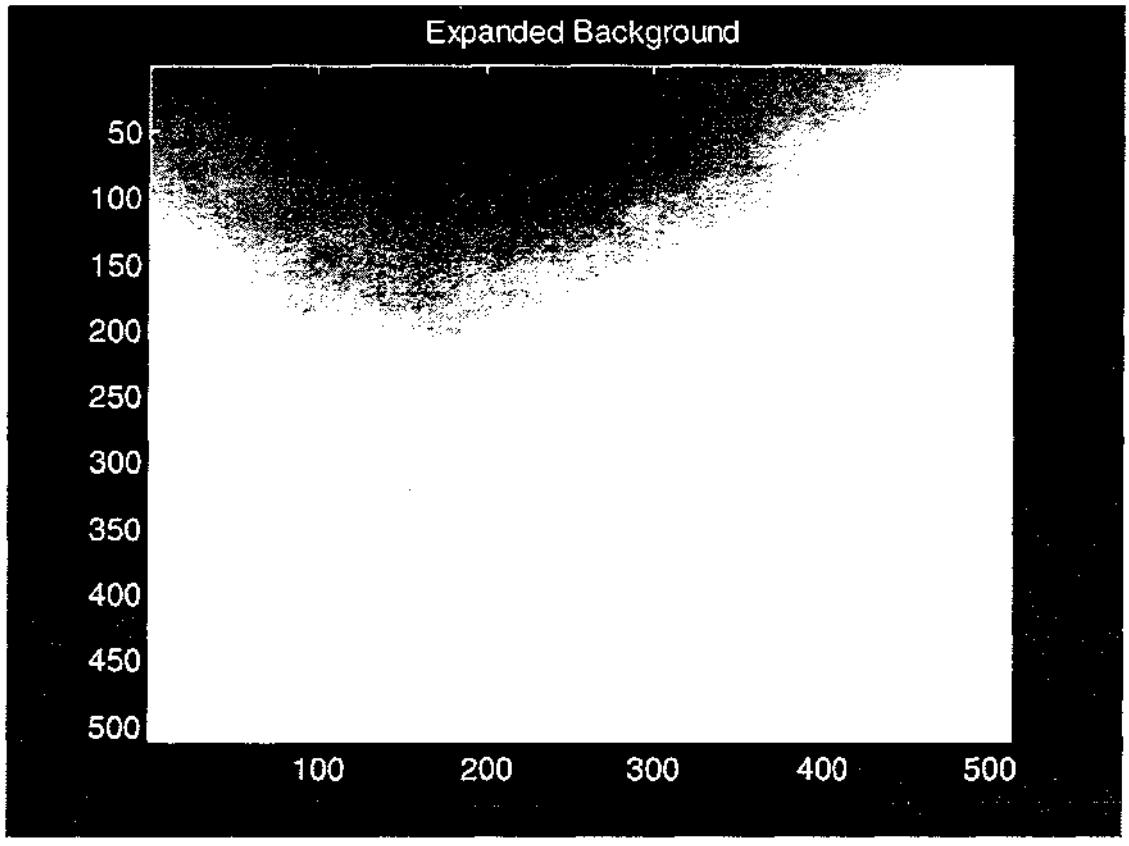


Figure 7-3

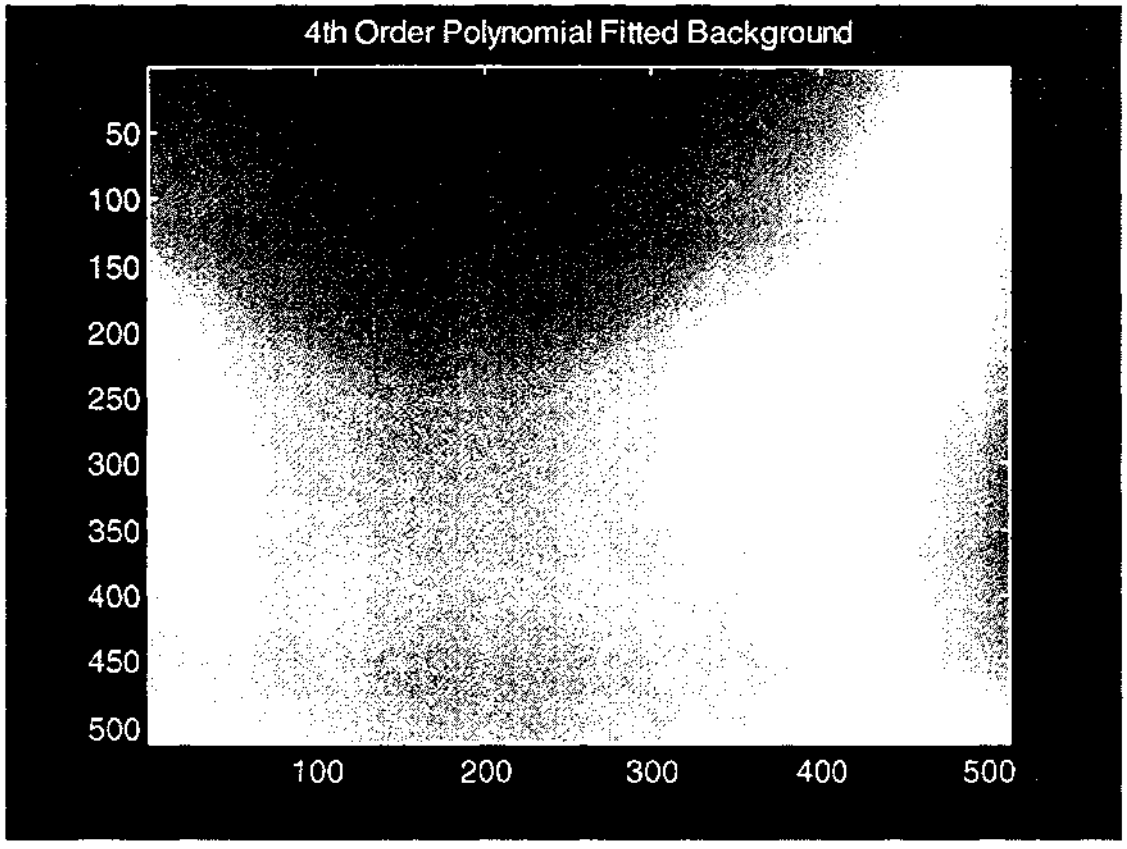


Figure 7-4

7.2 Remove illumination gradient

As discussed in Chapter 6 (section 3), non uniform illumination can affect the accuracy of the centroid calculation method which is used to accurately calculate the center position of the fiducial marks. In order to minimize this affect, two procedures were tested to eliminate any variation in the image due to variation in illumination across the image. The first method involved the capture of a second image consisting only of a blank background. The subtraction of this background image was found to substantially reduce the problem. This background image is shown in Figure 7-3, with expanded contrast to demonstrate the non-uniform illumination conditions that exist.

The second method operates solely on the one image of the calibration grid. In this method the image is divided up into 256 equally sized sub windows and a 2 dimensional 4th order polynomial function was fitted to the minimum background intensity level that occurred within each window. This function was then evaluated for every position in the image, the fitted background is shown in Figure 7-4. In a similar manner, the fitted background can then be subtracted from the image of the calibration grid to give a leveled image.

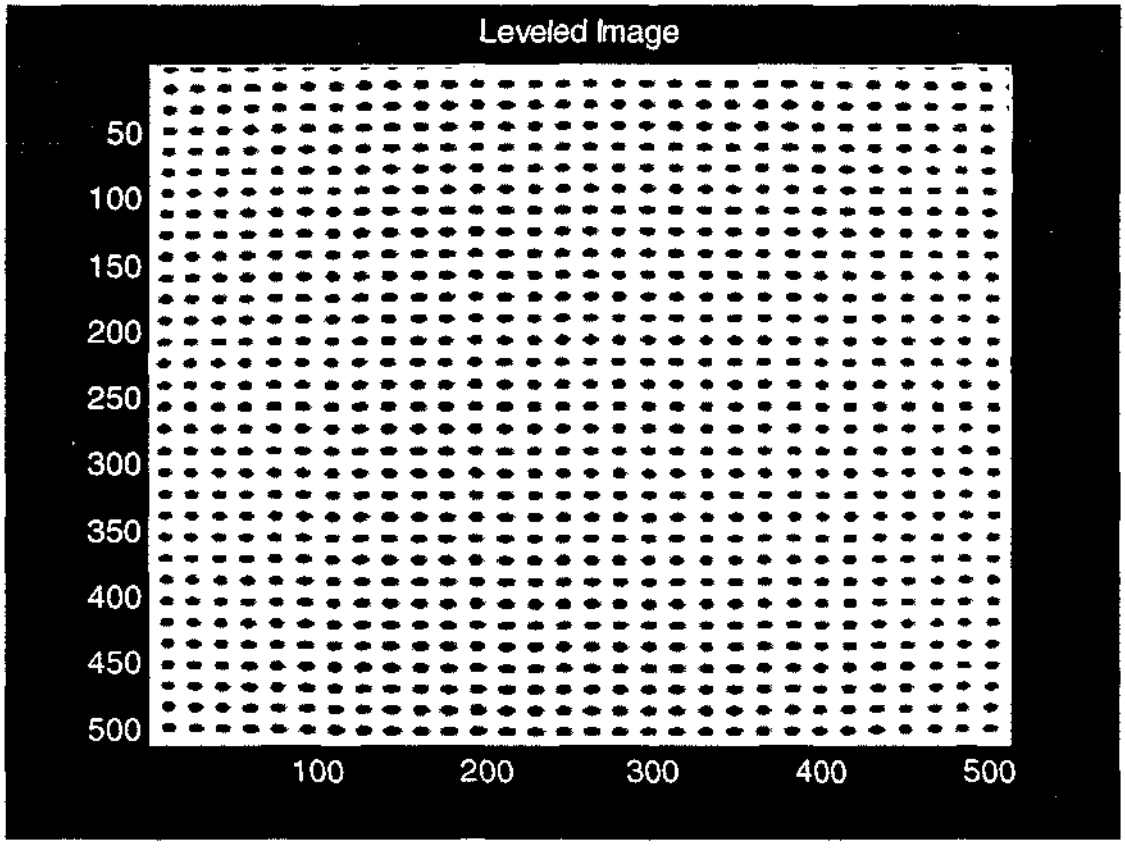


Figure 7-5

In both cases, when the background image is subtracted from the calibration chart, a visibly improved, leveled image can be obtained, and is shown in Figure 7-5. The disadvantage of the background image method is the extra labor required at the image acquisition stage. On the other hand far greater computation is needed in order to calculate a polynomial fitted background from a single image of the calibration chart.

