# Large Accurate Internal Camera Calibration using Rotation with Analysis of Sources of Error

G. P. Stein
AI Lab
MIT
Cambridge, MA, 02139

**Abstract**

*This paper describes a simple and accurate method for internal camera calibration based on tracking image features through a sequence of images while the camera undergoes pure rotation. A special calibration object is not required and the method can therefore be used both for laboratory calibration and for self calibration in autonomous robots working in unstructured environments.*

*Experimental results with real images show that focal length and aspect ratio can be found to within 0.15 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 this paper we perform a simple analysis to show to what extent do the various technical details affect the accuracy of the results. We show that having pure rotation is important if the features are derived from objects close to the camera. In the basic method accurate angle measurement is important. 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.*

*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.*

## 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. In practical terms this often means finding the parameters of the perspective projection model, focal length (or more precisely, principal distance) and the principal point. It also requires finding the parameters of a lens distortion model and finding the ratio of the spatial sampling of the image plane in the X and Y directions. 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.

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. Some knowledge can be obtained using uncalibrated cameras [7] [14], 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 internal and external camera parameters. The intermediate parameters can often be found quite easily [19] [9] and can be used in this form for some tasks. Finding the internal parameters from the intermediate parameters explicitly is a difficult task [19].

Work on motion vision and pose estimation often use the perspective projection model and assume that the internal camera parameters are known [13]. 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 and related work

We provide here a brief review of the various approaches to camera calibration. More extensive reviews of calibration methods appear in [12][18] [20]. Since the cameras used in machine vision are not designed to highly accurate specifications nor to be highly stable they might have to be calibrated frequently. Therefore, camera calibration must be kept simple and as quick as possible.

#### 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 [12] [23]. Outdoors, control points could be markings on the ground [18] or buildings whose positions can be verified from maps [19]. The typical statement of the calibration 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.

There is a problem that arises due to the interaction between the external and the internal parameters. The Manual of Photogrammetry [18] 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". The main problem is the error in the focal length. In [23] simulation results show that errors of about 0.5 percent can occur with even small amounts of additive Gaussian noise. We can expect larger errors when dealing with real cameras. 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 [23]. The problem manifests itself when one tries to use subsets of the camera parameters for other tasks. Taking multiple views of the calibration object from different camera locations can help solve this problem [18].

### 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.

In [4] the vanishing points of parallel lines drawn on the faces of a cube are used to locate the focal length and principal point. In [2] 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. Results using the plumb line method [3], which uses the images of straight lines for calibrating lens distortion, a method for finding the aspect ratio using the image of a sphere [16], are given in section 7 for comparison to the rotation method. More details on these experiments can be found in [20].

### 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. The rotation method developed here fits into this category.

Faugeras et al. [6] develop a method where a motion sequence of a 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 finding internal camera parameters. Gennery [10] also uses unconstrained motion. No information is given as to the accuracy of the results.

Methods that use camera rotation are presented in [1] [5] [11]. Basu [1] tracks the motion of contour lines in the image as the camera pans or tilts through small angles. No lens distortion model is used. The experimental results show that the focal length and scale factor can be found within 3 percent error. No experimental results were given for the principal point.

In [5] first the principal point is found and then using the principal point the focal length and scale factor are found. The principal point is found by tracing the motion of feature points as the camera rotates. It is shown than these traces are curves whose axes of symmetry pass through the principal point. Simulation results show that given a feature detector with accuracy of 0.2 pixels the principal point can be found to within a few pixels the focal length can be found to within 1 percent. The accuracy of the experimental results are not shown.

The work shown in [5] is the most closely related to this paper and a comparison is in order. The method presented in [5] uses three separate stages to recover the camera parameters. In the rotation method presented here, all the parameters are found in one step and is therefore more straightforward. The method for finding the focal length presented in [5] uses small angle approximations to predict the motion of points in the image while for signal to noise reasons, large rotations are required to make accurate measurements. In this paper the full 3D rotation equations are used which are accurate for large and small angles. The method for finding the radial distortion in [5] requires being able to rotate the camera accurately in two orthogonal directions and that these be parallel to the XY axes of the image plane. This is a rather complex mechanical setup. In the method presented here, one is only required to be able to rotate about the center of projection.

Recent work by Hartley [11] has shown that given a perfect perspective projection model it is possible in theory to recover both the camera parameters and the angle of rotation. This has been shown to work nicely in simulation but no accuracy results have been shown with real images.The problem with this method is that there is a strong coupling between the estimate of the focal length and the angle of rotation. For example, an increase in the estimate of the focal length can be compensated for by a decrease in the estimate of the angles. In practice we found that we could not achieve better than 4 percent accuracy in focal length estimates using the iterative method presented in [11].

