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.

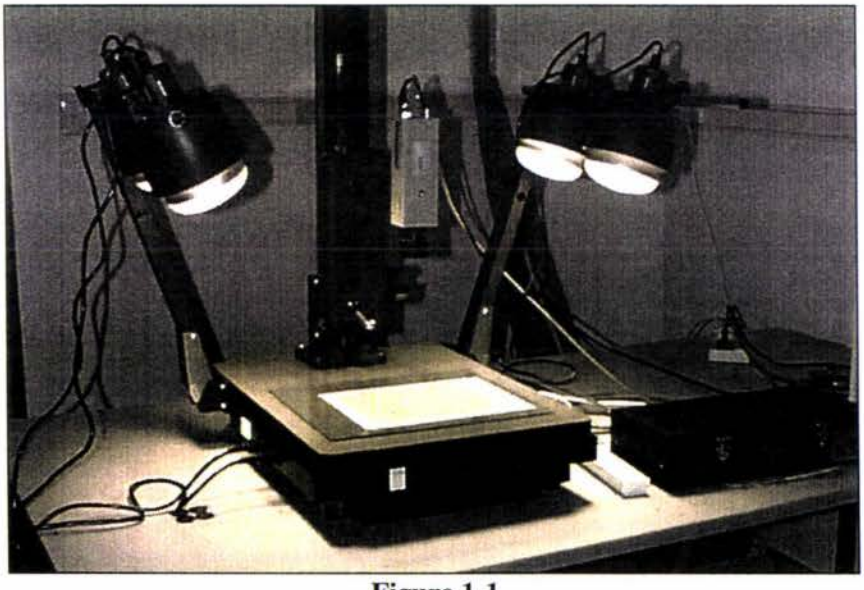


Figure 1-1

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.

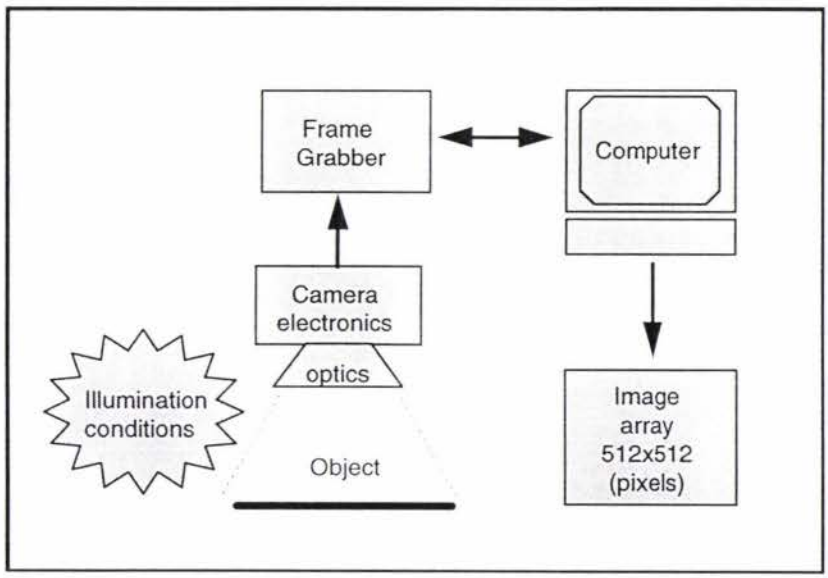


Figure 1-2

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.

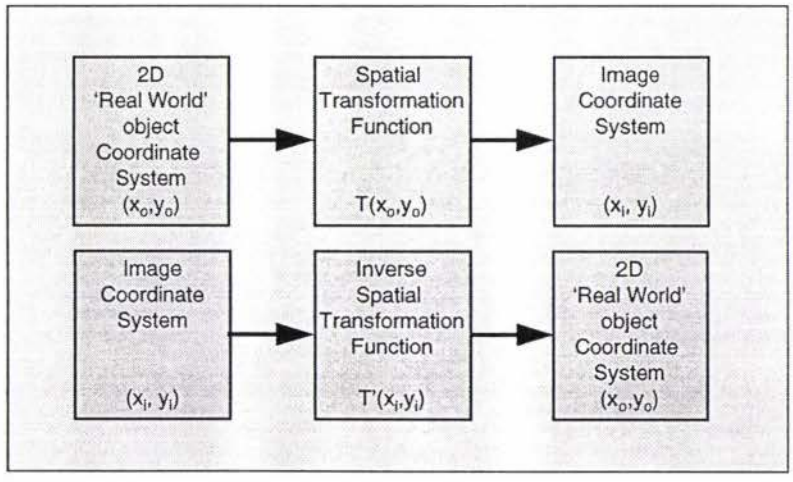


Figure 1-3

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.

