Internal Camera Calibration using Rotation and Geometric Shapes

by

Gideon P. Stein

Submitted to the Department of
Electrical Engineering and Computer Science
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE
in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1993

ABSTRACT

This paper describes a simple method for internal camera calibration for computer vision systems. It is intended for use with medium to wide angle camera lenses. With modification it can be used for longer focal lengths. This method is based on tracking image features through a sequence of images while the camera undergoes pure rotation. This method does not require a special calibration object. The location of the features relative to the camera or to each other need not be known. It is only required that the features can be located accurately in the image. This method can therefore be used both for laboratory calibration and for self calibration in autonomous robots working in unstructured environments. The method works when features can be located to single pixel accuracy, but subpixel accuracy should be used if available.

In the basic method the camera is mounted on a rotary stage so that the angle of rotation can be measured accurately and the axis of rotation is constant. A set of image pairs is used with various angular displacements. If the internal camera parameters and axis of rotation were known one could predict where the feature points from one image will appear in the second image of the pair. If there is an error in the internal camera parameters the features in the second image will not coincide with the feature locations computed using the first image. One can then perform a nonlinear search for camera parameters that minimize the sum of squared distances between the feature points in the second image in each pair and those computed from the first image in each pair, summed over all the pairs.

The need to accurately measure the angular displacements can be eliminated by rotating the camera through a complete circle while taking an overlapping sequence of images and using the constraint that the sum of the angles must equal 360 degrees.

The closer the feature objects are located to the camera the more important it is that the camera does not undergo any translation during the rotation. A method is described which enables one to ensure that the axis of rotation passes sufficiently close to the center of projection (or front nodal point in a thick lens) to obtain accurate results.

This paper shows that by constraining the possible motions of the camera in a simple manner it is possible to devise a robust calibration technique that works in practice with real images. Experimental results show that focal length and aspect ratio can be found to within a fraction of a percent, and lens distortion error can be reduced to a fraction of a pixel. The location of the principal point and the location of the center of radial distortion can each be found to within a few pixels.

In addition to the first method, a second method of calibration is presented. This method uses simple geometric objects such as spheres and straight lines to find, first the aspect ratio, then the lens distortion parameters and finally the principal point and focal length. Calibration is performed using both methods and the results compared.

Thesis Supervisor: Professor John Wyatt
Title: Professor, Dept. of Electrical Engineering
Thesis Supervisor: Professor Berthold Horn
Title: Professor, Dept. of Electrical Engineering

Acknowledgements

I wish to thank my thesis supervisors Profs. Berthold Horn and John Wyatt for their support and advice throughout working on this thesis. I wish to thank my fellow students whose help, suggestions and feedback were instrumental to this work. I especialy thank William (Sandy) Wells and Lisa Dron for sparing the time from their thesis work to discuss this project and David Beymer for the feature detection program used in this work.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-91-J-4038.

Support for Gideon P. Stein is provided by a fellowship from the National Science Foundation.

Contents

List of Figures

List of Tables

1 Introduction:

1.1 What is camera calibration ?

Internal camera calibration involves finding the relationship between image coordinates and ray directions in space, which are measured in the camera coordinate system. For all but very long focal lengths this relationship can be best described by the perspective projection model, the model of the perfect pinhole. When using this model, the parameters which need to be determined are the focal length and the principal point. To get accurate results a model for lens distortion must be added and the parameters of the lens distortion must be found. Since the spatial sampling in the X and Y directions are different one must find the ratio between them. This ratio is often referred to as the pixel aspect ratio or scale factor.

External camera calibration involves finding the position (translation and orientation) of the camera in some world coordinate system. External camera calibration is important in stereo vision where one needs to find the relative position of the coordinate systems of the two cameras and it is also important in hand-eye coordination in robots.

Camera calibration is important if we wish to derive metric information from the images, although qualitative information can obtained from uncalibrated cameras and a stereo camera pair. Work has been done to see how much knowledge can be obtained using uncalibrated cameras [5] [13], but the mainstream efforts in robot vision assume some means of calibration.

In some cases we do not need to find the camera parameters explicitly. In other words, the transformation is defined in terms of intermediate parameters which are combinations of the camera parameters. The intermediate parameters can often be found quite easily. It is possible to calibrate stereo systems in this way. In the case where there is no lens distortion, we could find the transformation between the 3D world coordinates and image locations in the image planes of the two cameras using homogeneous matrices [18] [7]. We might never bother to actually break down the transformation matrices into internal and external parameters. We could make accurate measurements of the location of one object relative to another object in the scene. If on the other hand we wish to find the location of the camera relative to the objects in the scene we would be required to find the external calibration parameters explicitly. This is a more difficult problem [18].

Work on motion vision and pose estimation often use the perspective projection model and assume that the internal camera parameters are known. They also assume that the parameters that can correct for lens distortion are known. In the case of motion vision the structure of the world and the position of the camera are unknown and will be determined using a calibrated camera. For such work, accurately calibrating the internal camera parameters is critical. A camera centered coordinate system is used, therefore the external calibration is not important.

1.2 Brief review of other methods

Cameras used for machine vision are different in many important respects from the metric cameras used in the field of photogrammetry and these differences affect our choice of calibration method. The cameras used in machine vision have a lower resolution than the metric cameras. The typical 'high resolution' video CCD camera only has a resolution of about 780×500 pixels and even nonstandard cameras only have resolutions of up to 4000×4000 pixels. The lower resolution and the lower dynamic range make some calibration techniques, such as stellar calibration, impossible to use and all techniques must assume that features are located with larger amounts of uncertainty than with metric cameras. The lenses used in computer vision systems are typically cheap and mass produced. On the one hand this means that they have significant lens distortion and on the other hand they are not thoroughly tested by the manufacturers and do not come with calibration curves.

Since the cameras used in machine vision are not designed to be highly stable they might have to be calibrated frequently. The consumers of machine vision camera systems are not photogrammetrists and do not have the skill, patience or interest in laborious calibration methods. Therefore, camera calibration must be kept simple and as quick as possible. We provide here a brief review of the various approaches to camera calibration. More extensive reviews of calibration methods appear in [11][17] and also in the appendix.

1.2.1 Methods that use known world points

Most of the standard techniques for camera calibration for machine vision use a set of calibration points with known world coordinates. These are sometimes called control points. In laboratories, control points can be obtained using a calibration object [11] [22]. Outdoors, control points could be markings on the ground [17] or buildings whose positions can be verified from maps [18]. If the aspect ratio is unknown then a three dimensional calibration object is required [11] (i.e. the points cannot be coplanar). As a 3D object one can use a planar object with feature points clearly marked, which can be moved accurately in the Z direction, perpendicular to the plane of the points.

The typical statement of the calibration problem is the following: given a set of control points with known world coordinates (X_i, Y_i, Z_i) and their location in the image (x_i, y_i) find the external and internal parameters which will best map the control points to their image points. The phrase 'best' or 'optimal' is vague and requires that some measure of optimality be defined. Most methods try to minimize the mean distance between the observed position of the features on the image plane and the position computed from the known world coordinates, the camera position, and the camera parameters.

There is a problem that arises due to the interaction between the external and the internal parameters. If the data are perfect then only the correct calibration parameters can reduce the mean square error to zero. If there is noise in the data, as there inevitably will be, then an error in the value obtained for an external parameter might be compensated for by an error in the value obtained for an internal parameter. The Manual of Photogrammetry claims that "the strong coupling that exists between interior elements of orientation [principal point and focal length] and exterior elements can be expected to result in unacceptably large variances for these particular projective parameters when recovered on a frame-by-frame basis". Despite this problem the calibration can still be used to accurately measure the position of an object in the workspace from it's image position in a stereo image pair. In [22], a simulation of camera calibration with noisy data resulted in an error greater or equal to 0.5% in both the focal length and the external parameters. This happened even with a simulation of a simple pinhole camera model with no lens distortion. Despite this error it was possible to use these camera parameters to locate objects in space with an accuracy limited only by the accuracy of the feature detection in the image. It is likely though, that only objects in the specific volume of space occupied by the original calibration object may be located accurately. As one tries to measure the position of objects further and further away from the calibrated volume of space the error could increase. The problem also manifests itself when one tries to use subsets of the camera parameters for other tasks.

The individual parameters can be found more accurately by using calibration points from a large volume of space and with a large variation in depth. One method [17] uses multiple views of the calibration object from different camera positions with camera internal parameters kept constant. For each new view one gets six new external parameters to be calibrated but the internal parameters are forced to be consistent with all the views.

1.2.2 Methods that use geometric properties

There are a variety of methods that use geometric objects whose images have some characteristic that is invariant to the actual position of the object in space and can be used to calibrate some of the internal camera parameters. These methods do not require knowing or finding the position of the object relative to the camera.

The plumb line method [2], which uses the images of straight lines for calibrating lens distortion, a method for finding the aspect ratio using the image of a sphere [15], and a new method which uses spheres to locate the principal point and estimate the focal length are described in detail in section 3 in this paper.

In [3] the vanishing points of parallel lines drawn on the faces of a cube are used to locate the focal length and principal point. In [1] planar sets of parallel lines are rotated around an axis not perpendicular to the plane. The motion of the vanishing points due to the rotation is used to find the principal point.

1.2.3 Methods that do not require known calibration points

Since the traditional methods of camera calibration use known world coordinates they are not suitable for self calibration of mobile robots. It would be a great advantage to be able to calibrate a camera using only feature coordinates in the image plane. This cannot be done with a single image. Instead it requires camera motion.

Gennery [8] proposes using nonlinear optimization to calibrate a stereo system where the unknown parameters are the focal lengths and the relative position of the two cameras. It is assumed that the camera has no radial distortion and that all the other camera parameters are known. No information is given as to the accuracy of the results. Faugeras et al. [4] develop a method where a motion sequence of a single camera moving in an unconstrained manner can be used to calculate the internal camera parameters. A camera model with no lens distortion is used. Simulations show that assuming zero mean Gaussian noise of only 0.01 pixels this method produces errors of over 9% in the focal length. This method is currently too inaccurate for real use.

The rotation method developed here is another method that does not require known world coordinates of the calibration points. It requires only feature correspondences from a set of images where the camera has undergone pure rotation. The internal parameters of the camera can be determined including radial lens distortion. It has proven to be robust and accurate to a fraction of a percent with real cameras and images.

1.3 Brief overview of the rotation method

The basic idea of the rotation method is very simple. Given any pair of images produced by a camera which has undergone pure rotation, if the internal camera parameters and the angle and axis of rotation are known, one can compute where the feature points from one image will appear in the second image of the pair. If there is an error in the internal camera parameters, the features in the second image will not coincide with the feature locations computed from the first image. The object of the calibration is to find camera parameters that will enable one to correctly predict the effects of camera rotation in some optimal manner. We use pairs of images where the camera has rotated around a constant axis of rotation. We have chosen as a cost function to be minimized, the sum of squared distances between the feature points in the second image in each pair and those computed from the first image in each pair, summed over all the pairs. The parameters are found by a straightforward nonlinear search.

The constraints imposed on the motion result in a calibration method which is very robust to data noise and does not require a good initial guess of camera parameters. For example, nominal values of the camera parameters can be used as the initial starting point for the optimization procedure.

Camera lens distortion is taken into account by first correcting the feature coordinates in both images for lens distortion and only then computing the rotated feature points and evaluating the cost function. The lens distortion parameters can be added to the list of parameters to be determined in the non-linear search.

1.4 What is in the rest of the paper ?

- Section 2 describes the camera model and notation used. It then goes on to describe in detail the effects of camera rotation and translation on an image.
- Section 3 describes calibration using geometric objects which is used as an alternative camera calibration procedure.

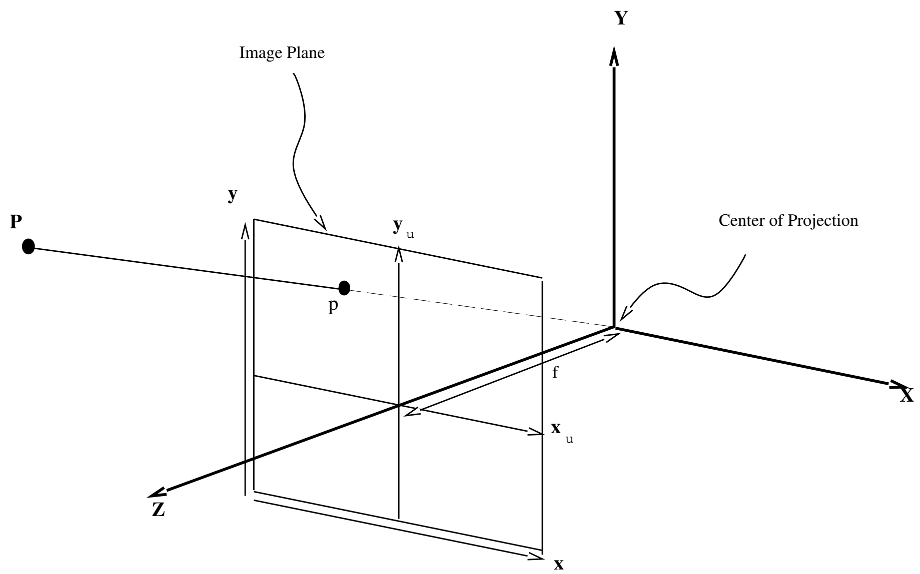


Figure 1: The camera coordinate system (X, Y, Z), the undistorted image coordinates (x_u, y_u) and the frame buffer coordinates (x, y)

- Section 4 describes the theory of the rotation method.
- Section 5 delves into the experimental details. Subsection 5.1 gives information about the cameras and the image acquisition hardware.
- Section 6 shows how, by rotating the camera through a full circle, one can avoid the need to measure the rotation angles.
- The results of the rotation method are presented and evaluated in section 7 and compared to the calibration methods using geometric objects.
- Section 8 discusses the main results and conclusions from this work.