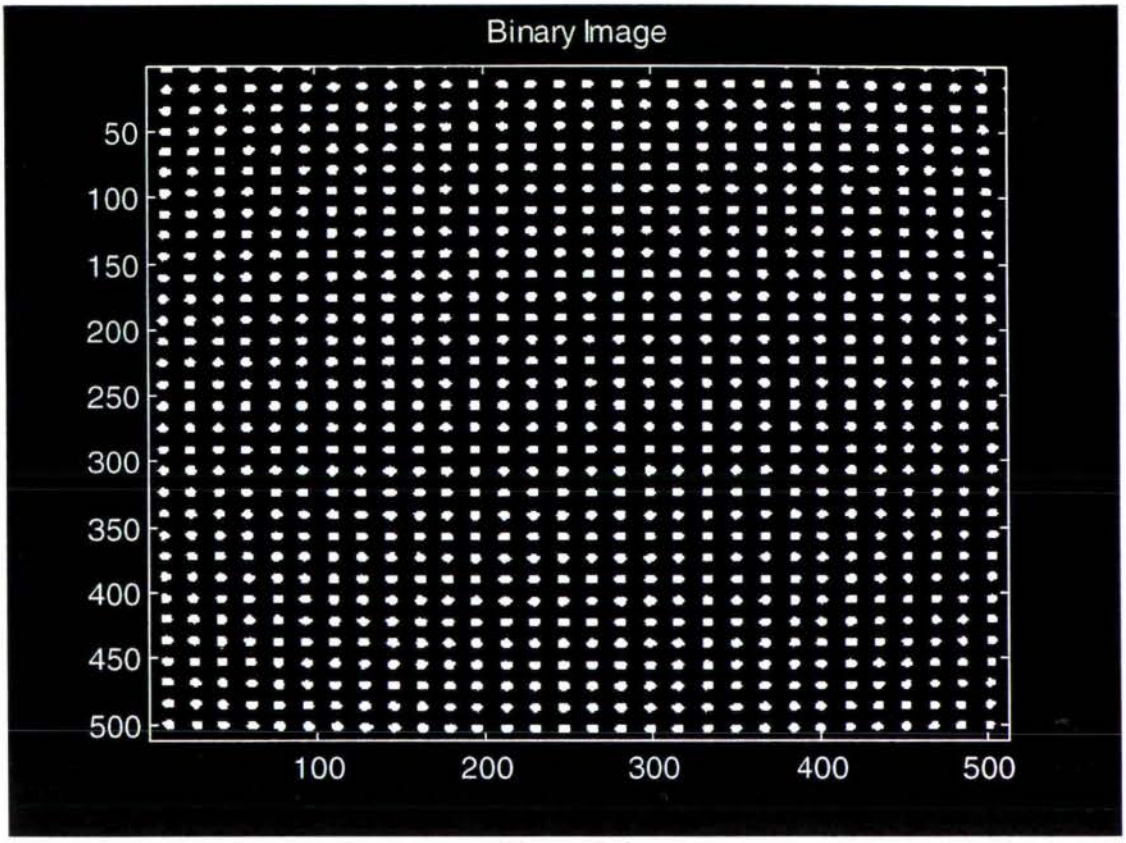


Figure 7-6

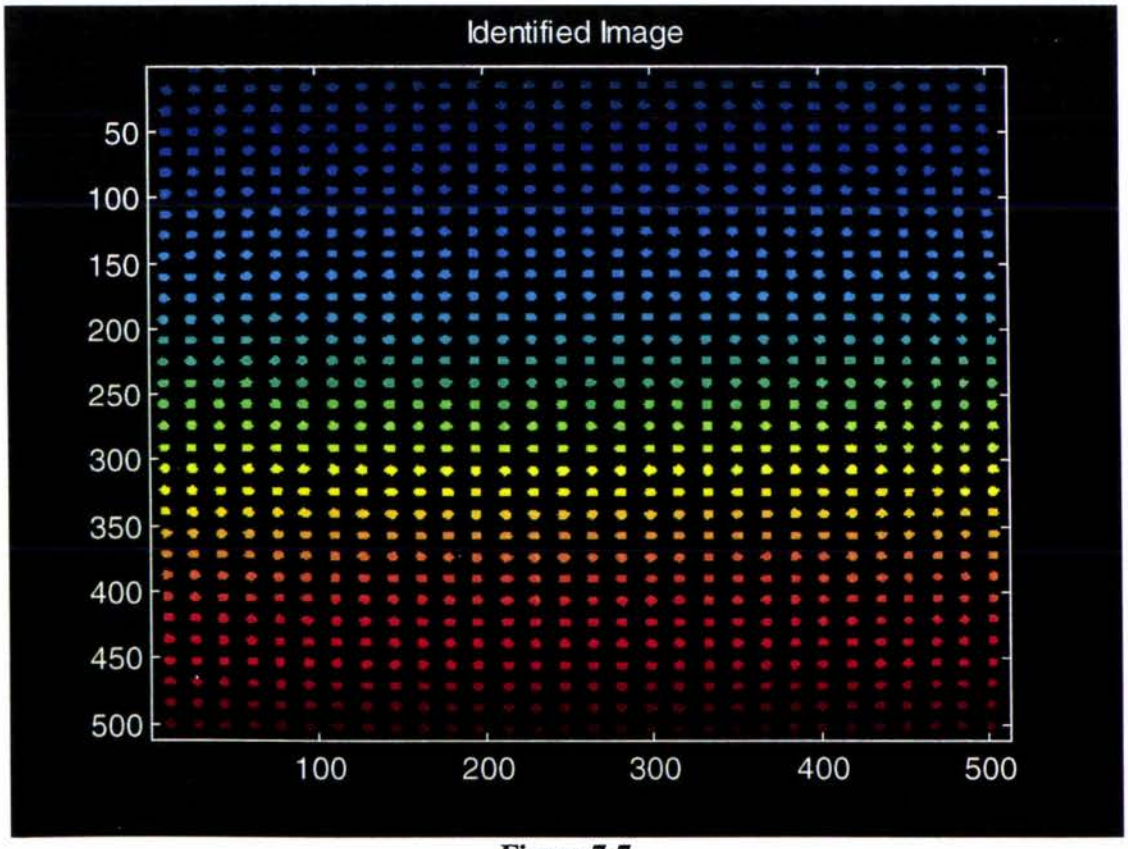


Figure 7-7

7.3 Segment fiducial marks

It is necessary to identify the region of interest associated with each fiducial mark as it appears in the image. This was achieved by globally thresholding the contrast expanded image at an intermediate intensity threshold of 125. The result of this is shown in Figure 7-5. As an alternative, a local thresholding procedure can be used to overcome problems that global thresholding has with excessive illumination gradients. In such a procedure, each pixel in the image is compared to mean intensity value determined from a window centered about the current pixel.

The binary image that results from the thresholding process consists of blobs associated with each of the fiducial marks. A MATLAB procedure was written to identify individual fiducial marks and to assign a unique label to each. This result of this process is illustrated in Figure 7-7, where false color is used to easily identify the discrete fiducial marks.

7.4 Determine fiducial mark position

Having located each fiducial mark, it is then possible to determine an estimated center position by calculating 1st order binary moments. This is accomplished by successively isolating each fiducial mark by obtaining a window around the central location and creating a subimage. The size of the square window was determined dynamically to be average area occupied by a fiducial mark. This is calculated to the square root of the total image size in pixels divided by the number of fiducial marks in the image.

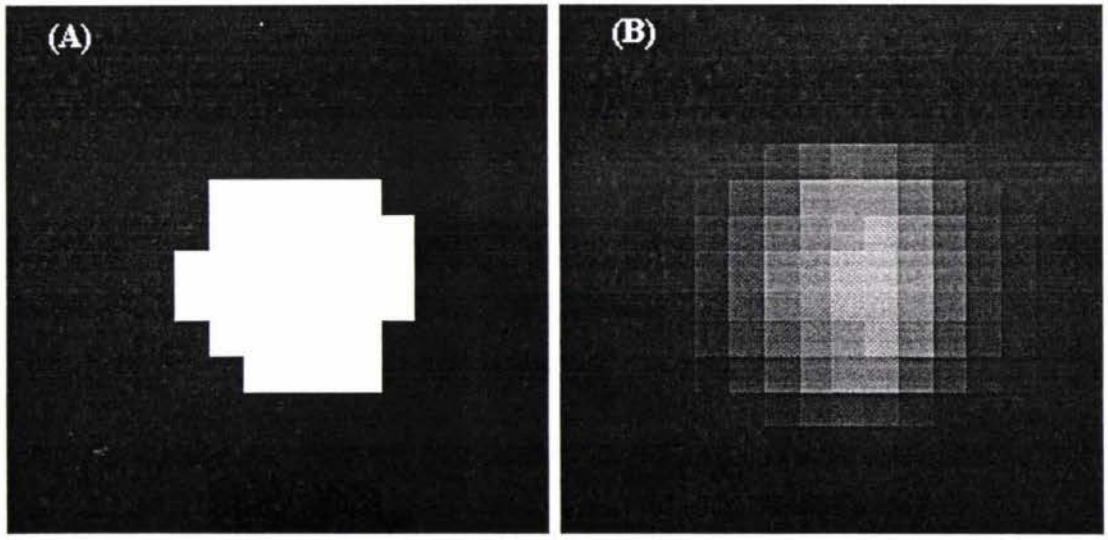


Figure 7-8

Having separated the regions of interest for each fiducial mark from the image, the Centroid Calculation Method (CCM), described in Chapter 6, was applied to each subwindow of the inverted, leveled image of the calibration chart. This allowed an improved estimate for the center positions of the fiducial marks in the image to be determined. Due to the finite size of the window, a non zero value for pixels present in the background will bias the calculated center position. For this reason it is necessary to subtract the minimum pixel intensity value from all pixels in the subimage. The desired effect of this is to set the background to zero.

One such window is shown in Figure 7-8. Figure 7-8(A) shows the binary image of the mark which was used to obtain an approximate center location of the fiducial mark, and Figure 7-8 (B) shows the gray scale intensity image that was used by the CCM in order to determine a grayscale weighted approximation of the centroids.

7.5 Match fiducial marks image positions to calibration grid.

Once all the fiducial mark positions have been determined accurately in the image it is necessary to match these coordinates in the image plane to their actual physical locations in the real world, two dimensional plane. The calibration chart consisted of a series of small filled circles whose centers were situated precisely at specific distances on a Cartesian plane like grid. Because the fiducial marks are all regularly spaced in the chart, they can be assigned coordinate values in (x,y) starting from (0,0) at the arbitrary center and progressing outwards towards the outside of the image. A specialized algorithm was designed to assign these real coordinate positions to the fiducial marks in the image which are not located in the same linear fashion due to the combined effects of lens and geometric distortions discussed in Chapters 2 & 3.