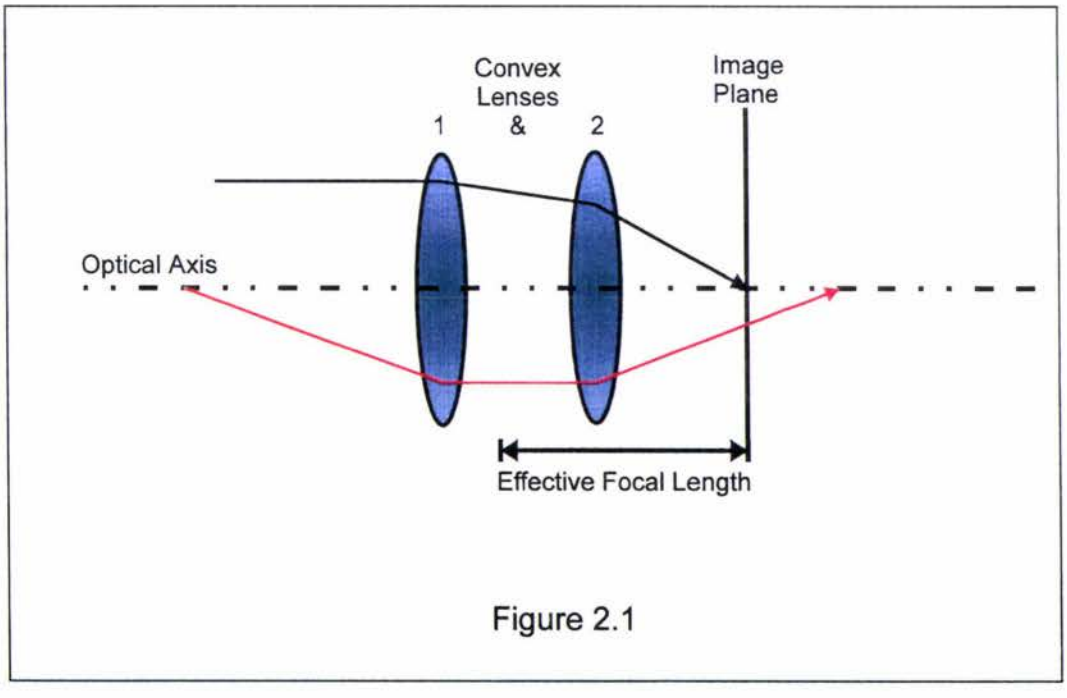


Figure 2.1

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.