2 Mathematical background

2.1 The perspective projection camera model with no lens distortion

We first build a camera centered Cartesian coordinate system. The image plane is parallel to the (X, Y) plane and at a distance f from the origin and the Z axis coincides with the optical axis (see figure 1).

Using perspective projection a point $\vec{P} = (X, Y, Z)$ in the camera centered coordinate system projects to a point $p_u = (x_u, y_u)$ on the image plane:

$$\begin{aligned} x_u &= fX/Z \\ y_u &= fY/Z \end{aligned} \tag{1}$$

The Z axis projects to the point $(0,0)$ in the image plane which is called the principal point (PP). Finally, $\vec{P}$ is converted to frame buffer coordinates:

$$\begin{aligned} x &= \frac{1}{S}x_d + c_x \\ y &= y_d + c_y \end{aligned} \tag{2}$$

where (c_x, c_y) is the principal point in frame buffer coordinates and S is the aspect ratio or scale factor of the image. The subscripts 'u' and 'd' are used to denote the undistorted and distorted image planes as used in the rest of the paper. We are now considering the case in which there is no lens distortion, therefore $x_d = x_u$ and $y_d = y_u$.

In order to understand the meaning of the parameter S it is important to understand the imaging process, which is the transfer of image information from the image plane to the computer memory. In CCD cameras, the image plane is a rectangular array of light sensor elements, which gives a discrete spatial sampling of the image plane. As in figure 1, we will call the vertical columns of the array Y and the horizontal rows X. The measure of light intensity falling on the sensors is read out sequentially, row after row and element after element along each row and transferred to the computer memory. The rate at which the sensor values are read is called the pixel clock frequency f_p.

The computer memory, which is often called the frame buffer, can also be viewed as an array. The rows of the CCD array are read out sequentially[1] into the frame buffer so that the i^{th} row in the frame buffers corresponds to the i^{th} row CCD array. In digital cameras each individual element along a row of the CCD array is digitized and transferred to the frame buffer, so the j^{th} element along each row in the frame buffer corresponds to the j^{th} element along the row of the CCD array. In regular video cameras the CCD values along each row are output as a time varying analog voltage signal which is resampled at a frequency f_s. Since the original pixel clock frequency is lost in the process it is no longer true to say that a specific element along a row in the frame buffer corresponds to a specific element along a row in the CCD array. In some cameras the pixel clock is available as a separate output or input and using these cameras it is possible to build systems that behave as digital cameras in this respect.

Since there is always a one-to-one correspondence between rows in the frame buffer and rows in the CCD array it is convenient to measure lengths such as the focal length in terms of vertical (or Y) pixels. In digital cameras the aspect ratio of the image S depends on the ratio of the vertical and horizontal distances between the centers of adjacent sensor elements in the image array. Let d_y and d_x be the vertical and horizontal distances respectively. Then:

$$S = \frac{d_x}{d_y} \tag{3}$$

In video camera systems S also depends on the ratio between the pixel clock frequency

[1] We do not go into the issue of interlaced and non-interlaced cameras.

and the sample clock frequency:

$$S = \frac{d_x}{d_y}\frac{f_s}{f_p} \tag{4}$$

The dimensions of the CCD array and the pixel clock frequency can be found from the camera data sheets provided by the manufacturers. The sample clock frequency is given by the manufacturer of the frame grabber system but because of the use of phase locked loops this frequency can vary [11].

2.2 Camera model with lens distortion

Off the shelf cameras and lenses which are often used for computer vision have large amounts of lens distortion. This is particularly noticeable in the case of wide angle lenses. The standard model for this distortion [17] takes into account both radial distortion and decentering distortion. This model gives the correction for lens distortion. In other words, it is a mapping from the distorted image coordinates, that are observable, to the undistorted image plane coordinates which are not physically measurable. The undistorted image plane coordinates are:

$$\begin{aligned} x_u &= x_d + \delta x \\ y_u &= y_d + \delta y \end{aligned} \tag{5}$$

where $\delta x, \delta y$ are the corrections for lens distortion and can be found by:

$$\begin{aligned} \delta x &= x_d(K_1 r_d^2 + K_2 r_d^4 + \cdots) + [P_1(r_d^2 + 2x_d^2) + 2P_2 x_d y_d][1 + P_3(r_d^2 + \cdots] \\ \delta y &= y_d(K_1 r_d^2 + K_2 r_d^4 + \cdots) + [P_2(r_d^2 + 2y_d^2) + 2P_1 x_d y_d][1 + P_3(r_d^2 + \cdots] \end{aligned} \tag{6}$$

and where:

$$r_d^2 = x_d^2 + y_d^2 \tag{7}$$

The first term in (6) is the radial distortion with parameters K_1 and K_2. The second term is the decentering distortion with parameters P_1,P_2 and P_3. Decentering distortion is due to misalignment of the individual lens elements and the non-perpendicularity of the lens assembly and the image plane.

We can derive an equivalent expression for the decentering lens distortion as follows. Suppose the optical axis of the lens is not perpendicular to the image plane. Then the optical axis will no longer pass through the principal point (which we have defined as the point where a line through the center of projection that is perpendicular to the image plane intersects the image plane) but through another point which we will call the point of best radial symmetry (c_{xr}, c_{yr}).

We perform a coordinate shift to the point of best radial symmetry by substituting $x'_d = x_d - c_{xr}$ and $y'_d = y_d - c_{yr}$ and calculate the radial distortion using (6):

$$\begin{aligned} \delta x &= x'_d(K_1 r_d^{'2} + K_2 r_d^{'4} + \cdots) \\ \delta y &= y'_d(K_1 r_d^{'2} + K_2 r_d^{'4} + \cdots) \end{aligned} \tag{8}$$

where, as in (6):

$$r_d^{'2} = x_d^{'2} + y_d^{'2} \tag{9}$$

Taking only the first coefficient of radial distortion, K_1 and expanding (8) we get:

$$\delta x = x_d K_1 r_d^2 - K_1[c_{xr}(r_d^2 + 2x_d^2) + 2c_{yr}x_d y_d] + K_1[x_d c_{yr}^2 + 2x_d c_{xr}^2 + 2y_d c_{xr} c_{yr} - c_{xr}^3 - c_{yr}^2 c_{xr}] \quad (10)$$
$$\delta y = y_d K_1 r_d^2 - K_1[c_{yr}(r_d^2 + 2y_d^2) + 2c_{xr}y_d x_d] + K_1[y_d c_{xr}^2 + 2y_d c_{yr}^2 + 2x_d c_{yr} c_{xr} - c_{yr}^3 - c_{xr}^2 c_{yr}]$$

The first term in (10) is the radial distortion term in (6). The second term is the decentering distortion term with $P_1 = -K_1 c_{xr}$ and $P_2 = -K_1 c_{yr}$. The third term is small since c_{xr} and c_{yr} are usually only at most tens of pixels. So the two forms are for our purposes equivalent.

2.3 Pure rotation using Rodriguez's formula

Let us rotate a camera in a rigid world environment around a rotation axis $\vec{W} = (w_x, w_y, w_z)$ through an angle of $(-\theta)$ degrees. In the camera coordinate system this is equivalent to saying that the world rotated through an angle of (θ) degrees around the axis $\vec{W}$. A world point $P = (X, Y, Z)$ in the camera coordinate system will move to $\vec{P'}$. $\vec{P'}$ can be found using Rodriguez's formula:

$$\vec{P'} = \mathbf{R}\vec{P} = \left[\cos\theta\mathbf{I} + \sin\theta\mathbf{Q} + (1-\cos\theta)\hat{W}\hat{W}^t\right]\vec{P} \quad (11)$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{Q}$ is the matrix given by:

$$\mathbf{Q} = \begin{bmatrix} 0 & -w_z & w_y \\ w_z & 0 & -w_x \\ -w_y & w_x & 0 \end{bmatrix} \quad (12)$$

The point $\vec{P'}$ projects to point $p' = (x', y')$ on the image plane where:

$$x' = f\frac{X'}{Z'} = f\frac{(r_{11}X + r_{12}Y + r_{13}Z)}{(r_{31}X + r_{32}Y + r_{33}Z)} \quad (13)$$
$$y' = f\frac{Y'}{Z'} = f\frac{(r_{21}X + r_{22}Y + r_{23}Z)}{(r_{31}X + r_{32}Y + r_{33}Z)}$$

Where r_{ij} is j^{th} element along the i^{th} row of the matrix $\mathbf{R}$. Multiplying through by $\frac{f}{Z}$ and substituting $x = f\frac{X}{Z}, y = f\frac{Y}{Z}$ gives :

$$x' = f\frac{(r_{11}x + r_{12}y + r_{13}f)}{(r_{31}x + r_{32}y + r_{33}f)} \quad (14)$$
$$y' = f\frac{(r_{21}x + r_{22}y + r_{23}f)}{(r_{31}x + r_{32}y + r_{33}f)}$$

We can conclude from (14) that the position of the point in the image after rotation depends only on the camera parameters, the rotation and the location of the point in the image before rotation. The 3D world coordinates are not required. In other words, if the camera parameters are known and the camera undergoes a known, pure rotation we can calculate the location of a feature in the new image given the coordinates of the feature in the image before rotation.

For rotation around the Y axis ($\vec{W} = (0, 1, 0)$), $\mathbf{R}$ becomes:

$$\mathbf{R} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \tag{15}$$

and equation (14) becomes:

$$\begin{aligned} x' &= f\frac{(\cos\theta x + \sin\theta f)}{(-\sin\theta x + \cos\theta f)} \\ y' &= f\frac{y}{(-\sin\theta x + \cos\theta f)} \end{aligned} \tag{16}$$

2.4 Rotation with translation and the importance of pure rotation

If the axis of rotation does not pass exactly through the center of projection then there will be some translation $\vec{T} = (t_x, t_y, t_z)$ in addition to the rotation around the center of projection. If d is the perpendicular distance from the center of projection to the axis of rotation then:

$$\|T\| = 2d\sin\frac{\theta}{2} \tag{17}$$

For $\theta = 60^o$, which is the largest angle of rotation used, we get $\|T\| = d$. For a translation vector $\vec{T} = (t_x, t_y, t_z)$ the location of an image point after rotation and translation will be:

$$\begin{aligned} x' &= f\frac{(r_{11}x + r_{12}y + r_{13}f + ft_x/Z)}{(r_{31}x + r_{32}y + r_{33}f + ft_z/Z)} \\ y' &= f\frac{(r_{21}x + r_{22}y + r_{23}f + ft_y/Z)}{(r_{31}x + r_{32}y + r_{33}f + ft_z/Z)} \end{aligned} \tag{18}$$

The location of the point is no longer independent of the depth and depends also on the translation vector. One can see from (18) that if the object is distant then the effect of the translation becomes negligible.

A few numerical examples help to get a feel for the problem. For the Chinon camera used in the experiments the pixel elements in the CCD array are approximately $6\mu m$ square and $f = 3$mm which is about 500 pixels. Suppose $\|T\| = 0.25$mm or about 40 pixels (we show in section 5.2 that this is an obtainable value) and suppose $Z = 1.0$m or 1.7×10^6 pixels. Then $\frac{ft_z}{Z} \leq 0.12$ pixels and $\arctan\left(\frac{t_z}{Z}\right) \leq 0.8$arc minutes which is a smaller angle than we can measure with the vernier scale of our rotation stage and at least two orders of magnitude smaller than angles used in the experiment. For an object 100m away, the translation and the corresponding error will be smaller by a factor of 100.