### 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. We take a set of images where the camera has rotated around a fixed axis. We can then find camera parameters that best predict the motion of the feature points in the set of images using 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.

The next section describes the camera model and notation used in this paper. It then goes on to describe in detail the effects of camera rotation and translation on an image. Section 4 delves into the experimental details. Subsection 4.1 gives information about the cameras and the image acquisition hardware. Section 5 shows how, by rotating the camera through a full circle, one can avoid the need to measure the rotation angles. Section 6 shows how the various sources of error affect the results. 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 model with 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).

The standard model for lens distortion [18] is a mapping from the distorted image coordinates, $(x_d, y_d)$, that are observable, to the undistorted image plane coordinates, $(x_u, y_u)$, which are not physically

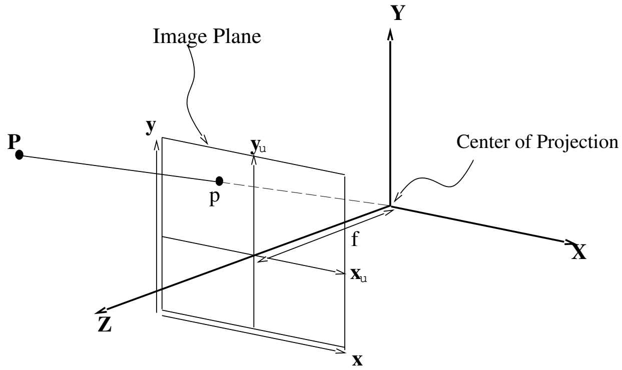



Figure 1: The camera coordinate system $(X, Y, Z)$, the undistorted image coordinates $(x_u, y_u)$ and the frame buffer coordinates $(x, y)$

measurable. The undistorted image plane coordinates are:

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

where:

$$r_d'^2 = x_d'^2 + y_d'^2 = (x_d^2 - c_{xr})^2 + (y_d^2 - c_{yr})^2 \tag{3}$$

It has been shown in [20] that allowing the center of radial distortion, $(c_{xr}, c_{yr})$ to be different from the principal point is equivalent to adding a term for decentering distortion as given in [18]. Finally, $\vec{P}$ is converted to frame buffer coordinates:

$$\begin{aligned} x &= \frac{1}{S} x_d + c_x \\ y &= y_d + c_y \end{aligned} \tag{4}$$

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. One can compute $S$ using the equation:

$$S = \frac{d_x}{d_y} \frac{f_s}{f_p} \tag{5}$$

where $d_y$ and $d_x$ are the center to center pixel spacing in the CCD array, $f_p$ is the camera pixel clock frequency and $f_s$ is the frame grabber sampling frequency. If any of these values are unknown, as in our case or are known to vary as in [12] then $S$ must be found by calibration. In the future we can expect digital cameras to become widely used and the scale factor will be known exactly.

### 2.2 Pure rotation using Rodriguez's formula

Using the camera coordinate system, let us say that the world rotates through an angle of $(\theta)$ degrees around the axis $\vec{W} = (w_x, w_y, w_z)$. 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:

$$\vec{P'} = \mathbf{R}\vec{P} = \left[\cos\theta\mathbf{I} + \sin\theta\mathbf{Q} + (1-\cos\theta)\hat{W}\hat{W}^t\right]\vec{P} \tag{6}$$

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

$$\mathbf{Q} = \begin{bmatrix} 0 & -w_z & w_y \\ w_z & 0 & -w_x \\ -w_y & w_x & 0 \end{bmatrix} \tag{7}$$

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

$$\begin{aligned} x' &= f\frac{X'}{Z'} = f\frac{(r_{11}X + r_{12}Y + r_{13}Z)}{(r_{31}X + r_{32}Y + r_{33}Z)} \\ y' &= f\frac{Y'}{Z'} = f\frac{(r_{21}X + r_{22}Y + r_{23}Z)}{(r_{31}X + r_{32}Y + r_{33}Z)} \end{aligned} \tag{8}$$

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 :

$$\begin{aligned} x' &= f\frac{(r_{11}x + r_{12}y + r_{13}f)}{(r_{31}x + r_{32}y + r_{33}f)} \\ y' &= f\frac{(r_{21}x + r_{22}y + r_{23}f)}{(r_{31}x + r_{32}y + r_{33}f)} \end{aligned} \tag{9}$$

We can conclude from (9) the well known fact 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.