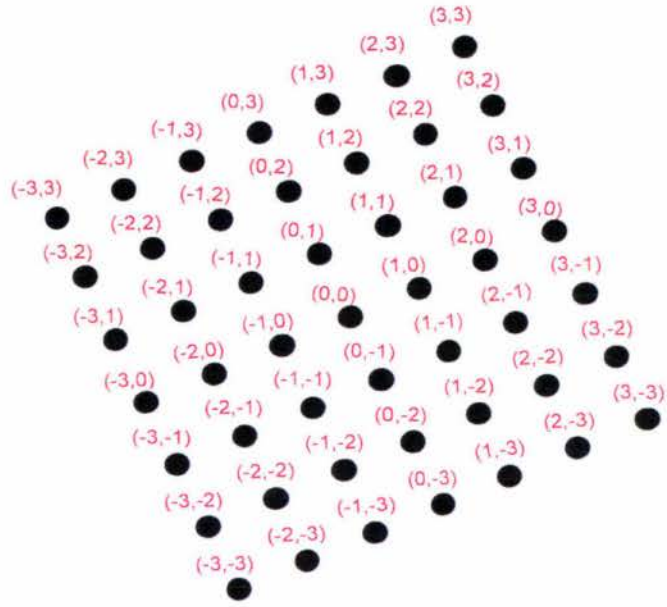


Figure 7-9

Coordinate assignment is illustrated in Figure 7.9 and was accomplished reliably by following a number of simple rules, implemented in the form of a MATLAB algorithm.

The fiducial mark located closest, by calculated Euclidean distance, to the center of the image was assigned as point (0,0). The second closest point to the center was assigned point (0,-1). Of the next three fiducial marks closest to the mark allocated (0,0), the mark most distant from (0,-1) is assigned the label (0,1). The remaining of the two points are arbitrarily labeled (-1,0) & (1,0).

The remaining fiducial marks in the calibration image are classified sequentially in increasing order of distance from the central mark (0,0). This is achieved by the fact that each new mark will already have a minimum of two previously classified neighbors. By examination of these neighbors it is possible to simply infer the correct classification of the current fiducial mark. In certain cases, due to fiducial marks occurring in the outer corners of the corners of the image, and also the need to reject fiducial marks that fall very close to the image edge, it sometimes happens that these points do no have 2 previously classified neighbors. If this situation arises it can lead to an incorrect classification of the current fiducial mark. For this reason the four

outermost fiducial marks that have been located using the Centroid Calculation Method are ignored.

By knowledge of the initial spacing of adjacent fiducial marks on the calibration chart (e.g.) 5 mm spaced grid, it is therefore possible to obtain two coordinate positions for each fiducial mark (i.e.) the untransformed coordinate position in 2D real world plane (x,y) and the subpixel transformed coordinates (x^T,y^T), in the spatially transformed image plane.

7.6 Obtain spatial transformation function

With the untransformed and transformed data available for every relevant fiducial mark in the image, it is possible to fit an N'th order polynomial equation to the data describing the spatial transformation function, as described in Chapter 5. By reversing the data sets in the analysis, it is possible to also derive the polynomial equation that describes the inverse transformation function. Given this equation, it is possible to calculate the actual real world positions of objects located in subsequent images that are acquired using the same camera/lens system setup. Examination of the residual error at this stage is an indication of how well the fitted polynomial function describes the combined distortion in the camera lens system.

7.7 Error assessment

It is straightforward to determine the error associated with fitting the polynomial in the least squares sense to the two sets of data. Obtaining an error term for the entire process is more difficult.

However, by re-imaging the calibration grid in a different orientation or by using a second calibration grid and applying the derived inverse transformation function, it is possible to obtain an estimate of the overall error in the calibration due to the both the Centroid Calculation Method and that involved in fitting the polynomial function. This can be achieved for both the x and y fitted equations, described by Equations 5.1 & 5.2, and a mean Euclidean error term calculated from the following relationship.

$$\varepsilon_{xy} = \sqrt{\varepsilon_x{}^2 + \varepsilon_y{}^2} \qquad \text{Equation 7.1}$$

where ε_{xy} = the Euclidean error

ε_x = x-axis component error

ε_y = y-axis component error

A mean Euclidean error term $\overline{\varepsilon_{xy}}$ can be calculated by averaging the individual Euclidean error term associated with each reverse checked fiducial mark position. Using this approach it was possible to investigate how the absolute error varied when using different calibration grids in terms of fiducial mark diameter. The error obtained at different camera aperture settings was also investigated.

The error was assessed in terms of maximum error present at any position in the image and also the RMS error across the entire of the image calculated at each fiducial mark location.

Chapter 8 -Results

8.1 Order of the fitted equation

Images were captured at an aperture setting f22 and image leveling was carried out by background image subtraction. The Centroid Calculation Method was used to determine the subpixel positions of the 3 mm diameter fiducial marks on a 10 mm evenly spaced grid. Polynomial equations were fitted to obtain the inverse spatial transformation function. As previously described in Chapters 3&4, two equations were obtained, one relating to the transformation of the x coordinate and the other to the y coordinate. The absolute Euclidean error of the residuals was calculated for each position in the calibration grid for polynomial equations fitted for first, second, third and fourth order polynomial functions. These results are displayed graphically in Figures 8-1, 8.3, 8.5, 8.7 respectively. In these graphs the magnitude of the absolute error is expressed on a color scale that ranges from black for low values of error through red, orange, yellow and finally white for the maximum error present on the grid. Positions of fiducial marks on the grid that occur too close to the image margins for analysis or are rejected by the distance criterion established in Chapter 7, are displayed in blue.

The manner in which the absolute Euclidean error varies with increasing radial displacement from the center of the calibration grid, in each case, is depicted graphically in Figures 8-2, 8-4, 8-6 & 8-8.

It can be seen from Figures 8-1 to 8-4 that major trends are present in the residuals of the first and second order polynomial functions. This indicates that these fitted equations do not explain the variation present in the positions of the transformed fiducial mark coordinate positions in the image of the calibration grid. Examining the mean and maximum absolute Euclidean errors of the residuals in each of these cases, Table 8-1, shows that only a small improvement in fit is obtained by fitting second order equations to the data, over a first order fit. A large improvement is achieved by fitting third order equations and a marginal improvement is obtained by fitting the fourth order polynomial equations.

Order of equation	Mean absolute error, $\overline{E}_{xy}$ (mm)	Maximum absolute error, E_{xy}^{max} (mm)
1°	0.4423	1.8170
2°	0.4034	1.6070
3°	0.1159	0.2530
4°	0.1148	0.2530

Table 8-1

In Figures 8-5 to 8-8, it can be seen that for both the third and fourth order cases, the residuals obtained appear to be randomly distributed and lack the obvious trends that were present for the lower order fitted polynomials. When examining the variance of residuals in these graphs it must be noted that due to the intrinsic nature of the square grid , fewer data points exist for both low values of radial displacement from the center, and higher values of radius where the fiducial marks are situated in the corners of the grid.

From the evidence determined experimentally, it would appear that the fourth order fitted polynomial functions are marginally better than the third order case.

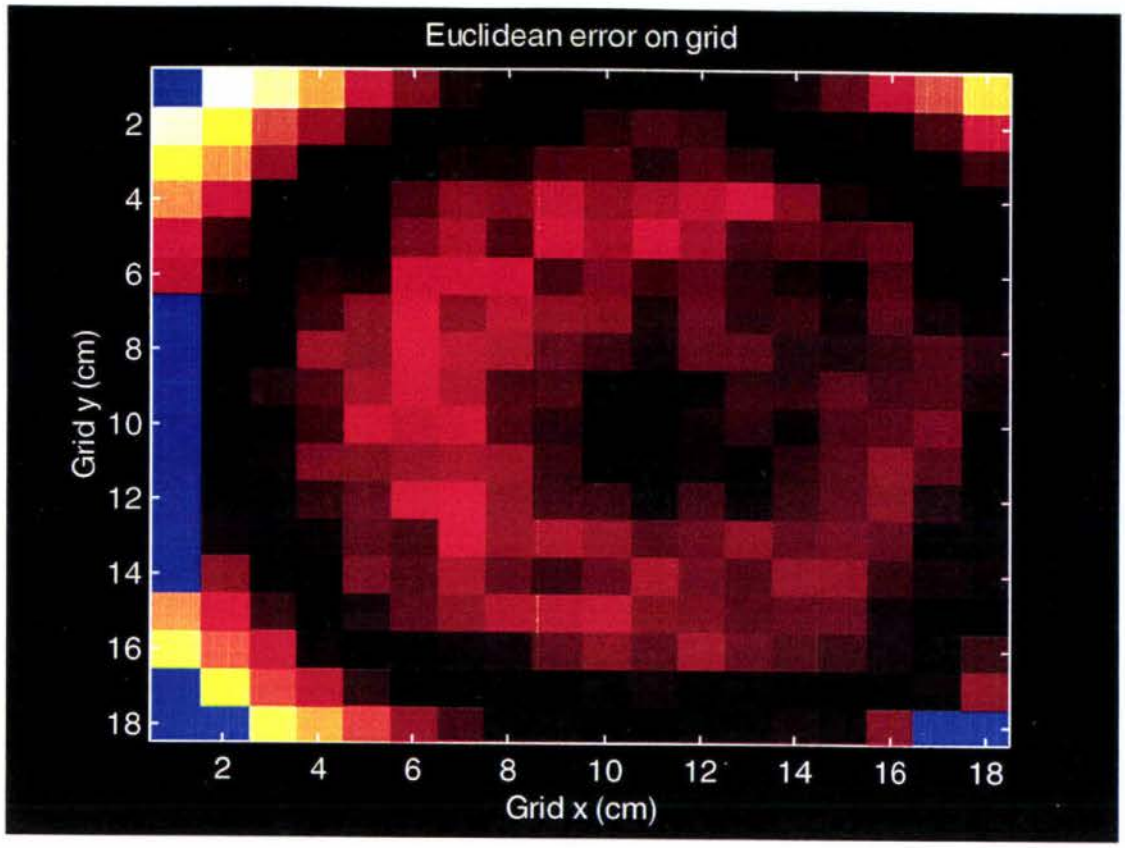


Figure 8-1 (First order polynomial fit)

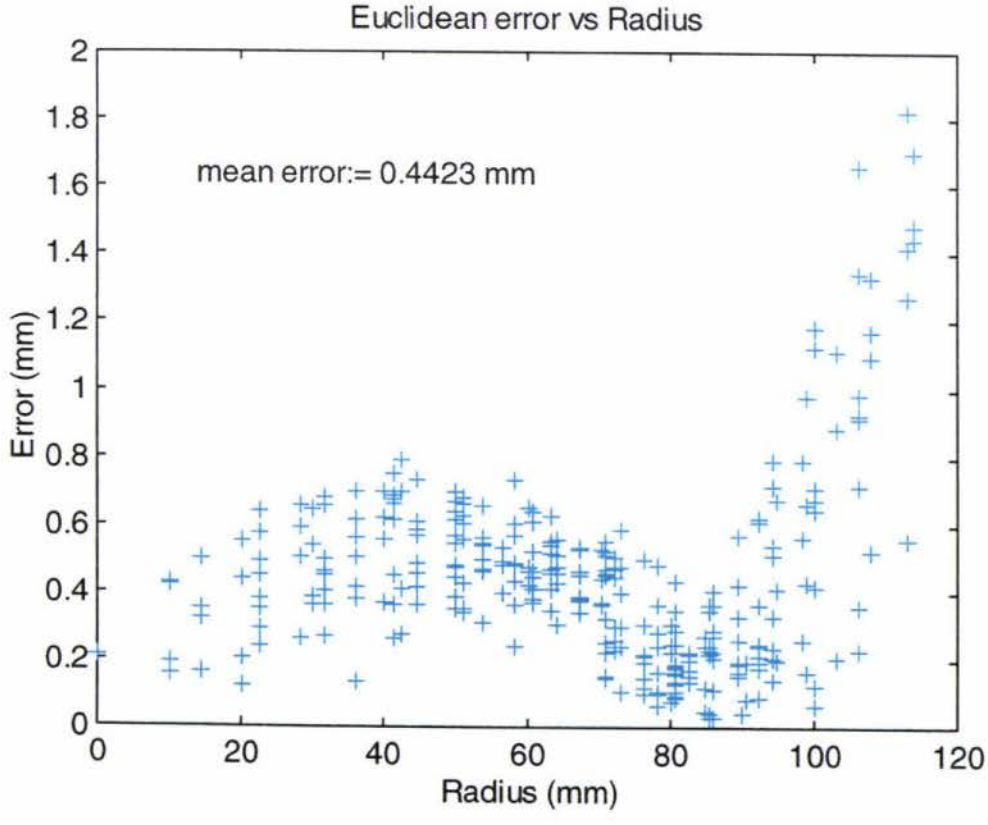


Figure 8-2 (First order polynomial fit)