3 Calibration using Geometric objects

3.1 Finding the aspect ratio using spheres

An interesting method for finding the aspect ratio was developed in [15]. We will make slight changes in the derivation mostly in order to fit in with notation used in this paper. The image of the occluding contour of a sphere under perspective projection is an ellipse. (The original paper claimed that it is a circle but this does not change the results.) In other words, the undistorted image coordinates fit the function:

$$\frac{(x_u - x_0)^2}{r_x^2} + \frac{(y_u - y_0)^2}{r_y^2} = 1. \tag{19}$$

The decentering lens distortion term in (6) is small compared to the radial distortion term. Taking only the first term of radial distortion, equation (6) becomes:

$$\begin{aligned} x_u &= x_d(1 + kr_d^2) \\ y_u &= y_d(1 + kr_d^2) \end{aligned} \tag{20}$$

where:

$$r_d^2 = x_d^2 + y_d^2 \tag{21}$$

Using the first term from the Taylor series expansion of $(1 + x)^{-1}$, equation (20) becomes:

$$\begin{aligned} x_u &= x_d(1 + kr_d^2) \approx \frac{x_d}{1 - kr_d^2} \\ y_u &= y_d(1 + kr_d^2) \approx \frac{y_d}{1 - kr_d^2} \end{aligned} \tag{22}$$

Solving (2) for x_d and y_d results in:

$$\begin{aligned} x_d &= S(x - c_x) \\ y_d &= (y - c_y) \end{aligned} \tag{23}$$

Substituting (23) in (22), then substituting the result in (19) and expanding results in a 4th order polynomial equation in x and y, where x and y are the image buffer coordinates:

$$a_0x^4 + a_1x^2y^2 + a_2y^4 + \cdots + (lower\, order\, terms) = 0 \tag{24}$$

The coefficients for the higher terms turn out to be:

$$a_0 = S^4k^2(r_y^2x_0^2 + r_x^2y_0^2 - r_x^2r_y^2) \tag{25}$$

$$a_1 = 2S^2k^2(r_y^2x_0^2 + r_x^2y_0^2 - r_x^2r_y^2) \tag{26}$$

$$a_2 = k^2(r_y^2x_0^2 + r_x^2y_0^2 - r_x^2r_y^2) \tag{27}$$

From (25) and (27) one can compute S:

Table 1: Finding the aspect ratio using image of a sphere

camera/lens	number of images	4th order fit		Fitting to Ellipse	
		mean S	σ_S	mean S	σ_S
Sanyo 8.5mm	6	1.2586	0.0015	1.2580	0.0012
Chinon 3mm	12	1.2469	0.0024	1.2472	0.0027

$$S = (\frac{a_0}{a_2})^{\frac{1}{4}} \tag{28}$$

Based on this analysis one can prescribe a method as follows. Take an image of a sphere close to the center of the image. Find the occluding contour of the sphere and fit to a 4th order polynomial in x and y. Divide the coefficient of the x^4 term by the coefficient of the y^4 term and take the fourth root to get S.

One might ask, "Why not simply assume that the image of the sphere in the distorted image plane is a circle and that the elongation in the frame buffer image is due to the aspect ratio alone ?". In other words, we might assume that the effects of the radial distortion and the effects of the perspective projection are very small. Then, one can fit the image of the sphere to an ellipse and use the ratio of the major and minor axis to calculate the scale factor. Our simulations, using realistic values of focal length and principal point, and our experimental results using real images show that both methods work equally as well.

The sphere used was a $1.25''$ billiard ball, which was painted with a mat black paint, on a white background. A Canny edge detector was used to find the occluding contour and the results were fitted to a 4th order polynomial. This procedure was repeated a number of times with the location of the sphere moved to random locations ± 25 pixels from the center of the frame buffer. The same edge data were fitted to an ellipse whose axes were parallel to the x, y axis.

Table 1 summarizes the results. The results using the two methods gave virtually the same results within a twentieth of a percent. The variance was similar and on the order of a few tenths of a percent. Results for the higher quality, higher resolution, Sanyo camera had a smaller variance.

There are points to note about the 4th order method:

- This method gives a very simple way of finding the aspect ratio to within a fraction of 1% accuracy. It is simple to perform and can be completely automated.
- The approximation in (22) assumes $kr_d^2 \ll 1$. Therefore when using lenses with large radial distortion, points on the sphere mustn't be too far from the image center. In other words the sphere must be centered close to the center of the image and mustn't be too large.
- There are some sphere locations where the coefficients a_0 and a_2 change sign. These produce spikes in the values of S.

3.2 Finding the radial distortion using straight lines

The image of a straight line under perspective projection is a straight line. In practice the image is not straight due to lens distortion. If one were to take the image of a set of straight lines, such as plumb lines, one could find correction parameters which would best straighten out the lines in the image using (6) or (8). The advantage of this method is that neither the position nor orientation of the lines relative to the camera nor the internal camera parameters such as focal length or principal point need to be known.

Brown [2] fully presents the mathematics of this method and uses plumb lines for the calibration of a high accuracy metric camera for both radial and decentering lens distortion. The camera used has a wide image format. The lens had a focal length of about 135 mm and viewing angle of about 60^o. The results in [2] show that the lens distortion parameters change both with the distance for which the lens is focused and with the actual object distance. In those results the radial distortion of the lens when focused at 3 ft and 6 ft were 60% and 80% of the distortion at infinity respectively.

The equations for determining the radial distortion δr_s for a lens focused at distance s can be computed from the radial distortion measured at two other focus distances, s_1 and s_2 using the equations derived in [2].

$$\delta r_s = \alpha_s \delta r_{s_1} + (1 - \alpha_s)\delta r_{s_2} \tag{29}$$

where

$$\alpha_s = \frac{s_2 - s}{s_2 - s_1}\frac{s_1 - f}{s - f} \tag{30}$$

Figure 2 shows the change in radial distortion as a function of the focus distance for the lens used in [2]. One can see that for distances less than 100 times the focal length the change might be significant. In our experiments the objects were no less than 50 focal lengths from the camera and so the radial distortion parameters were close to those at infinity.

Fryer [6] uses the plumb line method for the calibration of video cameras. Very sharp lines on a printed writing pad are used as the calibration pattern.

As the calibration pattern we used black bars against a white background. We used a pair of images, one with horizontal bars and one with vertical bars. Figure 3 shows one of the image pairs from the Chinon camera. The edge points were found to subpixel resolution and then clustered into lines using 8 way connection (i.e two edge points were clustered together if one was in any of the eight cells adjacent to the other). The x coordinates were corrected for the aspect ratio using the aspect ratio obtained in section 3.1. Given a set of distortion correction parameters, the edge points in each cluster were corrected for lens distortion. A straight line was fitted to each of the edges and the RMS distance of the points from line was computed. The best parameters are those parameters that reduce the RMS error to a minimum and could be found using non-linear optimization.

Fryer [6] used the decentering terms in (6) and it was therefore necessary to choose a point on the image as the principal point. We chose to use equation (8) which does

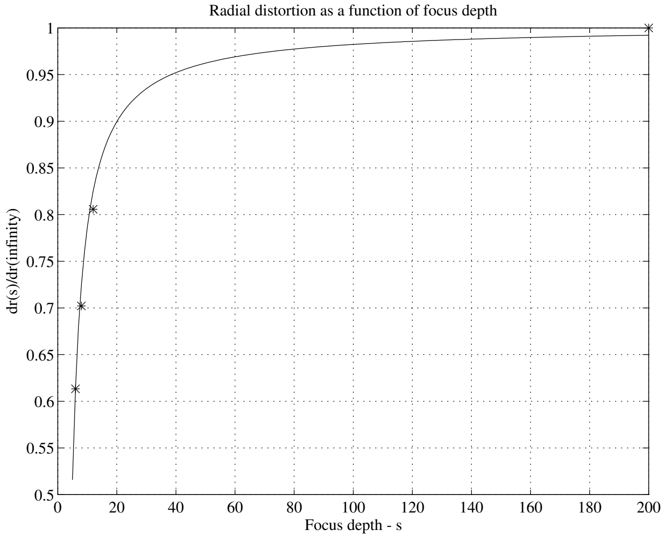


Figure 2: Radial distortion δr_s of a lens focused at varying depths s normalized by the distortion at $s = \infty$ (δr_{s_∞}) and for s in units of focal length. Theoretical curve fitted to data from [2].(*) denotes actual data .

not have a decentering term but accounts for the decentering distortion by using the point of best symmetry rather than the principal point as the center of the radial distortion. In this way we could perform nonlinear search to obtain all the distortion parameters and avoided the need to arbitrarily choose a principal point.

In some of our work we used the edges of black bars printed by a laser printer and attached to a flat acrylic sheet. This was a very convenient calibration object because it was easy to produce and one image would give a large number of lines (usually around 20) covering the whole image. The size of the laser printer page limited the size of the calibration object and this limited the maximum distance the calibration object could be placed from the lens and still fill the screen. If the calibration is too close then the radial distortion parameters found are not the distortion parameters for infinity. The maximum distance one can position the object and still fill the screen when measured in units of focal lengths depends on the size of the object and size of the imaging surface of the camera:

$$N = \frac{length\,of\,object}{length\,of\,image\,sensor} \tag{31}$$

The Chinon camera has a $\frac{1}{3}$ inch imaging sensor and the sheet could be positioned 60 focal lengths away. The experiment was repeated 6 times with the object at a distance of 0.2m. The results of using the plumb line method for the Chinon camera are presented in table 2 and table 3. Due to the large radial distortion both the 2nd order and 4th order coefficients of radial distortion (i.e K_1 and K_2) and the point of best symmetry (i.e. center of radial distortion) could be found with this method. By comparing the RMS error in table 2 and table 3 we can see that adding the 4th order term in (8) greatly improves the fit (reducing the RMS error by about 45%) and we

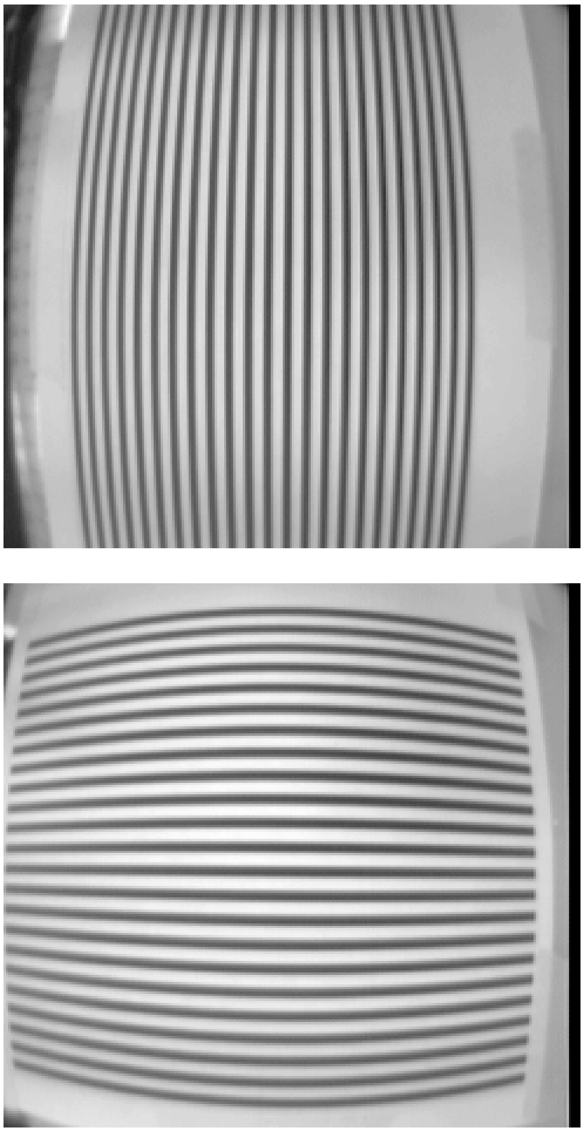


Figure 3: Typical images of vertical and horizontal bars taken with the Chinon camera and a wide angle lens. These images were used to calibrate the radial distortion parameters with the 'plumb line' method

can conclude that for this camera and lens both the 2nd order and 4th order terms in (8) are significant.

The Sanyo has a $\frac{2}{3}$inch sensor and so the sheet could only be positioned 25 focal lengths (0.15m) away and this was found to be too small. The radial distortion parameters found using an object placed 0.15m from the camera were significantly different than the radial distortion parameters found using an object placed 1m from the camera. Instead of bars printed on a sheet of paper, we used a 1m long ruler painted black and mounted on a white background. The ruler could be positioned 1m (or 120 focal lengths) away from the camera. A set of images was taken with the ruler at various positions in the image. About 10 vertical and 10 horizontal bars were imaged.

The experiment was repeated 3 times with the bar positioned at 1m, 0.5m and 0.2m and the radial distortion parameters found. Then the experiment was performed with the printed lines placed at a distance of 0.15m from the camera. The focus and aperture were held constant throughout. This was possible because a wide angle lens was used which had a large depth of field.