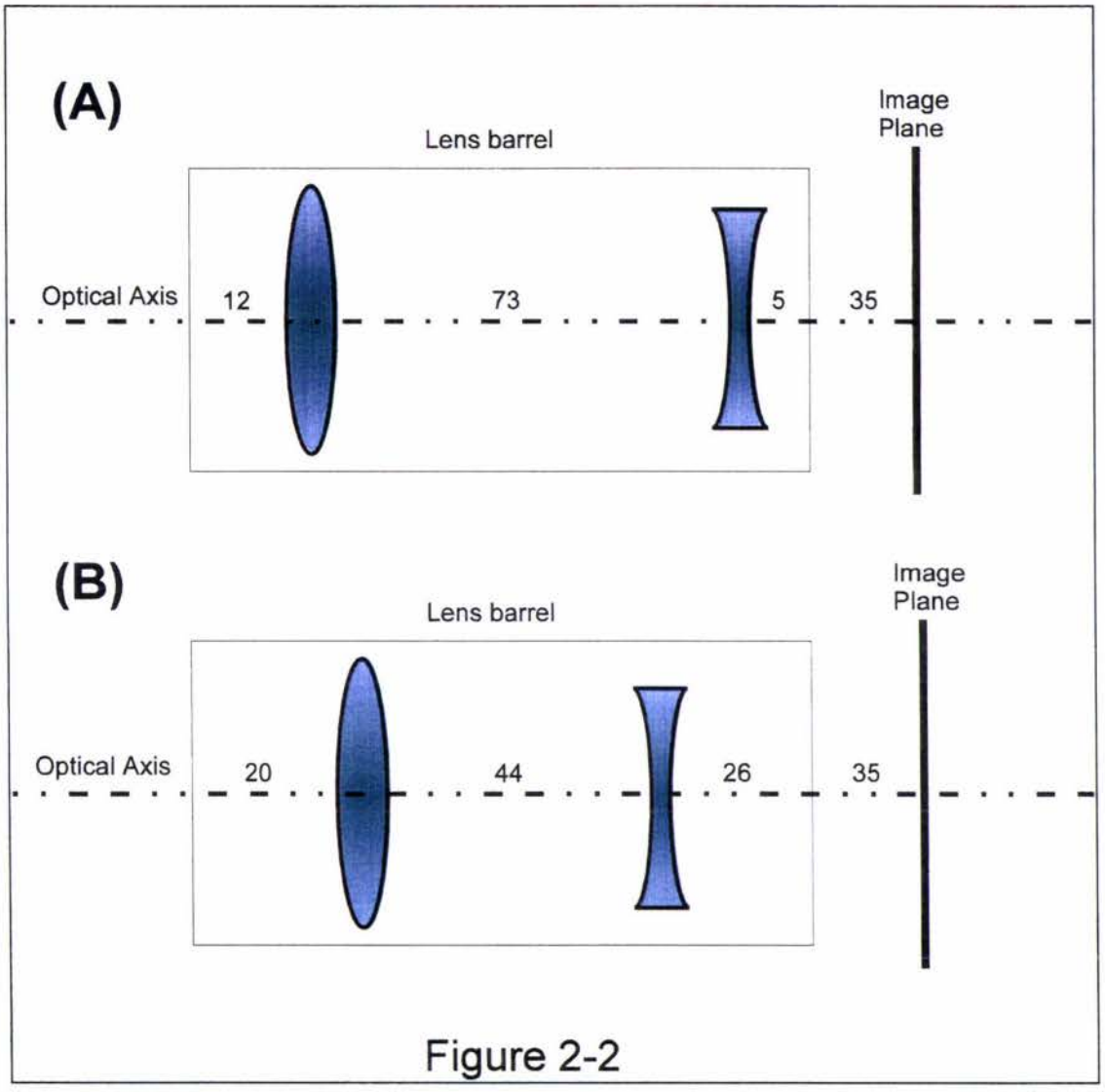


Figure 2-2

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.

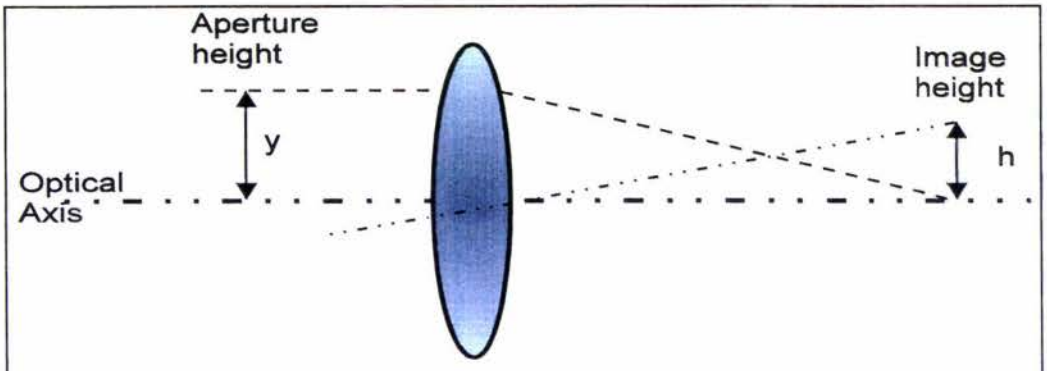
- Distortion of the image is present when the focal length of the lens varies as a function of field size or image height.

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'.

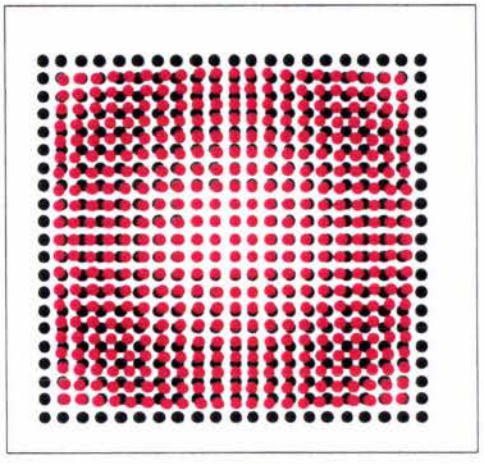


Fig. 2-3. Barrel distortion

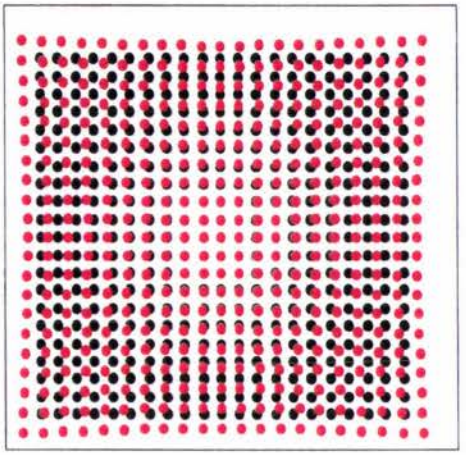


Fig. 2-4 Pin cushion distortion.