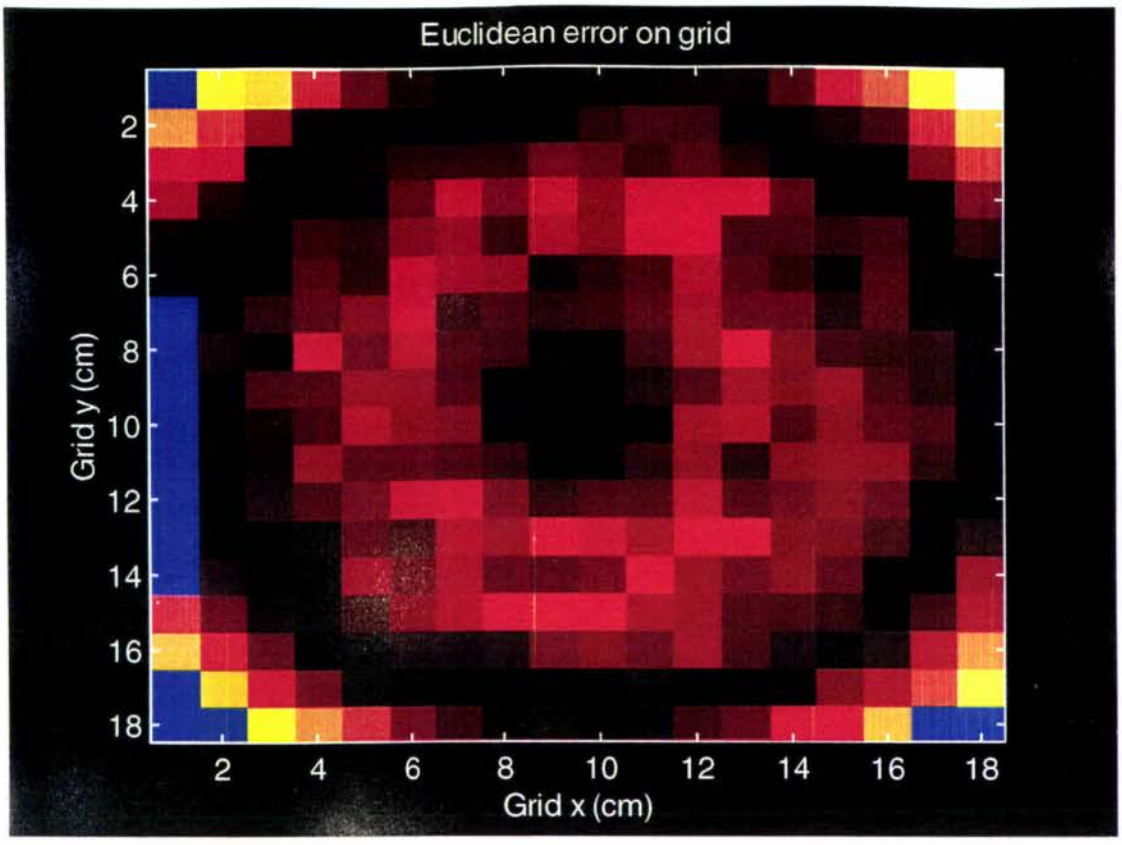


Figure 8-3 **(Second order polynomial fit)**

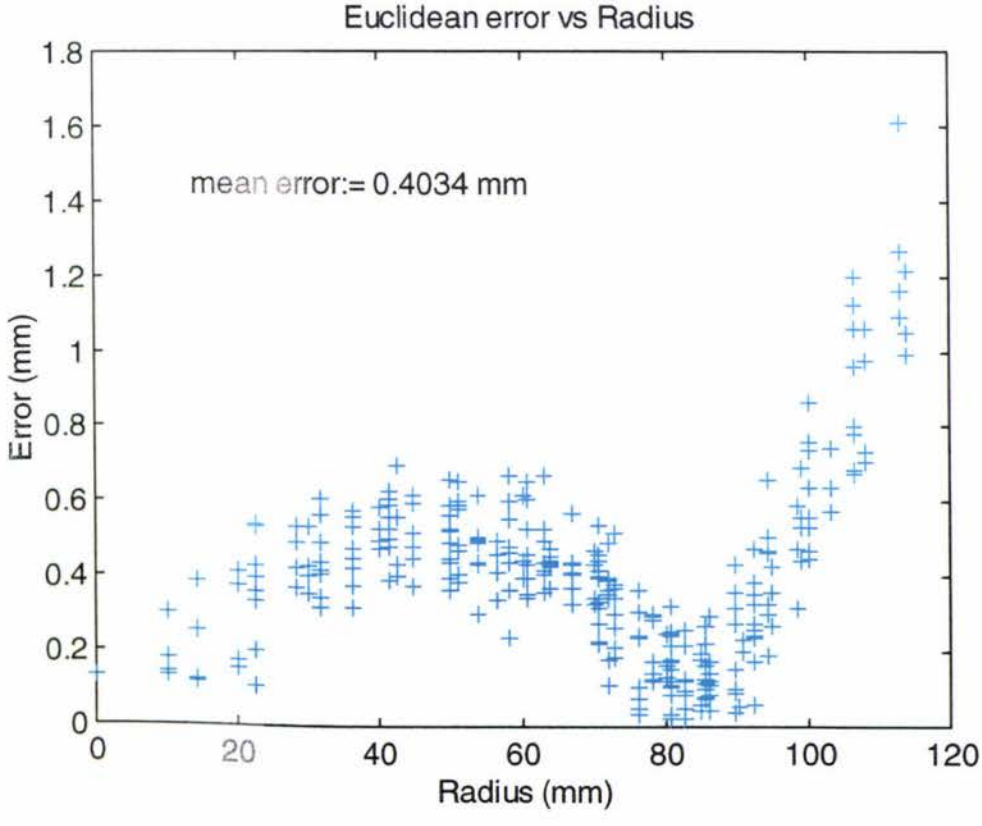


Figure 8-4 (Second order polynomial fit)

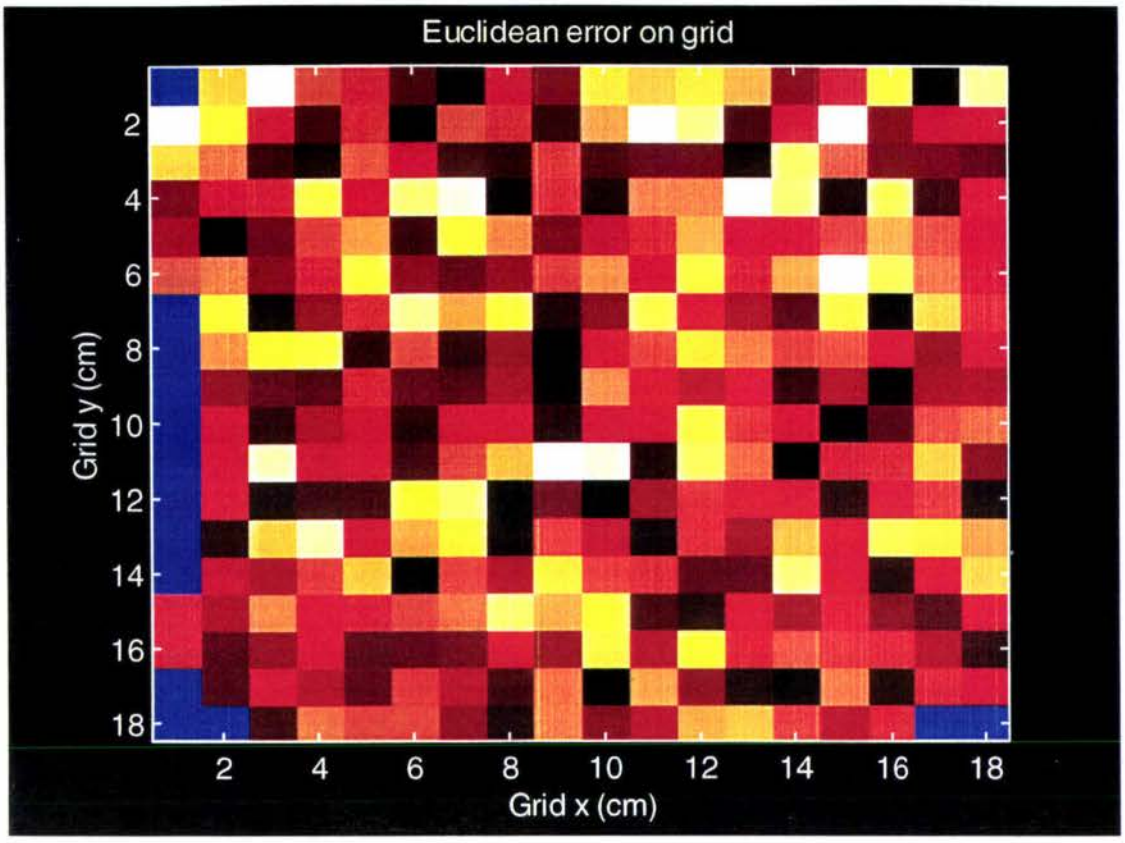


Figure 8-5 (Third order polynomial fit)