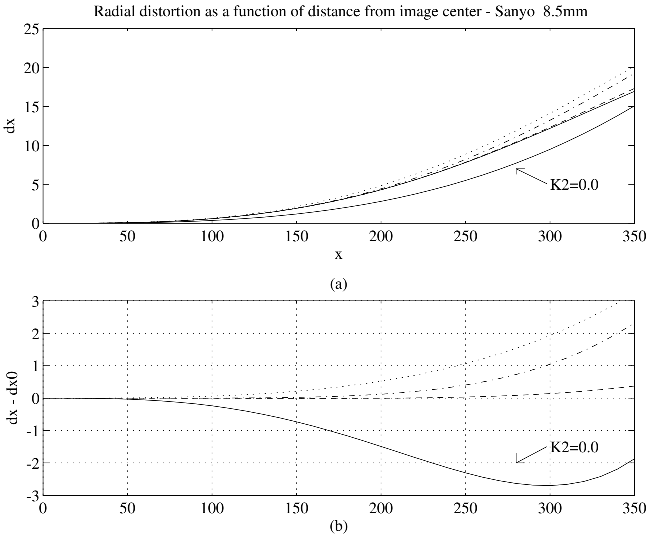
The results for the Sanyo camera are shown in table 4 and table 5. The drop in RMS error when we add the second radial distortion parameter is between 30 and 50 percent and we can conclude that the second radial distortion parameter gives a significant improvement. We can also see that the parameters vary with the distance to the calibration object with larger parameters obtained from nearer objects. Figure 4 shows the correction for radial distortion as a function of the distance from the center of the image using the results from table 4. In order to see the difference clearly we subtracted the correction value obtained from the images where the rod was placed at $1m$ from the camera from the correction values obtained in the other experiments (figure 4b). As we can see the size of the correction (δ_x) is larger using the radial distortion parameters obtained from straight objects placed closer to the camera. Also shown in figure 4 is the correction using a single radial distortion parameter. The correction values are over 2 pixels different.

The change in radial distortion as a function of object distance has an opposite sign from the results described in [2] and from the results shown in figure 2. Note that the results in figure 2 show the change in radial distortion due to a change in object distance and focus distance but other results in [2] show the same trend when focus distance is kept constant. We have no explanation for the difference between our results and those in [2]. Both directions of change can be obtained from (29) and it might depend on the lens used.

3.3 Finding the principal point and focal length using spheres

The image of a sphere under perspective projection is an ellipse whose axis of least inertia (i.e. major axis) is on a line passing through the principal point. This is shown geometrically in figure 5. This is true no matter where the ellipse is in the image or where the sphere is in the world (as long as it is in view).

The eccentricity of the ellipse is a function of the focal length, the distance of the center of the ellipse from the principal point and of the length of the major axis [20].



Legend:	
——	Ruler at 1m
- - -	Ruler at 0.5m
_._._	Ruler at 0.2m
.....	Laser print at 0.1m

Figure 4: (a) The correction for radial distortion (δ_x) plotted against the distance from the principal point for the Sanyo 8.5mm lens. Each line uses the correction parameters obtained from straight objects placed on planes at different distances from the image plane. Two parameters for radial (K_1,K_2) distortion were used. For comparison we also show the correction for radial distortion using a radial distortion model with only a single parameter (i.e $K_2 = 0.0$). For the single parameter model the calibration objects were on a plane 1m from the image plane. (b) The result of correction parameters (K_1,K_2) obtained from an object at 1m from the camera (i.e. the upper solid line in figure (a)) subtracted from the rest of the results in (a).

Table 2: Results of plumb line method for Chinon camera using two coefficients for radial distortion (K_1,K_2)

S=1.247					
Test image # Units:	c_{xr} $Pixels$	c_{yr} $Pixels$	K_1 $Pixels^{-2}$	K_2 $Pixels^{-4}$	RMS Error* Pixels
1	258.5	229.9	1.8132e-06	6.3707e-12	0.3668
2	258.5	229.9	1.8011e-06	6.4933e-12	0.3940
3	256.0	230.0	1.8000e-06	6.5353e-12	0.3668
4	257.6	229.6	1.7992e-06	6.3072e-12	0.3920
5	256.1	229.7	1.7721e-06	6.6650e-12	0.4100
6	256.7	229.9	1.8225e-06	6.2449e-12	0.4080
mean	257.2	229.8	1.801e-6	6.436e-12	
$\sigma/mean$	0.0043	0.0006	0.0095	0.0244	

(*) The error measured is the distance of each corrected point from the best straight line passing the set of points (see text).

Table 3: Results of plumb line method for Chinon camera using only one coefficient of radial distortion (K_1)

S=1.247 k2=0.0				
Test image #	c_{xr}	c_{yr}	K_1	RMS Error
1	258.3	229.3	2.4707e-06	0.7200
2	258.7	229.2	2.4765e-06	0.7536
3	255.6	229.2	2.4813e-06	0.7600
4	256.9	229.5	2.3545e-06	0.5892
5	255.3	229.5	2.3695e-06	0.6388
6	255.6	229.7	2.3912e-06	0.6248
mean	256.7	229.4	2.4240e-06	
$\sigma/mean$	0.0058	0.0009	0.0241	

Table 4: Results of plumb line method for Sanyo camera and an 8.5mm lens using two coefficients for radial distortion (K_1,K_2)

S=1.259						
Method	distance	c_{xr}	c_{yr}	K_1	K_2	RMS Error
Ruler	1m	244.9	245.2	6.0713e-07	-1.7296e-12	0.144
Ruler	0.5m	244.1	243.9	6.0243e-07	-1.6200e-12	0.138
Ruler	0.2m	243.5	242.7	6.0348e-07	-1.2584e-12	0.143
Laser	0.15m	242.7	242.1	6.6749e-07	-1.6051e-12	0.198
	mean	243.9773	242.8614	6.3438e-07	-1.5569e-12	
	$\sigma/mean$	0.0055	0.0055	0.045	-0.094	

Table 5: Results of plumb line method for Sanyo camera and an 8.5mm lens using only one coefficient of radial distortion (K_1)

S=1.259 k2=0.0					
Method	distance	c_{xr}	c_{yr}	K_1	RMS Error
Ruler	1m	253.1	245.2	3.5155e-07	0.270
Ruler	0.5m	255.6	241.7	3.5877e-07	0.263
Ruler	0.2m	246.2	241.7	4.3423e-07	0.202
Laser	0.15m	248.2	240.6	4.2083e-07	0.303
	mean	249.2	241.8	4.042e-07	
	$\sigma/mean$	0.017	0.0071	0.095	

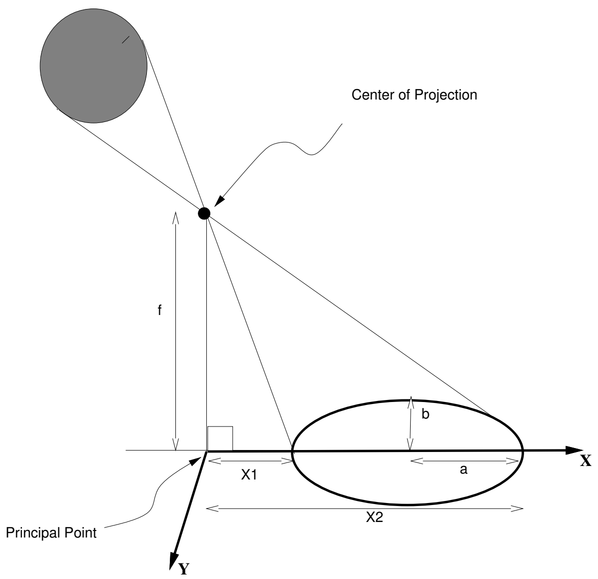


Figure 5: The image of a sphere under perspective projection is an ellipse.

With the eccentricity defined as:

$$e = \frac{\sqrt{a^2 - b^2}}{a} \tag{32}$$

we find that:

$$e = \frac{\sqrt{f^2 + x_1^2}\sqrt{f^2 + x_2^2} + x_1 x_2 - f^2}{\sqrt{f^2 + x_2^2}x_1 + \sqrt{f^2 + x_1^2}x_2} \tag{33}$$

where the geometric interpretation of the variables a,b,x_1,x_2 is shown in figure 5. Note that:

$$x_0 = \frac{1}{2}(x_1 + x_2) \tag{34}$$

and:

$$a = \frac{1}{2}(x_1 - x_2). \tag{35}$$

In theory, given the images of a set of spheres it would be possible to find the principal point, which is the intersection of the major axes. The focal length can then be found using the eccentricity of one of the ellipses, the principal point and (33).

The eccentricity of the ellipse is of the order of 10% for wide angle lenses. Nevertheless, in practice there are two other phenomena which are larger than this one. The radial distortion squashes the ellipse along it's major axis and the result looks like an ellipse whose minor axis passes through the center of the image. The aspect ratio of the image is the strongest phenomenon and makes the ellipse elongated along one of the image axes.

Figure 6 shows the actual image of a set of spheres. This image was created by placing 1.25″ diameter black spheres on a light box (the type used for viewing negatives and slides) in a lighted room. The light box eliminated the shadows under the spheres.

We first found the occluding contour of each sphere in figure 6 using an edge detector, then corrected the edge values for aspect ratio and radial distortion and finally fit the resulting points with an ellipse. The resultant ellipses and the major axes are shown in figure 7. Finding the intersection point is a simple linear algebra problem of finding a point with least RMS distance from all the lines.

While most of the ellipses pointed in the right direction a few were very far off. This was probably due to problems in the image processing such as shadows in the original image. The method for dealing with outliers was as follows. The intersection point was found using all the spheres. Then the distance from each line to the intersection was computed and the lines that were more than twice the median distance away were discarded. Then the intersection was recomputed using the remainder of the lines. In our experiment 24 out of 31 of the spheres were used and the rest were discarded. Figure 7 shows the PP found and the point of best symmetry used for correction of radial distortion. Table 6 shows the results. Note that the distance between the principal point and the point of best symmetry is over 3.5 pixels, which

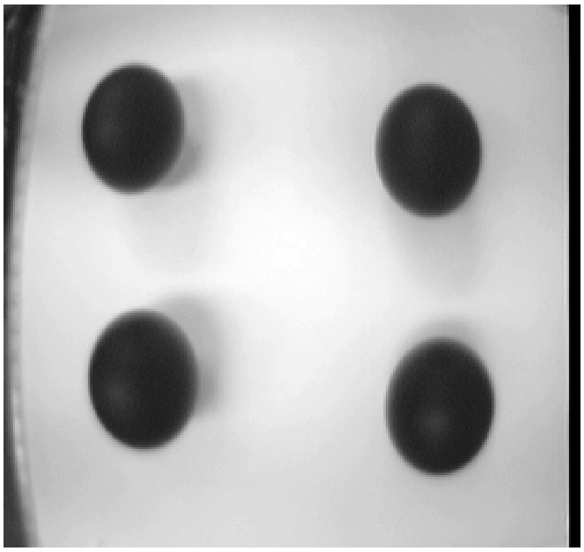


Figure 6: The image of 4 spheres used for locating the principal point.

Table 6: Results of using spheres for finding the principal point of the Sanyo 8.5mm camera. Note that the principal point coordinates are found in scaled image coordinates and the c_x must be scaled by $\frac{1}{S}$ to get the PP in image buffer coordinates.

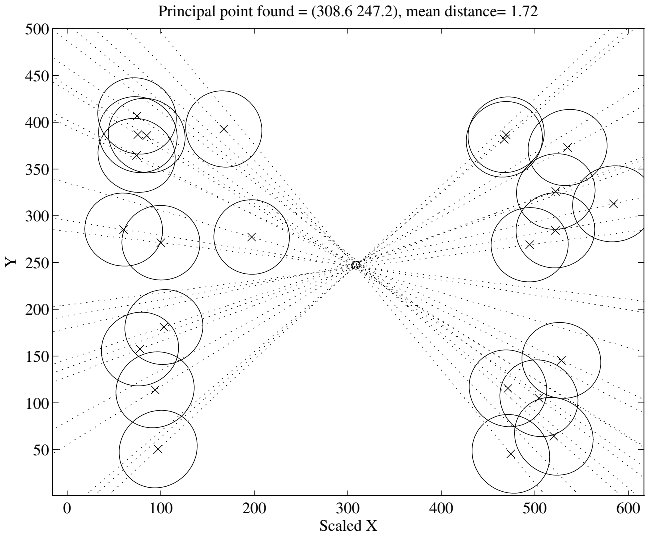
Correction parameters used	S=1.259	$K_1 = 6.024e - 07$	$K_2 = -1.62e - 12$	
for objects at 0.5m.	$c_{xr}S = 307.4$	$c_{xr} = 244.2$	$c_{yr} = 243.9$	
PP found with				
mean(d)=1.72	$c_x S = 308.6$	$c_x = 245.1$	$c_y = 247.2$	

is over twice the mean distance from the principal point to the lines. This means that the two points are distinctly different.