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.

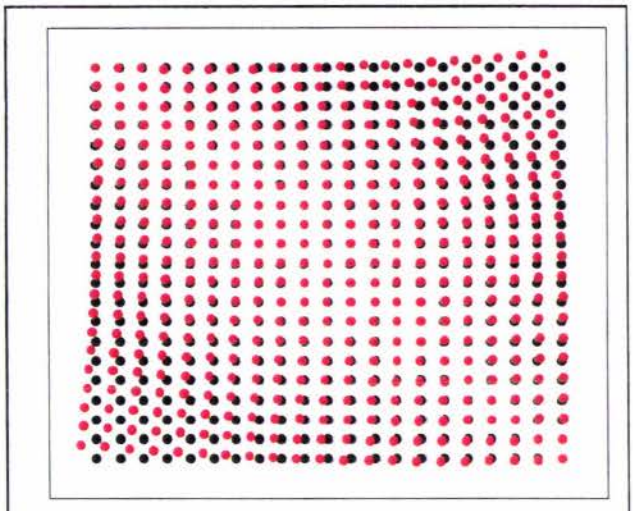


Fig. 2-5. Tangential distortion

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.

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

$$d_y(x,y) = (s_2 + p_2)x^2 + 2p_1xy + (s_2 + 3p_1)y^2 + k_1y(x^2 + y^2) \quad \text{Equation 2.14}$$

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

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

$$d_y(x,y) = g_2x^2 + g_3xy + (g_5 + g_4)y^2 + k_1y(x^2 + y^2) \quad \text{Equation 2.16}$$

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).

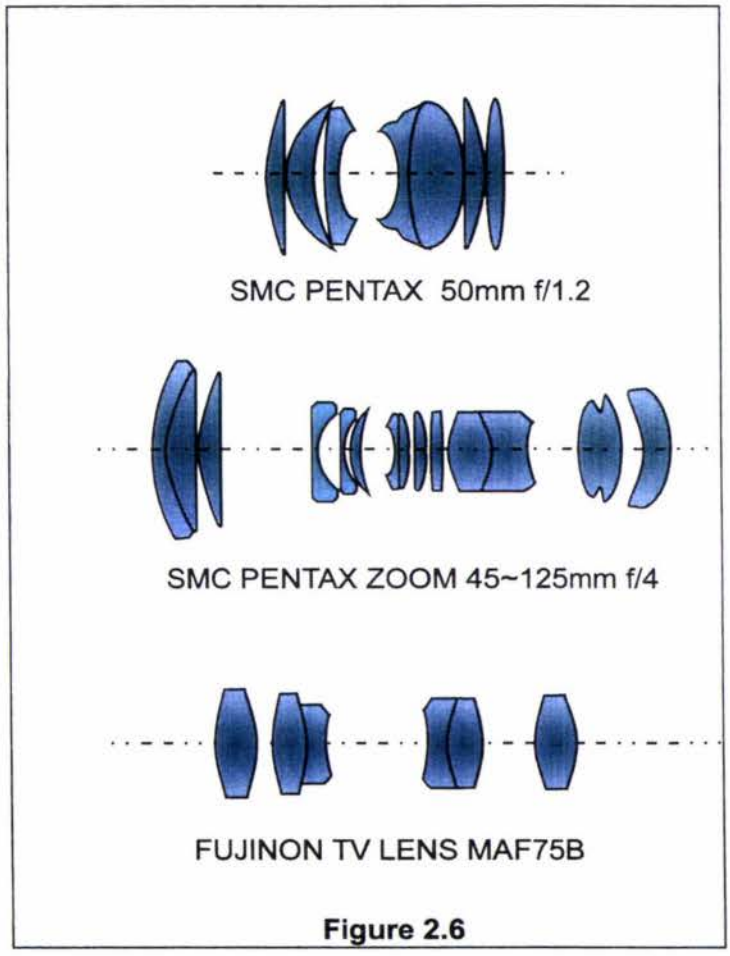


Figure 2.6

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.