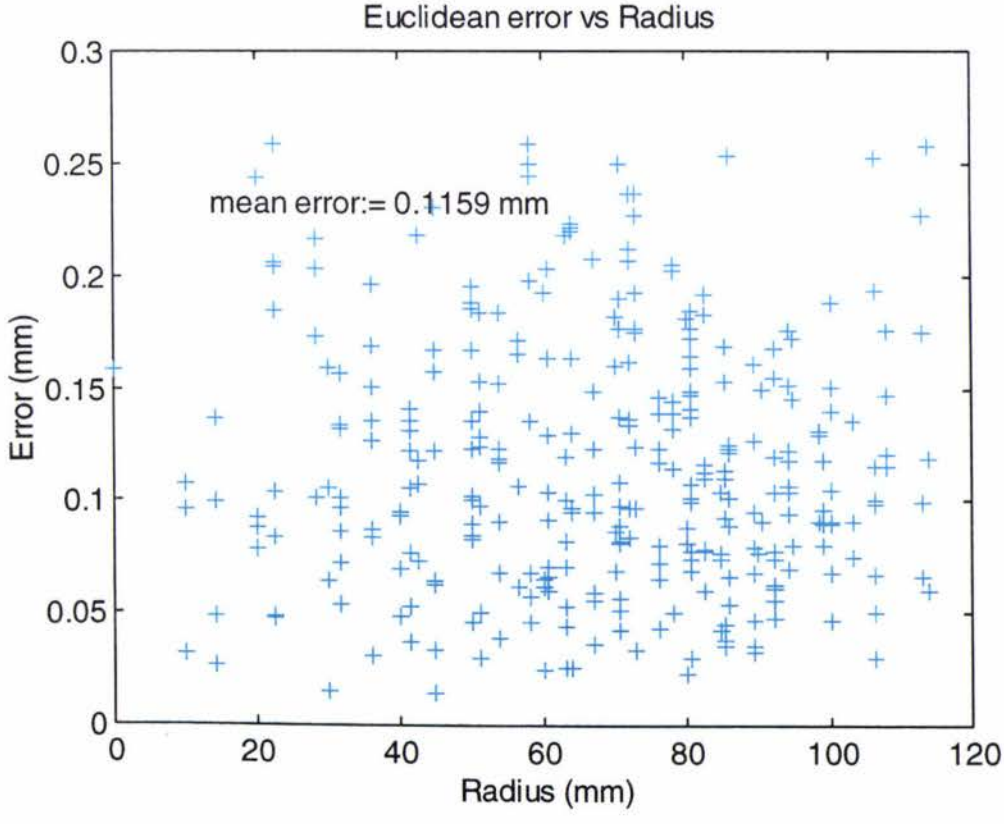


Figure 8-6 (Third order polynomial fit)

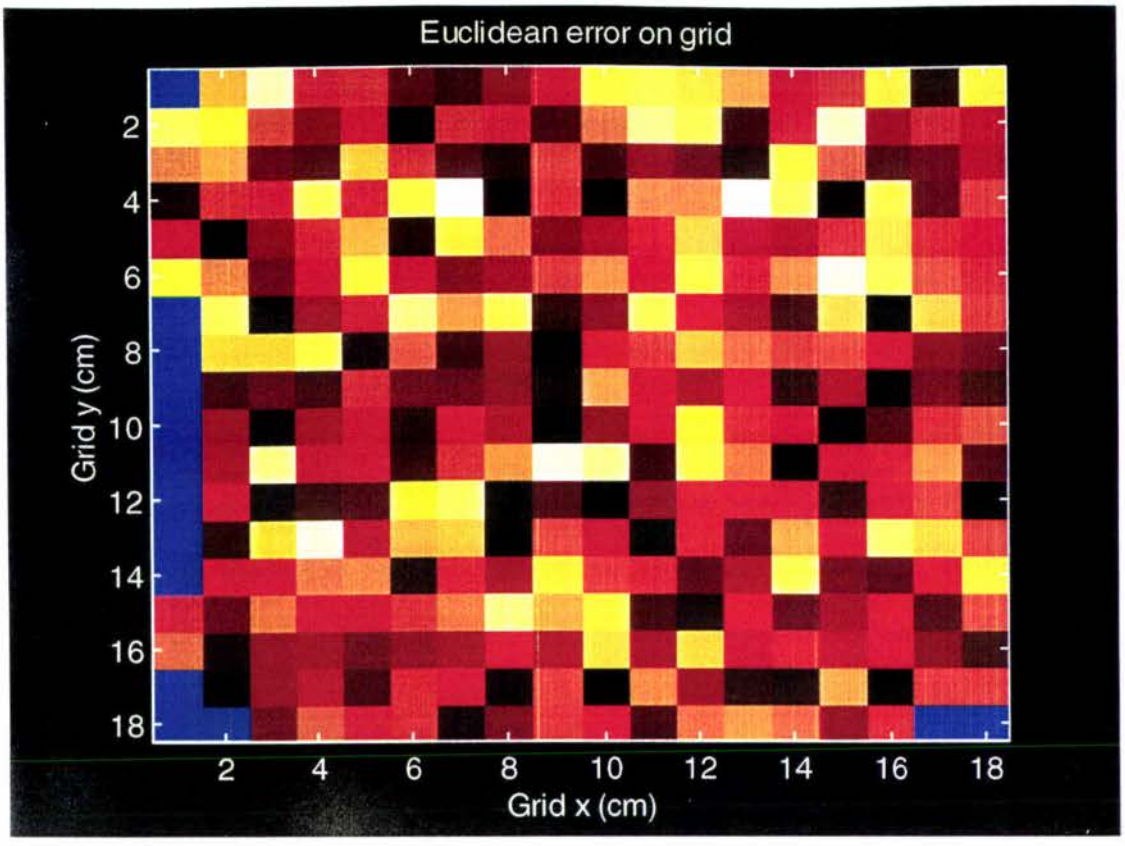


Figure 8-7 (Fourth order polynomial fit)

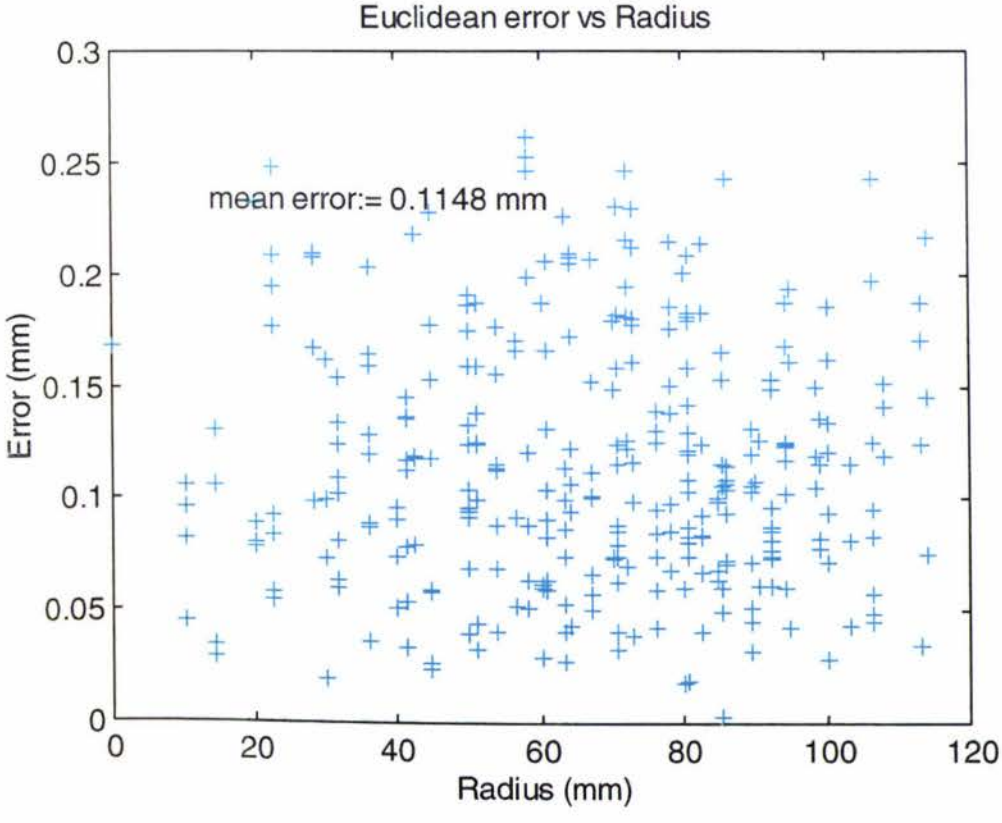


Figure 8-8 (Fourth order polynomial fit)

8.2 Fiducial mark diameter

The process was tested with images of calibration grids with a range of fiducial mark diameters, 1-5 mm, in steps of 1 mm. Images of the different grids were obtained under identical lighting conditions, at a height of 32 cm, and an aperture f22.

An estimate of the calibration error made by repeating the process on a second (verification) calibration grid. The results are displayed graphically in Figure 8-9. Calibration error is calculated as the mean absolute Euclidean error of the fiducial marks positions of the verification grid. Calibration error is displayed in the graph as points connected by a solid line. Mean absolute residual error from the polynomial fit, associated with the calibration grid, is displayed as a broken line joining data points. The calibration process was repeated for in each case for 1st, 2nd, 3rd & 4th degree polynomial fits with background image subtraction, a fourth order calculated background subtraction in the 4th order case, and with no background subtraction. In Figure 8-9, data series names are expressed as either absolute Euclidean error in polynomial fit, Pxy, or overall calibration error, Exy. In both cases the error description is immediately followed by the order of the polynomial fit and then the letters *bg*, if a background image subtraction was employed, *bgf*, if a fourth order fitted background was subtracted, or *nbg*, if a background image was not subtracted from the images of the calibration chart.