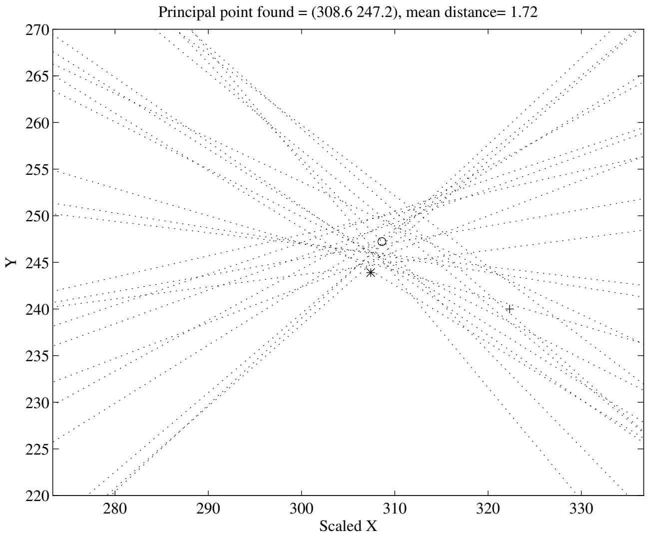
A focal length estimate was found from each ellipse by solving (33) numerically using bisection. Figure 8 shows the focal length estimate obtained from each ellipse. As one can see the standard deviation is very large (mean(f)=664.3 and $\sigma_f = 12.9$ pixels).

4 The rotation method - theory

The object of calibrating the internal camera parameters is to be able to determine, in camera coordinates, the direction to a point in space given the (x, y) coordinates of it's image. We can define this objective in another way. Using the calibrated camera model, we would expect to be able to determine what camera rotation would cause the image of a point in space to move from (x_1, y_1) to the new image coordinates (x_2, y_2). We note that the solution is not unique. This second definition of the objective is the basis of the rotation method. Given any pair of images where the camera has undergone pure rotation, if the internal camera parameters and the angle and axis



(a)



(b)

Figure 7: Finding the principal point using spheres (Sanyo 8.5mm) (a) The ellipses after correcting for aspect ratio and radial distortion. Note that all the X values have been scaled by $S = 1.259$. The major axes of the ellipses are shown. The lines intersect at the principal point which was found to be PP=$(308.6, 247.2)$. (b) Enlargement of the area around the intersection. 'o' marks the principal point, '*' marks the point of best symmetry of radial distortion $(307.4, 243.9)$ and '+' marks the center of the frame buffer $(322, 240)$.

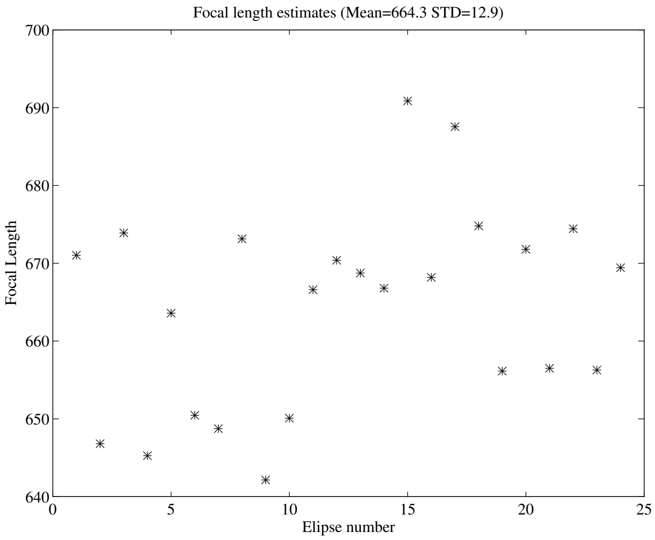


Figure 8: The estimated focal length obtained from each of the ellipses (mean(f)=664.3 and $\sigma_f = 12.9$). A Sanyo camera with an 8.5mm lens was used.

of rotation are known one can compute, using (14), where the feature points from one image will appear in the second image of the pair. If there is an error in the internal camera parameters, the features in the second image will not coincide with the feature locations computed from the first image. The object of the calibration is to find camera parameters which will enable one to best predict the effects of camera rotation in some optimal manner. We have chosen to minimize the sum of squared distances between the feature points in the second image in each pair and those computed from the first image in each pair, summed over all the pairs of images used.

To be more precise, M pairs of images are taken with the camera rotated at various angles. The axis of rotation is held constant for all the pairs. The relative angles of rotation are measured precisely. Corresponding features in each of the pairs are found and their coordinates extracted. There is no reason for the same number of features to be found in each image pair but that is what we did in practice. The number of features found will be denoted N.

First one must move to the undistorted image plane coordinates. Given the feature coordinates in the image buffer coordinates we can compute the feature coordinates in the distorted image plane using (2):

$$\begin{aligned} x^l_{d_{ij}} &= S(x^l_{ij} - c_x) \\ y^l_{d_{ij}} &= (y^l_{ij} - c_y) \end{aligned} \tag{36}$$

where (x^l_{ij}, y^l_{ij}) are the coordinates of feature i in image pair j. $l = 1, 2$ denotes whether it is the first or second image in the pair. We can use (8) to compute the feature coordinates in the undistorted image plane:

$$x^l_{u_{ij}} = x^l_{d_{ij}} + \delta x^l_{d_{ij}} \tag{37}$$
$$y^l_{u_{ij}} = y^l_{d_{ij}} + \delta y^l_{d_{ij}}$$

We can now define a cost function:

$$E = \sum_{i=1}^{N}\sum_{j=1}^{M}((x^{1'}_{u_{ij}} - x^2_{u_{ij}})^2 + (y^{1'}_{u_{ij}} - y^2_{u_{ij}})^2) \tag{38}$$

where $(x^{1'}_{u_{ij}}, y^{1'}_{u_{ij}})$ are the coordinates of $(x^1_{u_{ij}}, y^1_{u_{ij}})$ rotated using (14).

The task is now to find camera parameters $(f, c_x, c_y, S, K_1, K_2, c_{xr}, c_{yr})$ and the axis of rotation $\vec{W} = (w_x, w_y, w_z)$ that minimize E in (38). The parameters are found by a straightforward nonlinear search.

This is not the whole story. In order to obtain accurate values for the aspect ratio S it is necessary to use two sets of image pairs. In one the rotation is around the vertical (Y) axis and in the second the rotation is around the horizontal (X) axis. The rotation need not be precisely around the (X) or (Y) axis and the exact axis of rotation can be found through the minimization of E in (38).

5 The rotation method - experimental details

5.1 Hardware

Two B/W CCD cameras with standard video output were used for the experiments:

1. A Sanyo VDC3860 High Resolution Camera with 8.5 mm and 12.5 mm Comiscar lenses.

2. A Chinon CX-101 Low Resolution camera. This camera has a fixed focus 3 mm lens. The camera uses a $\frac{1}{3}$inch CCD and has an angle of view of 110^o along the diagonal.

The images were acquired using a Datacube framegrabber into 512×485 frame buffers and processed on a Sparc workstation.

For the rotation experiments the cameras were mounted on an XY stage with micrometer positioning (DAEDAL model 3972M). The XY stage was mounted on a precision rotary stage with a vernier scale that can measure angles to an accuracy of $\pm 1'$. $(1' = \frac{1}{60}^o)$. Experiments performed in section 5.5 show that an accuracy of better than $\pm 2'$ was actually achieved.

5.2 How to obtain pure rotation ?

As shown in section 2.4 it is very important that we obtain what is nearly pure rotation (i.e. the axis of rotation passes very close to the center of projection). To obtain pure rotation the camera is mounted on an XY stage that has a micrometer adjustment. This in turn, is mounted on the rotation stage. The camera can now be positioned so that the axis of rotation passes through the center of projection.

The test whether there is little or no translation is quite simple and involves the use of parallax. Let us suppose two points $\vec{P}_1, \vec{P}_2$ and the center of projection (0,0,0) all lie on a straight line. The image of the points $\vec{P}_1, \vec{P}_2$ both fall on the same point in the image plane. If the camera rotates about the center of projection then the three points will always lie on a straight line. On the other hand, if there is some translation of the center of projection then the three points will no longer be on a straight line and the images of the two points $\vec{P}_1, \vec{P}_2$ will separate. There is, of course, one special case. If the translation is along the line passing through the three points then they will stay on the same line. But since the translation is due to rotation about some point, the direction of translation continuously changes and therefore cannot be but momentarily aligned with the three points.

This test was performed experimentally as follows. The rotation was around the vertical axis. We took a white page with a set of vertical black lines, printed with a laser printer and mounted the page on a flat board and positioned it 1.0 m from the camera. We then took another page with one black line about one third the thickness of the previous ones and mounted it on a transparent acrylic sheet. The second page we positioned about 0.3 m from the camera in front of the first page so that the line in the second sheet seemed to be a perfect continuation of one of the lines in the first sheet. Figure 9 shows a diagram of the setup. We viewed the image live, through the TV monitor. Before the camera position was adjusted, when we rotated the camera, the lines on the front sheet would appear to move either 'faster' or 'slower' than the lines on the more distant sheet. By adjusting the camera position we could make it so that no relative motion could be detected. We then moved the camera 0.5 mm from the optimum position in both directions and the relative motion of the lines when the camera was rotated could be seen distinctly. From this we conclude that, with care, we could position the center of projection less than 0.5 mm from the axis of rotation. Figure 10 shows the images before and after rotation. It is hard to see the slight misalignment due to a rotation around a point 0.5 mm from the center of projection (Figure 10c). On the TV monitor it is quite clear. Figure 11 shows the join between the lines in Figure 10 in enlarged detail.

5.3 Features and feature detector

The rotation method does not require any special calibration object but it does require that features can be detected reliably in the image. We wished to have features that could be located with subpixel accuracy with the assumption that we could always reduce the accuracy of our measurements at a later stage. The exact details of the features and feature detector are not a critical part of this thesis and are presented here only for completeness.

There are a few improvements which would make the calibration process more automatic. The most important improvements are to use a feature detector which can operate in a natural environment and to automatically track the features.

In order to simplify the feature extraction process we used a black and white

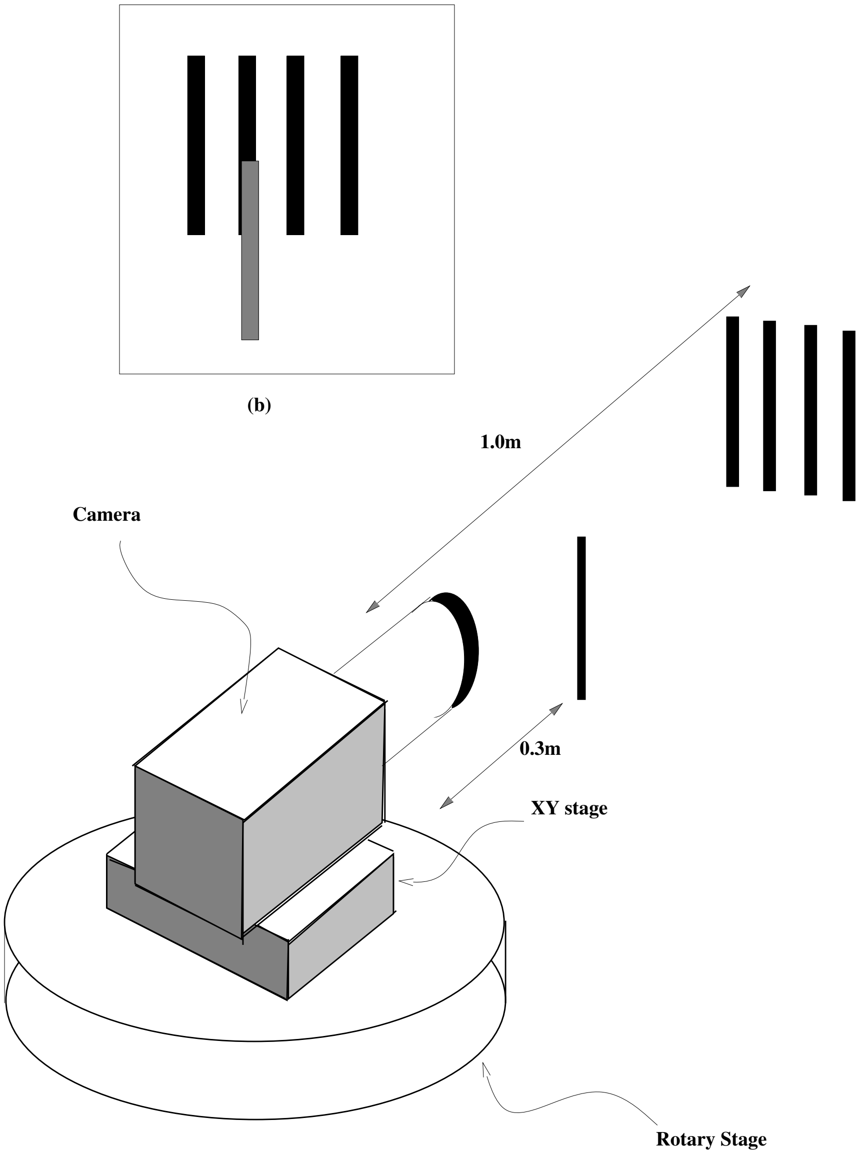
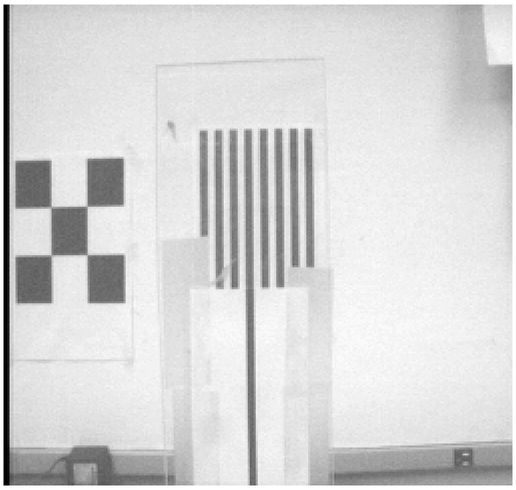


Figure 9: Digram of the setup used for testing for pure rotation. The camera is mounted on an XY stage that is mounted on a rotary stage. A set of black bars is positioned 1.0m from the camera. A single black bar is positioned about 0.3m from the camera. It's horizontal position is adjusted so that the close bar is aligned with one of the distant bars. The image is viewed on the TV monitor (b).



(a)

(b)

(c)

Figure 10: The image before rotation (a). The image after rotation when the camera rotation is pure rotation (b) and not pure rotation (c).

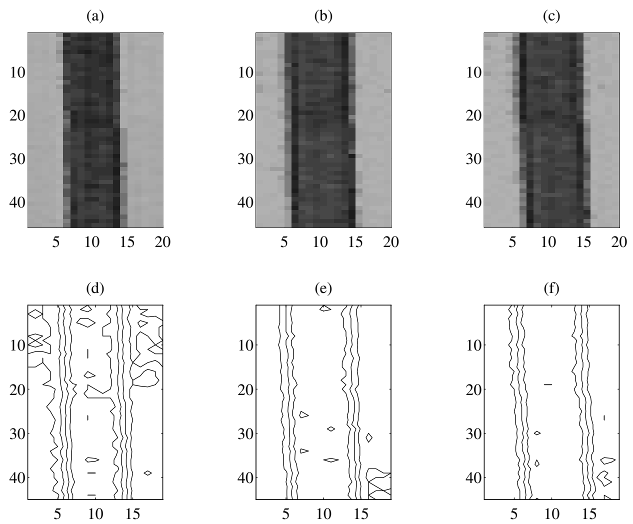


Figure 11: Enlarged detail of the image and contour plot of the grayscale values. The lines join at row 21 in all three images. One can see that in the image before rotation (a)(d) and in the image after rotation when the camera rotation is pure rotation (b)(e) the two lines are aligned. When the rotation is off by $0.5mm$ from pure rotation (c)(f) the nearer line (the lower half of the line) is slightly to the right.

checkerboard pattern printed on a single sheet of 8×11 inch paper as our calibration object. The features were defined as the corners of the black squares. A typical image is shown in figure 12. The calibration object occupies only a small section of the image and that area was identified and given to the feature extraction program by a human operator.

The program first uses a Canny edge detector to locate edge points in the image to subpixel accuracy and then uses a Hough transform to cluster edge points into lines. The program then finds the best straight line for each of the clusters which have a large number of points. It then finds the intersections and outputs the location of these intersections. One problem with this method is that there is an assumption that the lines should be straight but this is not a correct assumption due to radial distortion. This is not a major problem since the lines do not extend over large areas of the image and are therefore approximately straight. An improvement on this method would be to use a higher order curve for approximating the edge rather than a straight line.

5.4 Nonlinear optimization code

The camera parameters were found using a nonlinear optimization program based on the subroutine LMDIF from the software package MINPACK-1 [14]. This subroutine uses the array of residuals (e_i) to adjust the parameters to minimize the sum square error $\sum e_i^2$. This is a more powerful method than only using the sum itself. It calculates derivatives using forward differencing.

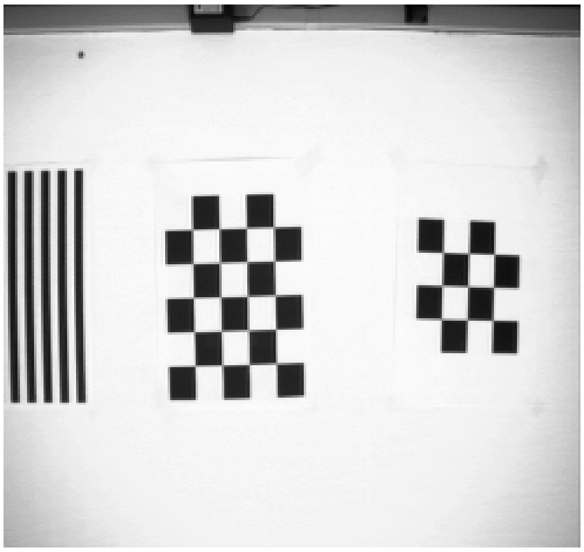


Figure 12: An example of one of the calibration images. The image was taken by the Sanyo 8.5mm camera. The left checker board could provide 20 feature points that are all at the corners of two black squares.

Each of the camera parameters could either be fixed to a constant value or could be allowed to be optimized. With none of the camera parameters fixed the program typically takes about 6 iterations to reach within 5 decimal places of the end result and 12 iterations to terminate. For 10 image pairs and 20 feature points per image this took 11 seconds on a SPARCstation ELC. No attempt was made to optimize the code.

5.5 Repeatability of angle measurement

The rotation stage has a vernier scale and can measure angles to an accuracy of $\pm 1'$. ($1' = \frac{1}{60}^o$). We were interested in testing to what accuracy one could actually position the camera. An image was taken at 225^o. The camera was rotated and then returned to the same measured angle of 225^o and a second image was taken. We used the Sanyo Camera with an 8.5 mm lens ($f = 630$ pixels) and the rotation was around the Y (vertical) axis.

The locations of 20 features were extracted from each of the images and compared. Let dx_i be the difference in the x coordinate of the i'th feature point between the first image and the second image and dy_i the difference in the y coordinate. The mean dx over the 20 feature points was 0.1775 pixels with a standard deviation of 0.0165 pixels and the mean dy was 0.0445 pixels with a standard deviation of 0.0051 pixels.

We used the camera parameters found in following sections but the nominal camera parameters could have been used just as well.

The angular error $\delta\theta$ in units of minutes arc can be calculated from the horizontal error d_x:

$$\delta\theta = \frac{60\arctan\left(Sd_x/f\right)180}{\pi} = \frac{60\arctan\left(1.259 * 0.1775/630\right)180}{\pi} = 1.22' \qquad (39)$$

The repeatability of angular measurements was between $\pm 1.22'$

5.6 The calibration method described step by step

1. The camera was mounted on a small XY stage that was mounted on top of a rotary stage. The rotary stage was securely fastened to an optical bench. The XY stage was adjusted so that the axis of rotation passed through the origin using the method described in section 5.2. The camera was mounted upright so that the rotation was around the Y axis.

2. A calibration object with easily detectable features, such as a checker board pattern was fastened securely to the wall approximately 1 m in front of the camera.

3. A set of 9 images was taken with the camera rotated at various angles. The angles were more or less evenly spaced over the whole range of angles in which the calibration object could be seen. The angle at which each image was taken was measured and recorded.

4. The feature detector was used to extract the locations of the features from each of the images. Depending on the lens and camera used, either 16 or 20 features were in the calibration object (see figure 12). Because of the spatial arrangement of the features, the feature detector always detected the features in the same order so that finding feature correspondence between images was automatic.

5. Five or six pairs of images were randomly selected from the set of nine images and the angle of rotation for each pair was calculated from the measured angles.

6. The camera was mounted on its side so as to enable rotation around the X axis and steps 1 through 6 were repeated to produce a second set of image pairs.

7. A combined data file was created from all the pairs of images using both axes of rotation.

8. The nonlinear optimization program was used to find the camera parameters which best accounted for the motion of the feature points.

6 Calibration by rotating a complete circle.

6.1 Theory

If the angles of rotation can be measured accurately, then the technique described above can yield accurate estimates of the camera parameters, especially for the focal length. If the angle measurement has a zero mean random error then one can increase the accuracy by using a large number of image pairs. If, on the other hand, there are systematic errors in the angle measurements, then there will be a proportional error in the focal length estimate and the other estimates will also be off by various amounts.

This problem can be eliminated entirely by using a sequence of images taken as the camera rotates through 360 degrees. Within a large range around the correct angle and focal length, the focal length estimate is a monotonically decreasing function of the angle measurement. This means that given an estimate of the angle, one can obtain an estimate of the focal length and vice versa, given a guess at the focal length, one can obtain an estimate of the angle of rotation. While this can be shown to be true if all the other camera parameters are held constant it is not so clear that this holds when we obtain the other camera parameters, each time, by optimization as in the previous section. Nevertheless, experimental results show this to be true.

Let us assume that one has a sequence of n images taken as the camera rotates through a whole circle and the i^{th} image has some area of overlap with the $i+1^{st}$ image and that there are some detectable features in each of those areas of overlap. Let us then take the first image in the sequence to be also the last. Each of the i^{th} and $i+1^{st}$ images can be used as an image pair. For a given guess at the focal length we can obtain an estimate for the angles between the two images in each of the image pairs. We then sum up these angles. If the focal length was guessed correctly then the sum should be 360 degrees. Since each of the angles, and hence the sum of the angles, is a monotonic function of our guess of f it is simple to find the correct f iteratively using simple numerical methods such as bisection or Brent's method [16].

Figure 13 shows the angles obtained from a real image sequence as a function of the focal length. The camera parameters do not stay constant as we vary our guess of f. As an example Figure 13d shows the values of the radial distortion parameters K_1 and K_2 obtained using various values of focal length.

6.2 Experimental details

The Chinon camera was mounted on the rotary stage as in section 5.2. The camera was positioned so that the axis of rotation passed as near as possible to the center of projection. A sequence of overlapping images was taken as the camera was rotated to the left in a full circle. Features in the left of the i^{th} image corresponded to features in the right of the $i+1^{st}$ image. Figure 14 shows a typical image pair.

Ideally a set of calibration patterns would be set up in a ring around the camera so that each pair of adjacent patterns could be seen in the camera's field of view simultaneously. For practical reasons we found it easier to move the patterns in a leap frog manner. The pattern used in the first image was attached to the wall and not moved. It was then featured in the last image as well.

Although the camera used had a wide angle lens and a rotation of 60 degrees would still enable overlap between the images, the camera was not always rotated as much as it could be. If we had done so it would be hard to distinguish between the effects of focal length and radial distortion. So we initially made a few relatively

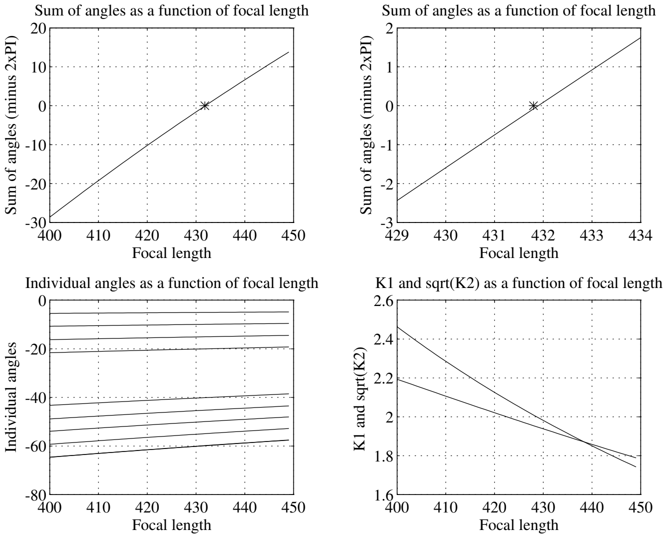


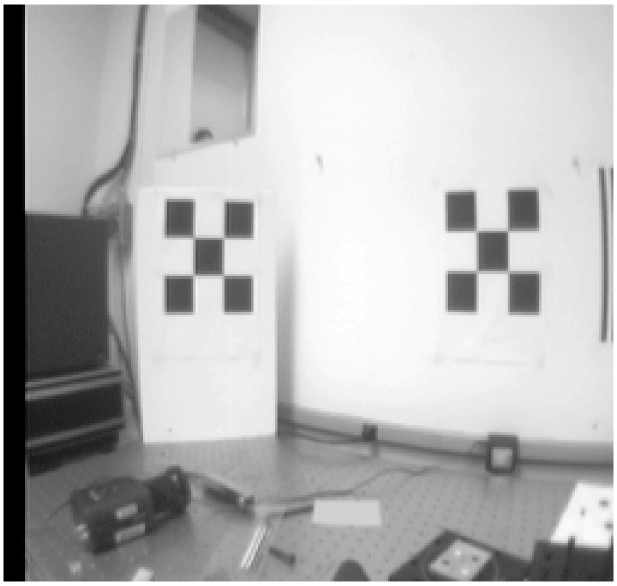
Figure 13: The sum of the angles and the individual angles of rotation between the images in the sequence as a function of the focal length use for the calculation. An enlarged view of the area around the correct focal length (upper right). The values of the radial distortion parameters K_1, K_2 also vary monotonically with the focal length (lower right).

small rotations (5-20 degrees) and we then used large rotations to complete the circle. The images were taken at the following angles: 225, 215, 210, 195, 155, 110, 50, 350, 295, 245, 225, - the same image was used as first and last.