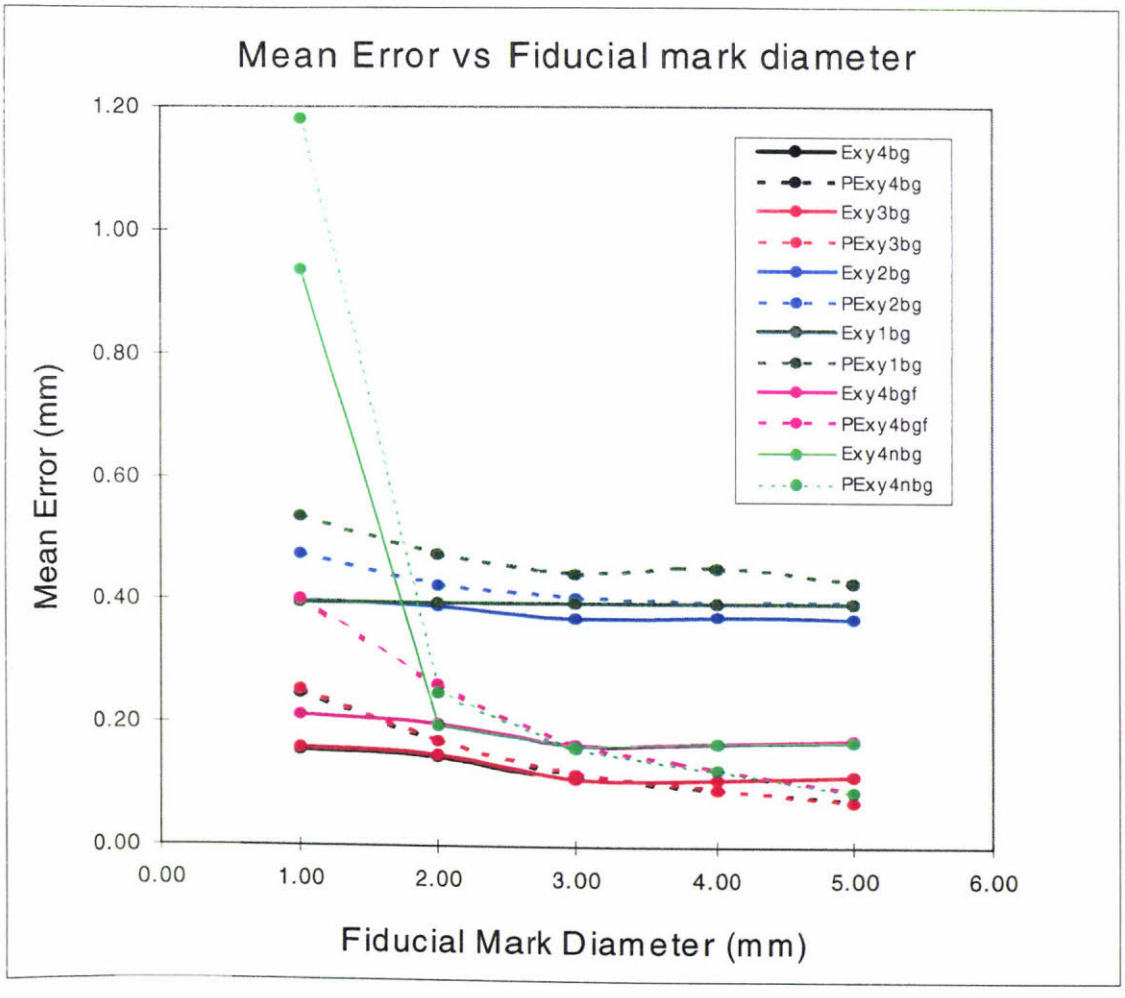


Figure 8-9

From Figure 8-9, it can be seen that a shallow minimum occurs at a fiducial mark diameter of 3 mm, the center of the range of diameters tested. Calibration error is seen to increase as the diameter is both increased and decreased from this point.

For the case in which no attempt was made to level the image in terms of illumination gradient, it can be seen that residual error from the polynomial fit and the overall calibration error is seen to increase dramatically for the fiducial mark diameter of 1 mm. This sudden increase is not seen for the case of background image subtraction or the fitted background subtraction.

8.3 Aperture control

Images of the calibration grid were captured for the range of aperture settings, f2, f2.8, f4, f5.6, f8, f11, f16, f22. The illumination level during image capture was controlled so as to maintain a constant mean intensity value in each image for the range of aperture settings. A separate verification image was acquired in order to calculate a mean calibration error associated with the process. The results for this are displayed graphically in Figure 8-10.

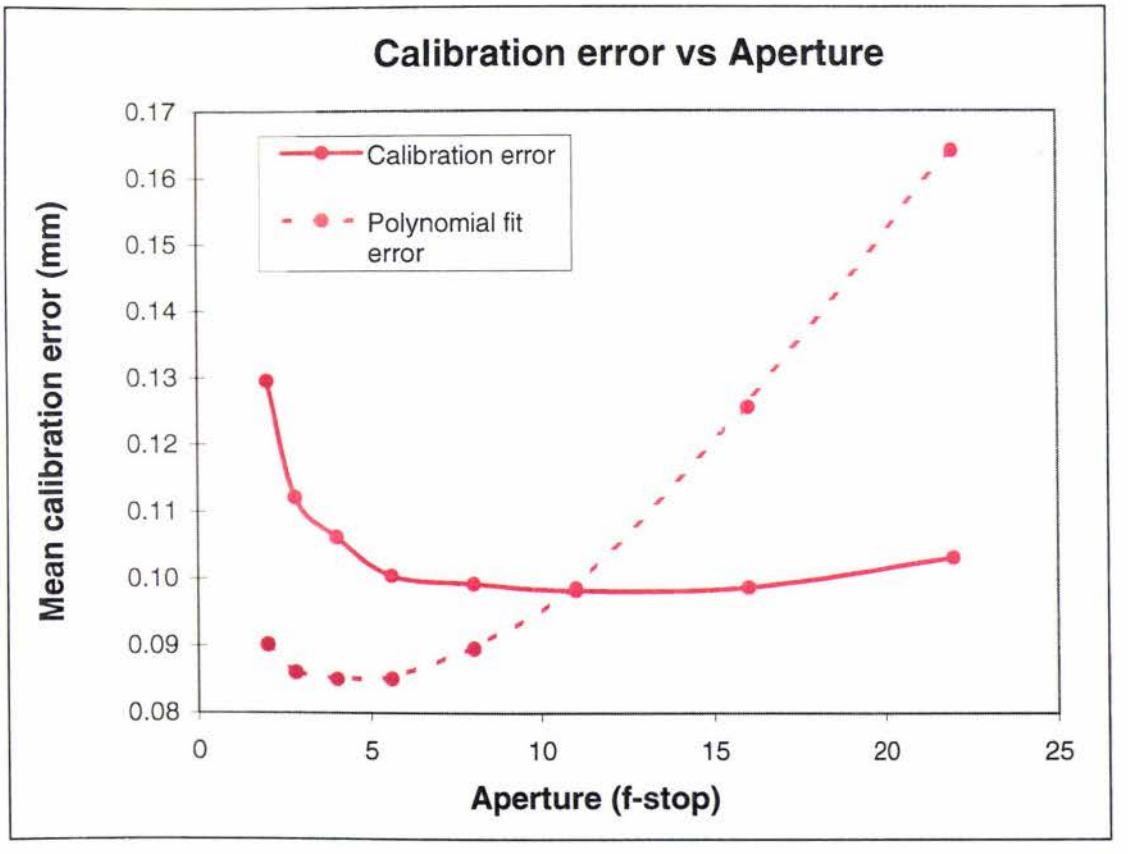


Figure 8-10

From the obtained data, it can be seen that there is a general reduction in polynomial residual error that corresponds to an increase in the physical aperture of the camera lens (reduction in f-stop). For values of f-stop below f4, the polynomial residual error is seen to increase.

In terms of overall calibration error, there is a rapid increase for larger sizes of physical aperture, shown by low f-stop values below about f8. As the physical aperture is reduced, above f11, there is evidence to suggest that calibration error gradually increases.

8.4 The fitted Polynomial

Figure 8-11 displays the results obtained for a 4th order fitted polynomial for the image of the 10 mm spaced grid of 3 mm diameter fiducial marks, with background image subtraction. The image was obtained with the Fujinon TV lens aperture setting at f22 at an object distance of 32 cm. The tabulated data is the calculated coefficient associated with each term of the polynomial equation. Two equations are given, Xg=f(x,y), which relates coordinates in the image to Xg, the x-axis coordinate in the calibration grid and Yg=f(x,y), which relates coordinates in the image to Yg, the y-axis coordinate in the calibration grid.

	Xg=f(x,y)	Yg=f(x,y)
x^4	0.00000000002442	-0.00000000002334
x^3y	0.00000000000066	-0.00000000000781
x^2y^2	-0.00000000001565	0.00000000000406
xy^3	-0.00000000003727	-0.00000000002099
y^4	0.00000000003448	-0.00000000001407
x^3	-0.00000010471837	0.00000000120125
x^2y	-0.00000000164988	0.00000009995105
xy^2	-0.00000009900279	0.00000000628606
y^3	0.00000000590162	0.00000010690793
x^2	-0.00000285546020	-0.00000233341107
xy	0.00000712534334	0.00000101647834
y^2	-0.00000240774552	-0.00000746935936
x	-0.34711921218081	-0.00155629361882
y	-0.00082956538101	0.36224411861554
k	4.98265272380295	-0.26838706501972

Figure 8-11

8.5 Execution time

The image calibration procedure was carried out on a Pentium-100 MHz (32MB) computer system running in MATLAB version 4.2c under Windows 95. Processing time required for the calibration of the 512 x 512 pixel image of the control grid, having approximately 320 fiducial marks, was about 6.5 minutes. This included the processing of a verification image for the determination of an overall estimate of calibration error.

Chapter 9 -Discussion

- **Order of the fitted equation**

The residual error from the fitted polynomial equation does not represent the total error present in the calibration process because it does not take into account the error associated with the centroid determination process. To calculate an overall calibration error, it is necessary to apply the polynomial transformation to the centroid positions from a second verification image and to compare the calculated positions to actual positions on the grid to give an absolute calibration error.

By varying the degree of the fitted polynomial and examining the residuals and by looking at the mean absolute calibration error as estimated by calculating a verification image (Figure 8-9), it can be seen that a 4th order equation most closely approximates the spatial transformation function that maps the positions of fiducial marks in an image to their true positions in two dimensional space on the calibration grid. This transformation takes into account factors of distortion in the camera lens system, as well as geometric warping due to camera alignment and orientation which exhibit gradual variations across the image and can be accurately modeled by the fitting of low order polynomial equations. The result is confirmed by the absence of obvious trends present in the residuals that would suggest systematic error present in the model (Figure 8-8).

At an aperture setting of f22, and with background image subtraction, the maximum absolute Euclidean error obtained for the 4th order fitted polynomial was 0.25 mm, (Table 8-1). On the 18 cm square grid, and when aligned with an image axis, this represents a maximum subpixel accuracy of 0.71 pixels. The mean subpixel accuracy (0.115 mm) when aligned to an image axis gives an average error of 0.33 pixels.

- **Fiducial mark diameter**

Figure 8-9 shows the results obtained for variation in fiducial mark diameter as well as adopting background image subtraction and fitted background subtraction for the elimination of illumination gradients which are known to affect the accuracy of the Centroid Calculation Method. From the graph, it can be seen that the calibration error is significantly reduced for all fiducial mark diameters by increasing the order of the fitted polynomial function. For the case where no background subtraction is performed, it can be seen that the accuracy of the calibration process is most adversely affected for the fiducial mark with the least diameter, 1 mm. This is because, for fixed centroid subwindow dimensions, there is a far lower ratio of object pixels to background pixels for small fiducial marks. Hence the object exerts a reduced influence on the calculated centroid position, and background variation intensity has greater affect on the average position. This is clearly demonstrated in Figure 8-9, for the data where no subtraction at all is performed.

For the case in which a fourth order equation is fitted for the background and subtracted, it can be seen that the high amount of error associated with the 1 mm diameter fiducial mark, is eliminated. This is because the 4th order fitted background removes the slowly changing intensity gradient that most greatly affects small fiducial mark position estimations. For fiducial marks greater than 1 mm in diameter, calibration error is not significantly different from the case where background subtraction is not performed at all.