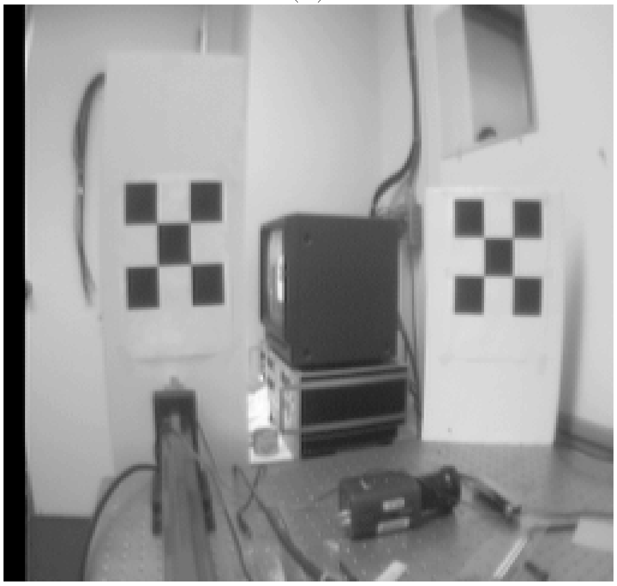
7 Rotation method results and comparison with results obtained using geometric objects

Four experiments were performed with the Sanyo camera and the 8.5mm lens. For each experiment two set of images were used, one with horizontal rotation and one with vertical rotation. Each set was made up of nine images taken at various angles of rotation. From each set we selected five pairs of images. We used these sets of images to calibrate the camera using the measured angles. We then performed a fifth, similar, experiment with a 12.5mm lens. Between each set of images the camera position was adjusted so that the axis of rotation passed through the center of projection.

With the Chinon camera we performed two experiments using measured angles. We then performed a third experiment (#6 in table 7) where the camera was rotated a full circle and we took an overlapping set of images. This set was used for the full circle rotation method described in section 6. For this third experiment we used only a single set with horizontal rotation. Since this meant that the aspect ratio could not be recovered reliably, we used the average aspect ratio obtained from the first two experiments. The results of the rotation experiments are shown in table 7. For



(a)



(b)

Figure 14: Typical image pair for the full circle rotation method. Image (a) was taken at 195^o and image (b) at 155^o.

Table 7: Calibration results - using rotation

	Sanyo Camera									
#	$f_{[mm]}$	$f_{[pixels]}$	cx	cy	cxr	cyr	k1	k2	sf	RMS error
1	8.5	632.4	251.2	243.4	254.1	234.9	5.23e-07	-9.08e-13	1.2583	0.278858
2	8.5	632.4	248.4	245.9	252.6	238.3	5.57e-07	-1.03e-12	1.2576	0.354125
3	8.5	631.8	254.1	240.3	254.8	243.4	5.19e-07	-7.92e-13	1.2567	0.248890
4	8.5	631.3	243.5	248.4	250.1	234.0	5.59e-07	-1.39e-12	1.2552	0.197770
	$Mean^*$	631.95	249.3	244.5	252.9	239.2	5.3958e-07	-1.0303e-12	1.2569	
	σ/mean	0.0008	0.0180	0.0142	0.0083	0.0148	0.0403	-0.2511	0.0011	
5	12.5	928.4	244.3	249.9	242.4	233.7	2.239e-07	5.188e-13	1.25741	0.120
	Chinon Camera									
#		f	cx	cy	cxr	cyr	k1	k2	sf	RMS error
6		433.5	254.0	225.5	254.2	223.2	1.91e-06	5.41e-12	1.24664	0.43
7		433.2	249.3	225.3	248.9	221.5	1.88e-06	5.43e-12	1.24721	0.40
8		431.9	256.3	225.3	254.7	225.7	1.923e-06	3.836e-12	(1.247)	0.922

* Mean of experiments 1 though 4.

Table 8: Calibration results - using geometric shapes

#	$f_{[mm]}$	$f_{[pixels]}$	cx	cy	cxr	cyr	k1	k2	sf	
	Sanyo Camera									
1	8.5	664.3	245.1	247.2	244.9	245.2	6.07e-07	-1.73e-12	1.2586	
	Chinon Camera									
2	3	n/a	n/a	n/a	257.2	229.8	1.80e-6	6.44e-12	1.247	

comparison, the results obtained using geometric shapes are summarized in table 8.

We have no ground truth values for the various camera parameters but we can check to see if the results are repeatable. For the Sanyo 8.5 mm the standard deviation of the focal length is 0.5 pixels or 0.08%. The standard deviation in aspect ratio, S is 0.11%. When a 12.5mm lens is used the aspect ratio is the same, as one would hope. The mean for the rotation experiments is 1.2569 which agrees well with the 1.2586 obtained using the spheres. The standard deviation in focal length of the chinon camera between experiments 6,7 and 8 is 0.85 pixels or about 0.17% of the focal length. The difference between the aspect ratio obtained in experiment 6 and 7 is 0.05% and the mean is 1.2469 which is the same as obtained using spheres.

In order to see the significance of these values we compare the error in measuring the angle to a point in space due to the variance in the focal length and principal point and the error in measuring the angle due to an error in feature location. Let us assume that the correct principal point is at $(0, 0)$. Let us assume the point P_1 projects to the principal point $(0, 0)$ and that point P_2 projects to a point $(x, 0)$. The angle between the rays to P_1 and P_2 is:

$$\alpha = \arctan \frac{x}{f} \tag{40}$$

If we have an error in feature location of Δx then we will measure the angle as:

$$\alpha_2 = \arctan \frac{x + \Delta x}{f} \tag{41}$$

If we have an error in the focal length of Δf then we will measure the angle as:

$$\alpha_3 = \arctan \frac{x}{f + \Delta f} \tag{42}$$

If we have an error in the location of the principal point of ΔC_x then we will measure the angle as:

$$\alpha_4 = \arctan \frac{x - \Delta C_x}{f} - \arctan \frac{-\Delta C_x}{f} \tag{43}$$

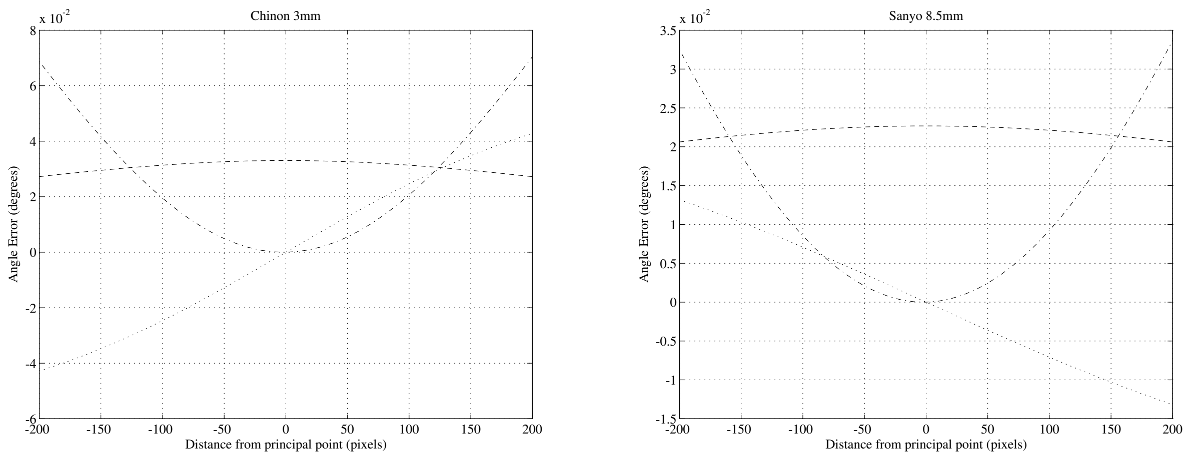
Figure 15(a) and Figure 15(b) show the error in angle measurements due to an error of 0.25 pixels in feature location and compare it to the measurement error due to an error of Δf or of ΔC_x in the focal length or principal point respectively, where Δf was taken as 1 standard deviation of the focal length: 0.85 pixels for the Chinon camera and 0.08% for the Sanyo camera. ΔC_x was taken as 1 standard deviation of the principal point: 4 pixels for the Sanyo camera and 3 pixels for the Chinon camera. The error in angle measurement due to the errors in focal length and principal point is of about the same magnitude as the measurement error due to a feature location error of 0.25 pixels.

The variance in the values obtained for the radial distortion parameters K_1 and K_2 using the rotation method is large but the increase of the value of K_1 is compensated for by the decrease in the value of K_2. Figure 16(a) shows the correction for radial distortion, (δ_x), for the 8.5mm lens as a function of the radial distance from the principal point using the radial distortion parameter values obtained with the plumb line method and with the rotation method. Since we don't know the correct answer, the δ_x computed using the parameters obtained from the first rotation experiment is used as ground truth and subtracted from the results obtained using the plumb line method and the other rotation experiments. The results are plotted in Figure 16b. Figure 16c and Figure 16d show similar results for the Chinon camera.

One can see in Figure 16 that the radial distortion coefficients K_1 and K_2 obtained from the rotation and plumb line experiments give the same correction for radial distortion with the difference being less than 0.5 pixels for $r < 200$. Using only one term for radial distortion, with the coefficient K_1 obtained by the plumb line method, the correction differs by over 2 pixels for the Chinon camera. Thus it is significantly different and shows that two coefficients must be used. The graphs diverge for $r > 200$ because the rotation method, as implemented, uses features located along the axes and therefore there is only data for r less than about 200. This can be corrected by using features near the corners of the images as well.

8 Discussion

The rotation method of calibration effectively calibrates the internal parameters of the camera. Using the calibrated camera one can calculate the direction of a ray



(a) (b)

Legend:	
- - -	Error of 0.25 pixels in feature location
.....	Error of $1\sigma_f$ in focal length (0.2% Chinon and 0.08% Sanyo)
_._._	Error of $1\sigma_{Cx}$ in principal point (3 pixels Chinon and 4 pixels Sanyo)

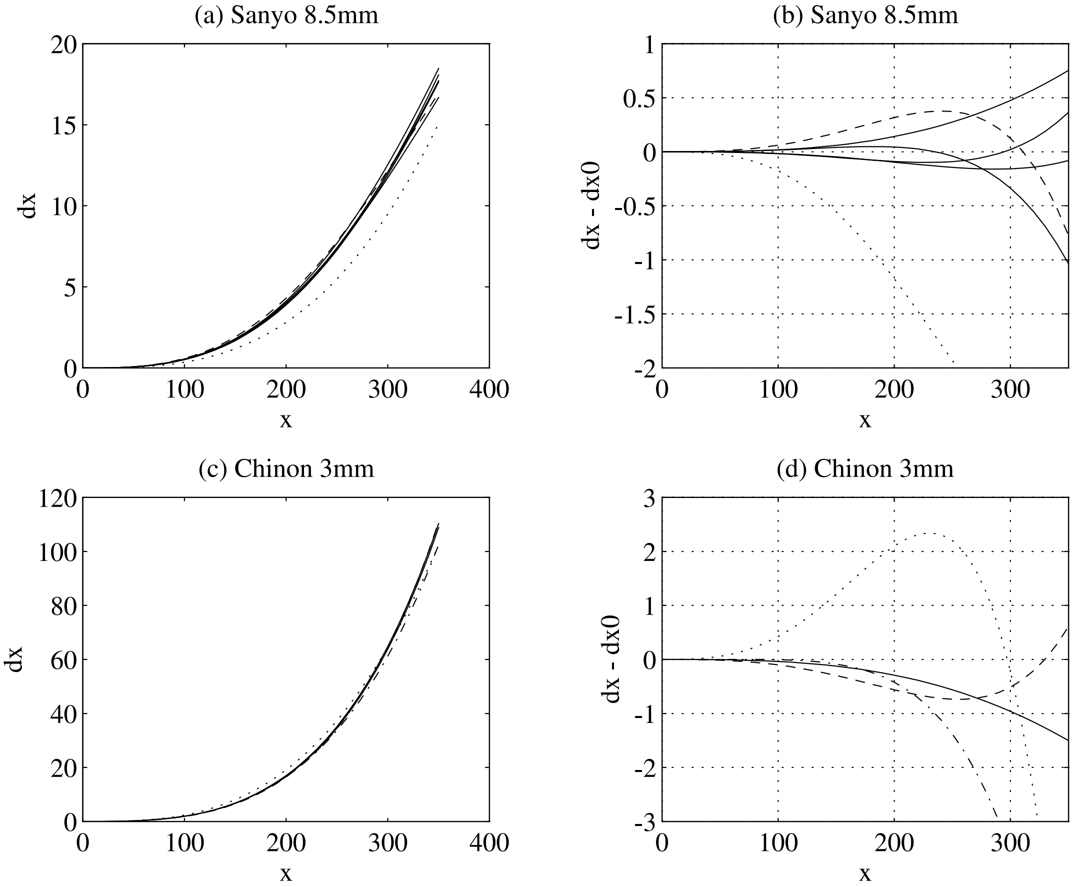
Figure 15: The error in measuring angles between a point imaged at $(0,0)$ and a point imaged at $(x,0)$ due to an error in feature location, focal length or principal point for (a) the Chinon camera or (b) the Sanyo camera.

to a point in space given the coordinates of the image. This direction is given in the camera coordinate system. The calibration method was designed specifically to provide internal camera calibration rather than both internal and external calibration parameters for two reasons. The first is that for our future work in motion vision we are interested only in the internal camera parameters; specifically in the ability to accurately compute the image after camera rotation. The second is that by avoiding the need to calibrate external parameters we avoid the problems involved in decoupling the internal and external parameters.