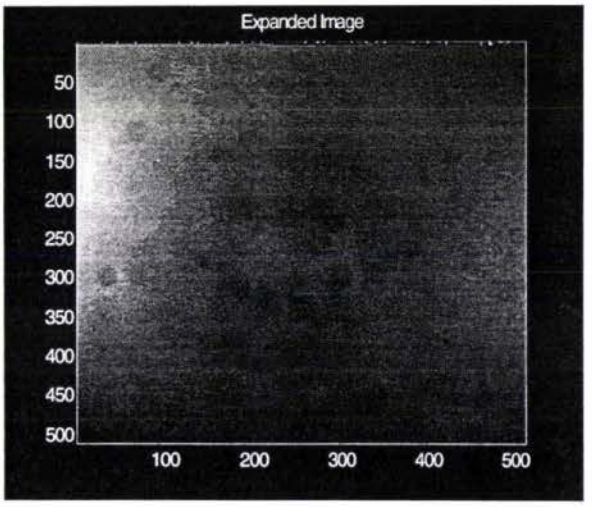


Figure 9-1

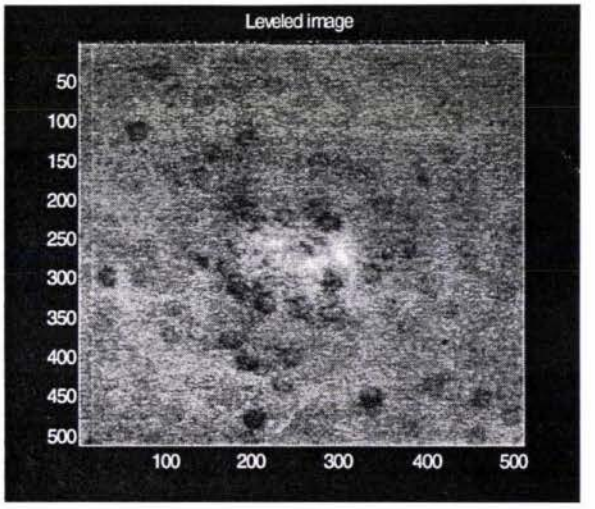


Figure 9-2

When a separate background image is acquired, and subtracted from the calibration image, there is a significant reduction in the calibration error for all values of fiducial mark diameter, when compared to the fitted background subtraction. The magnitude of this difference is approximately 0.06 mm. The most likely explanation for this difference is the fact that for the 4th order fitted background, only the slowly changing variation in background intensity is modeled. It is not possible for a low order (4th) polynomial equation to take into account rapid changing localized variations in intensity. These higher spatial frequency variations may be due to noisy photo sites on the CCD sensor, spot illumination discrepancies, or dirt on the glass cover used to maintain the flatness of the calibration chart during image acquisition. These variations were identified in the expanded image of the background and are displayed in Figure 9-1. Figure 9-2 illustrates the effect when a two dimensional 4th order polynomial is fitted to the background image (shown in Figure 9-1) and subtracted. It can be seen from this image that the low frequency intensity variation has been successfully removed from the image but the higher spatial frequency intensity variations, present in the background image, still remain after the fitted background has been subtracted. The light region situated at the center of the image is a result of light re-reflected from the camera lens back down onto the chart.

This is not the case when a background image is captured and used to level the image of the control chart by subtraction. This is because the rapid changing component is also captured in the background image and is effectively eliminated from the leveled image during subtraction. For this reason, the subtraction of an acquired background image, for the purpose of intensity gradient removal, is the preferred option.

In terms of fiducial mark diameter, for this particular image size, approximately 18 cm x 18 cm for a 512x512 pixel image, there appears to be an optimal fiducial mark diameter of about 3 mm. As the diameter is reduced beyond this size it was found that background illumination gradients have an increasing effect and increases the calibration error.

As well as this, it is also possible that conditions prescribed for the elimination of systematic error in the Centroid Calculation Method due to undersampling, will not be satisfied for smaller fiducial mark diameters. (e.g.) the estimated centroid location of the sampled fiducial mark will not equal the actual position (Section 6.6.1).

It is also inevitable that as fiducial mark diameter is increased, beyond a certain size, adjacent fiducial marks will begin to interfere with each other resulting in a loss of accuracy in the centroid determination and leading to an increase in calibration error. A fiducial mark having a diameter of 3 mm in this case has a pixel-diameter in the image of about 8.5 pixels.

- **Control of aperture**

Although lens distortion present in an image is not affected by changing the physical aperture setting (Section 2.2), this is not the case with fiducial mark position determination. The effect that changing aperture has on the fitted polynomial residuals and overall calibration error is shown in Figure 8-10.

At large physical apertures (f2), the residual error from the 4th order polynomial fitting process is relatively low (~90μm). On the other hand the overall calibration error is observed to be relatively high. This indicates that the error associated with centroid location of fiducial marks at large apertures is high. This phenomenon is almost certainly associated with the loss of image definition that can be attributed to primary lens aberration that predominate at large apertures (e.g.) spherical aberration, coma, astigmatism (Section 2.1).

Combined with this, systematic error associated with the calculation of the centroid (Section 6.6.1) (Alexander & Ng, 1991) may also contribute to error at large apertures due to spatial undersampling of fiducial marks.

As the physical size of the aperture is reduced (f8-f11) the accuracy of the calibration procedure shows a significant improvement. A calibration error 'minimum' occurs at an aperture setting , f11, which is probably due to a reduction in the effects of lens aberration.

Despite increasing residual error in the fitted polynomial at smaller physical apertures, the overall calibration error remains relatively low. This occurrence is a good indication that the fitted equation more precisely models the true inverse spatial transformation function of the camera lens system, that maps positions in the image to physical positions in the real world.

The increase in calibration error that is observed at f16 and f22 is most likely the increase in error due to the increasing diffraction of light at the aperture. Predicted by theory and described by Kingslake (1992), increasing diffraction at the aperture reduces image definition according to the following equation:

$$\text{Resolution (lines per mm)} = \frac{1600}{F-number} \qquad \text{Equation 9.1}$$

This relationship describes the maximum achievable system resolution in lines per millimeter due to diffraction corresponding to a certain physical aperture defined by its f-number. This figure is termed the "diffraction limit" of resolution and if a lens reaches this theoretical limit, it is called diffraction limited. Error due to diffraction is dominant at f22 and in all likelihood also significant at f16 as well.

It must be noted that these results correspond to the particular lens under test and that other lenses will give varying results depending on quality of the lens, focal length and f-number.

- **Error assessment**

In terms of the overall accuracy achieved by this calibration process it is necessary to look at the components of error that contribute to the overall calibration error that is being evaluated.

At the first step there is the error involved in the construction of the calibration grid itself. The grid is accurately created using the precision drafting software CorelDraw. However the charts themselves are printed at a nominal precision specified at 600 dpi on a HP LaserJet printer. For a grid size of about 18 x 18 cm and an image size of 512 x 512 pixels this amounts to a resolution of 0.12 pixels horizontal and vertical resolution for each calibration chart imaged. This equates to about 0.17 pixels when the maximum uncertainty is aligned with the image axis or 0.06 mm Euclidean error on a two dimensional calibration chart. In our calibration

procedure two charts were employed, firstly one for derivation of the spatial transform function and then the second for verification purposes, giving a maximum uncertainty of 0.34 pixels, or 0.12 mm.

The overall calibration error obtained for the calibration of the Fujinon TV lens was about 0.10 mm. This value compares favorably with the error estimate based on uncertainty of fiducial mark position location in the actual control chart. (0.12 mm).

The overall benefits of the calibration process can be illustrated by examining the results obtained for the 10 mm spaced grid of 3 mm diameter fiducial mark, captured at f22, and leveled using a background image subtraction process. Without a calibration process, the best that one could hope to achieve would be by assuming a linear transformation, and ignoring the non-linear effects of lens or geometric distortion. In this case the greatest residuals occur near the outside of the image where the effects of distortion are greatest (Figure 8-2). In this situation the maximum absolute error is 1.8 mm (or 5.1 pixels when aligned with the image axis). The mean error is 1.25 pixels. When the image is calibrated with 4th order polynomial equations the maximum error obtained is 0.27 mm (0.77 pixels) and the mean error 0.115 mm (0.33 pixels) (Figure 8-8).

An additional improvement of about 15% can be obtained by acquiring the image at a lens aperture of f11, hence avoiding error attributable to diffraction.

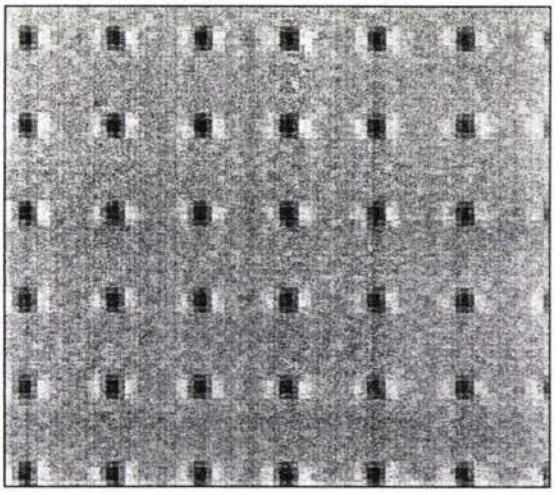


Figure 9-3

- **Calibration of other cameras**

In addition to the black & white camera trialed, two single chip color cameras were tested but extensive image errors were obtained due to color smearing between adjacent photosites on the CCD chip. When the color image was converted to an intensity only component, or grayscale image, the result was groups of high intensity pixels adjacent to fiducial marks, as illustrated in Figure 9-3. The intensity of these pixels were significantly brighter than the image background introducing large errors into the method used for the calculation of the centroid of the fiducial marks. Whilst this problem could be reduced by using reduced shutter speeds on cameras with electronic shutters it was concluded that due to the absence of color image information in this particular application that this procedure was more ideally suited for use with black & white cameras. The use of higher quality 3-chip color cameras would also alleviate this problem, but these were not tested.

Chapter 10 -Conclusion

An automated camera calibration procedure was successfully developed based upon image analysis methods and implemented using Mathworks MATLAB software. The process was used to determine a spatial transformation function that mapped fiducial mark positions, determined from an image of a calibration chart obtained from a video camera, into the real world positions represented on the chart.

A set of polynomial equations was calculated in an attempt to explain variation in the model attributable to lens distortion and geometric distortion due to camera alignment and perspective.

The effect of altering the degree of the fitted polynomial on error was investigated. No significant reduction in error was achieved by increasing the order of the fitted polynomial equation above order 4.

The effect of altering fiducial mark size on error was investigated. For the Fujinon TV lens being tested at a focal distance of 32 cm and a field size of approximately 18 cm square, an optimal fiducial mark diameter was determined to be 3 mm. Increased calibration error was obtained for fiducial mark diameters both greater and less than this figure.

The effects on calibration error of varying aperture were investigated. Greater calibration error observed at low aperture settings was attributed primarily to lens aberration, and possibly also systematic error in the Centroid Calculation Method due to spatial undersampling in the image. Increased diffraction, resulting in loss of definition at the reduced aperture, probably explained an the increase in error that was observed at aperture setting, f22. An optimum aperture was determined to be f11, for this particular lens.

The use of this camera calibration procedure has resulted in a large increase in accuracy for position determination or measurement from an image. When the non-linear effects of geometric or lens distortion are ignored, the maximum observed error was seen to be as high as 5.10 pixels compared to the maximum error in a 4th order calibrated image of 0.77 pixels. Mean error was observed to decrease from 1.25 to 0.33 pixels in the calibrated case.

The mean error obtained as a result of the calibration closely approached the actual uncertainty present in the physical calibration chart.

The computation time required for the calibration of an image of a control chart having 320 control points, including the calculation of a verification image, was found to be 6.5 minutes.

The advantage of the automated camera calibration procedure is that it is accurate and fast, unlike manual methods that are tedious, time consuming and prone to error.

11. Appendix

Guidelines for using the Calimage program.

The Calimage program was written in MATLAB m-file script language and where possible compiled to produce executable MEX-file code. The program was run under MATALB version 4.2c using the Image Processing Toolbox on a Pentium 100MHz computer with 32Mb RAM.

The Calimage program requires a 512 x 512 pixel image of the control grid, a image of the background (if background subtraction is to be used as a leveling process), and an additional verification image of the calibration grid if this procedure is required.

1) The program initially asks the user to define the method of image leveling to be performed. The options are (1) Bkgrnd image, in which the user is later asked to specify a background image for subtraction, (2) Bkgnd4 fit, in which case a 4th order polynomial will be fitted to the background and subtracted, or (3) No Bgnd, in which case there will be no leveling of the image.
2) If the first option was selected in the previous step the user will now be asked to specify a background image file and path.
3) In the next step the program will prompt the user to enter the order of the polynomial to be fitted for the spatial transformation function.
4) The program will next prompt the user to enter the name and path of a verification image to be used to calculate the overall calibration error.

Results

The polynomial function coefficients are returned for the equations: Xg=f(x,y) and Yg=f(x,y). The order that the coefficients are given are in the order of decreasing powers of x and y. For the 4th order fitted polynomial, the order of the coefficients will be:

$$x^4, x^3y, x^2y^2, xy^3, y^4, x^3, x^2y, xy^2, y^3, x^2, xy, y^2, x, y, k$$

The Mean Euclidean Error (mm), for fitted polynomial, is given as well as the Mean Euclidean Error of calibration (mm).

Plots are also given for distribution of polynomial residuals on the grid, and versus radius from the center of the grid, and also calibration error on the verification grid.

A sample output from the program is given in Figure A1. This output corresponds to images obtained for the 10 mm space grid of 3 mm diameter fiducial marks obtained for the Fujinon TV lens at an aperture of f22, and utilizes background image subtraction and a verification image to calculate overall calibration error.

Sample Output for 'CALIMAGE' MATLAB program

» ***calimage***
Select background subtraction
Bgrnd image(1) Bgnd4 fit(2) No Bgnd(3):***1***
Backgnd image, (Path\file.bmp): ***c:\matlab\masters\lens\sgrid000***
Enter order for polyfit:***4***
Verification Image (Path\file.bmp) ***c:\matlab\masters\lens\sgrid03r***
Enter Grid Seperation (mm):***10***
Enter test image (path\file.bmp) ***c:\matlab\masters\lens\sgrid003***

Polynomial coefficients for Xg=f(x,y)
ans =
0.00000000002442
0.00000000000066
-0.00000000001565
-0.00000000003727
0.00000000003448
-0.00000010471837
-0.00000000164988
-0.00000009900279
0.00000000590162
-0.00000285546020
0.00000712534334
-0.00000240774552
-0.34711921218081
-0.00082956538101
4.98265272380295

Polynomial coefficients for Yg=f(x,y)
ans =
-0.00000000002334
-0.00000000000781
0.00000000000406
-0.00000000002099
-0.00000000001407
0.00000000120125
0.00000009995105
0.00000000628606
0.00000010690793
-0.00000233341107
0.00000101647834
-0.00000746935936
-0.00155629361882
0.36224411861554
-0.26838706501972

Mean Euclidean error for fitted polynomial
mez = 0.11481780357207
Mean Euclidean error of calibration
mez = 0.10728933384477

Figure A1

12. Bibliography

Aghalan, H., Schaper, C., and Kailath, T. (1993). "Machine vision techniques for subpixel estimation of critical dimensions." *Optical Engineering*, 32(4), 828-839.

Alexander, B., and Ng, K. (1991). "Elimination of systematic error in subpixel accuracy centroid estimation." *Optical Engineering*, 30(9), 1320-1331.

Brown, D. (1966). "Decentering distortion of lenses." *Photogrammetric Engineering*, 32(3), 444-462.

Brown, D. (1971). "Close-Range Camera Calibration." *Photogrammetric Engineering*, 37(8), 855-866.

Curry, S., and Baumrind, S. (1986). "Calibration of an array camera." *Photogrammetric Engineering and Remote Sensing.*, 52(5), 627-636.

Ethrog, U. (1984). "Non-metric camera calibration and photo orientation using parallel and perpendicular lines of the photographed objects." *Photogrammetria*, 39, 13-22.

Faig, W. (1975). "Calibration of close range photogrammetric systems-mathematical formulation." *Photogram. Eng. Remote sensing*, 41(12), 1479-1486.

Fillard, J. (1992). "Subpixel accuracy location estimation from digital signals." *Optical Engineering.*, 31(11), 2465-2471.

Fillard, J., M'timet, H., Lussert, J., and Castagne, M. (1993). "Computer simulation of super-resolution point source image detection." *Optical Engineering*, 32(11), 2936-2944.

Fraser, C. and Shortis, M. (1992). "Variation of Distortion within the Photographic Field." *Photgrammetric Engineering and Remote Sensing*, 58(6), 851-855.

Fryer, J. (1986). "Distortion in a zoom lens." *Aust. J. Geod. Photogram. Surv.*(44), 49-59.

Fryer, J. (1987). "Lens Distortion: the forgotten component of digital photogrammetry." *Applications of close range photogrammetry*, Univ. Melbourne, Australia, 65-77.

Fryer, J., and Brown, D. (1986). "Lens distortion for close range photogrammetry." *Photogrammetric journal of remote sensing*, 52(1), 51-58.

Fryer, J., and Fraser, C. (1986). "On the calibration of underwater cameras." *Photogrammetric Record*, 12(67), 73-85.

Fryer, J., and Mason, S. (1989). "Rapid Lens Calibration of a Video Camera." *Photogrammetric Engineering and Remote Sensing*, 55(4), 437-442.

Gibson, D., and Gaydecki, P. (1995). "Definition and application of a Fourier domain texture measure: Applications to histological image segmentation." *Computers in biology and medicine.*, 25(6), 551.

Grosky, W., and Tamburino, L. (1990). "A unified approach to the linear camera calibration problem." *IEEE transactions on pattern analysis and machine intelligence*, 12(7), 663-671.

Havelock, D. (1989). "Geometric precision in noise-free digital images." *IEEE transactions PAMI*, 11(10), 1065-1075.

Jain, A. (1989). *Fundamentals of digital image processing*, Prentice Hall, Englewood Cliffs, NJ.

Juvells, S., Vallmitjana, A., Carnicer, A., and Campos, J. (1991). "The role of amplitude and phase of the Fourier transform in the digital image processing." *American Journal of Physics.*, 59(8), 744-748.

Kingslake, R. (1989). *A history of the photographic lens*, Academic Press, Inc., London.

Kingslake, R. (1992). *Optics in photography*, SPIE -the International Society for Optical Engineering, Washington.

Lichti, D., and Chapman, M. (1996). "CCD camera calibration using the finite element method." *SPIE the International Society for Optical Engineering*, 2598, 34-43.

Lyvers, E., Mitchell, O., Akey, M., and Reeves, A. (1989). "Subpixel measurements using a moment-based edge operator." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(12).

Magill, A. (1955). "Variation in distortion with magnification." *Journal of research of the National Bureau of Standards*, 54(3), 135-142.

McLean, G. (1993). "Image warping for calibration and removal of lens distortion." *IEEE Pacific Rim Conference on communications, computers and signal processing.*, 170-173.

Melen, T., and Balchen, J. (1994). "Modeling and calibration of video cameras." *Conference of the SPIE -the International Society for Optical Engineering.*

Mikhail, E., Akey, M., and Mitchell, O. (1984). "Detection and sub-pixel location of photogrammetric targets in digital images." *Photogrammetria*(39), 63-83.

Murai, S., Matsuoka, R., and Okuda, T. (1984). "A study of analytical calibration of non-metric cameras." *International archives of photogrammetry*, 25(5), 570-579.

Nomura, Y., Sagara, M., Naruse, H., and Ide, A. (1992). "Simple Calibration Algorithm for High-Distortion-Lens Camera." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11), 1095-1099.

Peli, T., and Malah, D. (1982). "A study of edge detection algorithms." *Computer Graph. Image Processing*, 20, 1-21.

Pilarski, S., Ivanov, A., and Kameda, T. (1994). "On minimizing aliasing in scan based compaction." *Journal of electronic testing*, 4(1), 83.

Ray, S. (1992). *The Photographic Lens*, Butterworth-Heinemann Ltd.

Reichel, L. (1991). "Fast QR decomposition of Vandermonde-like matrices and polynomial least squares approximation." *Siam J. Matrix Anal. Appl.*, 12(3), 552-564.

Robotics International of the SME (1984). *The Third Annual Applied Machine Vision Conference Proceedings.*, Elsevier Science Publishers, Chicago, Illinois.

Russ, J. (1995).*The Image Processing Handbook,* 2nd Ed., Boca Raton, Fla.: CRC Press.

Schatz, D. (1993). "Subpixel accuracy in machine vision software." *Sensors*, October, 75-76.

Seitz, P. (1988). "Optical superresolution using solid-state cameras and digital signal processing." *Optical Engineering*, 27(7), 535-540.

Shih, S., Hung, Y., and Lin, W. (1995). "When should we consider lens distortion in a camera calibration." *Pattern Recognition*, 28(3), 447.

Shih, S. H., Lin, W. (1993). "Accurate linear technique for camera calibration considering lens distortion by solving an eigenvalue problem." *Optical Egineering*, 32(1), 138-149.

Stanton, R., Alexander, J., Dennison, E., Glavich, T., and Hovland, L. (1987). "Optical tracking using charged-coupled devices." *Optical Engineering*, 26(9), 930-938.

Tian, Q., and Huhns, N. (1986). "Algorithms for subpixel registration." *Computer Vision, Graphics and Image Processing.*, 35, 220-233.

Tsai, R. (1987). "A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV camera and lenses." *IEEE journal Robotic Automation*, RA-3(4), 323-344.

Valkenburg, R. (1994). "An investigation of subpixel feature localization for precision measurement." *Image and Vision Conference*, Massey University.

Walker, B. (1994). *Optical Engineering Fundamentals.*, McGraw-Hill, Inc.

Wang, J., and Huang, H. (1994). "Film digitization aliasing artifacts caused by grid line patterns." *IEEE transactions on medical imaging.*, 13(2), 375.

Weng, J., Cohen, P., and Herniou, M. (1992). "Camera Calibration with distortion models and accuracy evaluation." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10), 965-980.

Wolberg, G. (1990). *Digital Image Warping*, IEEE Computer Society Press, Washington [DC].

Wong, K., and Ho, W. (1986). "Close range mapping with a solid state camera." *Photogrammetric Engineering and Remote Sensing.*, 52(1), 67-74.

Wu, X. (1991). "An efficient anti-aliasing technique." *Computer graphics.*, 25(4), 143.