The rotation method is simple to use because it does not require known world coordinates of control points nor does it require very high precision feature extraction and it can be automated easily. Thus, it is suitable for autonomous robots working in unstructured environments or as a quick method for calibrating a camera in the laboratory.

The new method for finding the principal point and focal length using the images of spheres gives a value for the principal point which is in the range of values obtained in the rotation experiments. Repeated experiments (not reported here) show that the principal point can be recovered with a repeatability of ± 2 pixels. This is better than that given by the rotation method. The focal length estimate has too large a variance to be useful. It is possible that with improved lighting conditions and better image processing techniques more accurate estimates of the focal length can be found.

This method fits into our effort to avoid needing to know or find out the external



Legend:

—–	K_1 and K_2 from the different rotation experiments.
_._._	K_1 and K_2 from full circle rotation (Chinon only).
- - -	K_1 and K_2 from plumb line method.
.....	K_1 from plumb line method with $K_2 = 0$.

Figure 16: (a) The correction for radial distortion δ_x plotted against the distance from the principal point for the Sanyo 8.5mm lens. (b) δ_x derived using the average rotation experiment parameters subtracted from the rest of the results in (a). (c) The correction for radial distortion δ_x plotted against the distance from the principal point for the Chinon 3mm lens. (d) δ_x from the first rotation experiment subtracted from the rest of the results in (c).

camera parameters, such as the location of the calibration object. It uses a simple calibration object and can use a large number of images which can be combined in an easy manner to give a least squares estimate of the principal point. A major disadvantage is that it it requires knowing other internal camera parameters, such as the radial distortion parameters, and errors in these parameters can cause errors in the estimates of the focal length and principal point.

A Appendix: Review of calibration methods

A.1 Methods that use known world points

The standard methods of camera calibration for machine vision use sets of points with known position in some world coordinate frame and their corresponding image coordinates. These 3D points are called control points. The typical statement of the calibration problem is the following: given a set of control points with known world coordinates (X_i, Y_i, Z_i) and their location in the image (x_i, y_i) find the external and internal camera parameters that will best map the control points to their image points.

The calibration techniques used in machine vision are usually referred to as *Analytical Methods* in the photogrammetric literature [17] and are used for 'on the job' calibration, as opposed to the calibration methods used in camera and lens manufacture. The standard approach solves the nonlinear projection equations by iterative techniques. Since, as we have argued in section 1.2.1, errors could occur in the estimation of the internal parameters due to the coupling between internal and external parameters, it is a good idea to use multiple views of the calibration object from different camera positions. This is especially a good idea if one wishes to determine the internal parameters accurately. This can be performed using *Simultaneous Multi-Frame Analytical Calibration* [17].

Initially, the machine vision literature did not take into account lens distortion. The imaging process can then be described using homogeneous matrices[18]. Given a point $P = (X, Y, Z, 1)$ in homogeneous form the imaging process can be described as post multiplying P by a matrix $\mathbf{M}$:

$$(sx, sy, sz, s) = P\mathbf{M} \tag{44}$$

and the image coordinates (x, y) can be recovered by dividing the first and second term by the fourth. Strat[18] shows that the how to compute matrix $\mathbf{M}$ knowing the internal and external camera parameters. Sutherland[19] shows how to obtain the matrix $\mathbf{M}$ in a least squares way using the known world coordinates and their corresponding image coordinates.

The more difficult problem is obtaining the individual camera parameters from the matrix $\mathbf{M}$. This involves solving nonlinear equations. Ganapathy[7] shows a non-iterative method for their solution and a more geometrically intuitive method is described in [18].

Lenz and Tsai[21][11] and Weng et al.[22] provide methods for determining the camera parameters in the presence of lens distortion. Lenz and Tsai[21] assume only one parameter of radial distortion and no decentering distortion. First they assume the principal point is at the center of the image buffer and find the external parameters and focal length using the *radial alignment constraint*[21] and then find the radial distortion parameter, K_1. Then, using these parameters, they compute an error term by projecting the known world coordinates onto the image plane. In the next stage they use this error term as a cost function for a nonlinear search for the principal point coordinates.

Weng et al.[22] include decentering lens distortion in their camera model. They solve the problem of calibration by dividing the parameters into two groups. The first group includes the external parameters and the projection parameters. The second group includes the lens distortion parameters which are initially set to zero. The procedure is first to find the parameters in the first group and then using those parameters find the lens distortion parameters. Then using the new lens distortion parameters they recompute the parameters in the first group. This procedure is repeated till convergence.

Both the techniques of Lenz and Tsai[21][11] and the techniques of Weng et al.[22] result in camera parameters which enable very accurate 3D measurements. Lenz and Tsai[21][11] do not evaluate the accuracy of the individual camera parameters. The results in [22] show that in simulation the errors in focal length are above 0.5%.

A.2 Methods that use geometrical properties of objects

There are a variety of methods that use the vanishing points of parallel lines to determine the external and internal camera parameters. As shown in [9, pages 54-58] [3] the perspective projection of a set of parallel lines that are not parallel to the image plane is a set of lines that appear to meet at one point on the image plane. This point is called the vanishing point. The classic example of this phenomenon is the parallel tracks of a railway line which appear to meet at some point far off in the distance. The reader is pointed to [9] [3] for a formal discussion of the subject but a few properties of vanishing points are noted [3]:

1. The vanishing points associated with the sets of lines that are all parallel to a given plane, all lie on the same line in the image. This is called the vanishing line.

2. Given the vanishing points of three sets of mutually orthogonal lines, the orthocenter of the triangle with vertices in the three vanishing points is the principal point.

3. Given the vanishing points of two sets of orthogonal lines (x_1, y_1) and (x_2, y_2) it is shown that:
$$x_1 x_2 + y_1 y_1 + f^2 = 0 \tag{45}$$

4. If the camera moves, the motion of the vanishing points in the image plane depends only on the camera rotation not the camera translation. The vanishing

points of three non coplanar sets of parallel lines fully determine the rotation matrix.

Caprile and Torre [3] use these facts to determine the principal point and focal length of the camera and also to determine the camera rotation in an image sequence. The image aspect ratio must be determined by some other means. No effort was made to deal with radial distortion. The experiment was repeated an unspecified number of times and the principal point found varied over a range of 2 or 3 pixels. The focal length varied over a range of 0.5% and 0.3% for 8mm and 16mm lenses respectively. The calibration object used was a cube with parallel lines drawn on it's faces.

Beardsley [1] shows that if a plane is rotated around an axis which is not perpendicular to the plane, then the vanishing point produced by a set of parallel lines on the plane, draws a conic section in the image plane. Suitable positioning of the plane and the axis of rotation will produce an ellipse. The major axis of the ellipse passes through the principal point and the eccentricity of the ellipse depends on the position of the center of the ellipse and the focal length. Given a set of ellipses it is possible to find the aspect ratio, principal point and focal length. Experimental results show that the focal length can be found within 2%.

A.3 Methods that do not require known calibration points

It would clearly be useful to be able calibrate the camera using the image itself without the need of a special calibration object, in other words, to calibrate the camera using the locations of features in the image without knowing the 3D locations of those points in the world.

If the internal camera parameters are known we can find the relative position of two cameras in a stereo pair up to a scale factor in the translation. Following [9, pages 144-147] given the feature locations of a set of image points in the two cameras (x_{1i}, y_{1i}) and (x_{2i}, y_{2i}) the following system of equations must be solved:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}^t \left[\mathbf{R} \begin{pmatrix} x_{1i} \\ y_{1i} \\ f_1 \end{pmatrix} \times \begin{pmatrix} x_{2i} \\ y_{2i} \\ f_2 \end{pmatrix} \right] = 0, \; i =, \ldots, N \tag{46}$$

where (X, Y, Z) and $\mathbf{R}$ are the relative translation and rotation respectively between the camera 1 and camera 2.

Equation (46) can be solved iteratively. Longuet-Higgens[12] provides a non-iterative solution to this problem. If the focal lengths of the cameras are unknown but all the other internal camera parameters are known then Hartley[10] shows that the focal lengths and the relative position of the cameras can be found using feature point locations alone. It is also shown [10] that no more internal camera parameters can be deduced from the set of two images.

Faugeras et al. [4] show that given a motion sequence of 4 images where the motion is unknown and unconstrained it is possible in theory to recover the internal camera parameters assuming a perfect pinhole model. The method currently suffers very badly from noise in feature location. Simulation results with 0.01 pixel error give focal length errors of over 9%. This makes it impractical for real use.

References

[1] **Beardley, P., et al**, "Camera Calibration Using Multiple Images" In *Proceedings of the Second European Conference on Computer Vision*, 312-320, Santa Margherita Ligure, Italy, May (1992)

[2] **Brown, D.C.**,"Close -range Camera Calibration" *Photogrammetric Engineering* **37** 855-866 (1971)

[3] **Caprile, B. and Torre, V.**, "Using Vanishing Points for Camera Calibration" *International Journal of Computer Vision***4**, 127-140 (1990)

[4] **Faugeras,O.D., et al.**, "Camera Self-Calibration : Theory and Experiments" In *Proceedings of the Second European Conference on Computer Vision*, 321-334, Santa Margherita Ligure, Italy, May (1992)

[5] **Faugeras,O.D.**, "What can be seen in three dimensions with an uncalibrated stereo rig" In *Proceedings of the Second European Conference on Computer Vision*, 563-578, Santa Margherita Ligure, Italy, May (1992)

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

[7] **Ganapathy,S.**, "Decomposition of transformation matrices for robot vision" *Pattern Recognition Letters***2**, 401-412 (1984)

[8] **Gennery, D.B.**, "Stereo-camera Calibration" in *Proc. Image Understanding Workshop*, 101-108, Nov. (1979)

[9] **Haralick, R.M. and Shapiro, L.G.**, *Computer and Robot Vision* Addison-Wesley Publishing Company, (1993)

[10] **Hartley, R.I.**, "Estimation of Relative Camera Camera Positions for Uncalibrated Cameras" In *Proceedings of the Second European Conference on Computer Vision*, 563-578, Santa Margherita Ligure, Italy, May (1992)

[11] **Lenz, R.K. and Tsai, R.Y.**, "Techniques for Calibration of the Scale Factor and Image Center for High accuracy 3-D Machine Vision Metrology" *IEEE Trans. Pattern Anal. Machine Intell.* **10**,713-720 (1988)

[12] **Longuet-Higgins, H.C.**, "A computer algorithm for reconstructing a scene from two projections" *Nature* **293**,133-135 (1981)

[13] **Mohr, R. and Arbogast, E.**, "It can be done without camera calibration" *Pattern Recognition Letters* **12**,39-43 (1991)

[14] **More, J.J., et al.**, "User Guide for Minpack-1" Argonne National Laboratory, Argonne, Illinois (1980)

[15] **Penna, M.A.** "Camera Calibration: A quick and Easy Way to Determine the Scale Factor" *IEEE Trans. Pattern Anal. Machine Intell.* **13**,1240-1245 (1991)

[16] **Press, W.H. et al.**, "Numerical Recipes in C" Cambridge University Press, (1988)

[17] **Slama, C.C. ed**, *Manual of Photogrammetry*,4th edition, American Society of Photogrammetry (1980).

[18] **Strat,T.M.**, "Recovering the Camera Parameters from a Transformation Matrix" in *Proc. DARPA Image Understanding Workshop*, 264-271, Oct. (1984)

[19] **Sutherland, I.E.**, "Three-Dimensional Data Input by Tablet" in *Proceedings of the IEEE*, **62**,453-458 (1974)

[20] **Thomas, G.B.**, "Calculus and Analytic Geometry" 4th Edition, Addison-Wesley, Reading, MA, (1969)

[21] **Tsai, R.Y.** "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses" *IEEE Journal of Robotics and Automation***RA-3** 323-344 (1987)

[22] **Weng, J et al.** "Camera Calibration with Distortion Models and Accuracy Evaluation" *IEEE Trans. Pattern Anal. Machine Intell.* **14**,965-980 (1992)