### 2.3 Rotation with translation

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. In that case the location of an image point after rotation and translation will be:

$$\begin{aligned} x' &= f\frac{(r_{11}x + r_{12}y + r_{13}f + ft_x/Z)}{(r_{31}x + r_{32}y + r_{33}f + ft_z/Z)} \\ y' &= f\frac{(r_{21}x + r_{22}y + r_{23}f + ft_y/Z)}{(r_{31}x + r_{32}y + r_{33}f + ft_z/Z)} \end{aligned} \tag{10}$$

The location of the point is no longer independent of the depth and depends also on the translation vector. This issue is discussed further in section 6.2.

## 3 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 of rotation are known one can compute, using (9), 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 and the rotation, 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 (4):

$$\begin{aligned} x^l_{d_{ij}} &= S(x^l_{ij} - c_x) \\ y^l_{d_{ij}} &= (y^l_{ij} - c_y) \end{aligned} \tag{11}$$

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 (2) to compute the feature coordinates in the undistorted image plane:

$$\begin{aligned} x^l_{u_{ij}} &= x^l_{d_{ij}} + \delta x^l_{d_{ij}} \\ y^l_{u_{ij}} &= y^l_{d_{ij}} + \delta y^l_{d_{ij}} \end{aligned} \tag{12}$$

We can now define a cost function:

$$E = \sum_{i=1}^{N}\sum_{j=1}^{M}\left((x^{1'}_{u_{ij}} - x^2_{u_{ij}})^2 + (y^{1'}_{u_{ij}} - y^2_{u_{ij}})^2\right) \tag{13}$$

where $(x^{1'}_{u_{ij}}, y^{1'}_{u_{ij}})$ are the coordinates of $(x^1_{u_{ij}}, y^1_{u_{ij}})$ rotated using (9).

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 (13). 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. We expect that in the future, with the use of digital cameras, the aspect ratio will be known and then only one set of images is required. In any case, 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 (13).

# 4 Experimental details

## 4.1 Hardware

Three B/W CCD cameras with standard video output were used for the experiments. A Sanyo VDC3860 High Resolution Camera with 8.5 mm and 12.5 mm Comiscar lenses. 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. A Pulnix WV-CD 50 high resolution camera with a 16 mm lens. The images were acquired using a Datacube frame-grabber into $512 \times 485$ 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 $\pm 1'$. ($1' = \frac{1}{60}^o$). In practice we found that the repeatability of angular measurements was between $\pm 1.22'$ [20].

## 4.2 How to obtain pure rotation ?

As shown in section 2.3 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 micrometer adjustment of the XY stage is a bit of an over kill.

The test whether there is little or no translation is quite simple and involves the use of parallax. This test was performed experimentally as follows. The rotation was around the vertical axis. We positioned a set of vertical black lines, printed with a laser printer and mounted on a flat board, 1.0 m from the camera. We then took a single, vertical black line, about one third the thickness of the previous ones and mounted on a transparent acrylic sheet. This was positioned 0.3 m from the camera in front of the first page so that the single line seemed to be a perfect continuation of one of the lines in the first set. Figure 2 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. For more details see [20].

## 4.3 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 sub-pixel accuracy

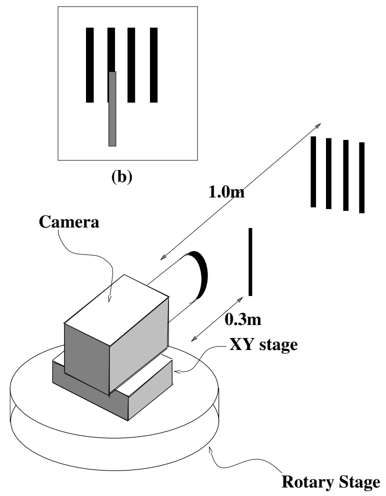



Figure 2: Digram of the setup used for testing for pure rotation. The camera is mounted on an XY stage that is mounted on a rotary stage. A set of black bars is positioned 1.0m from the camera. A single black bar is positioned about 0.3m from the camera. It's horizontal position is adjusted so that the close bar is aligned with one of the distant bars. The image is viewed on the TV monitor (b).

with the assumption that we could always reduce the accuracy of our measurements at a later stage. The details of the features and feature detector are not critical and are presented fully in [20].

In order to simplify the feature extraction process we used a black and white checkerboard pattern printed on a single sheet of $8 \times 11$ inch paper as our calibration object. The features were defined as the corners of the black squares. A typical pair of images is shown in figure 4.

## 4.4 Nonlinear optimization code

The camera parameters were found using a nonlinear optimization program based on the subroutine LMDIF from the software package MINPACK-1 [15]. This subroutine uses a modified Levenberg-Marquart algorithm adjust the parameters so as to minimize the sum square error $\sum e_i^2$. It calculates derivatives using forward differencing.

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 SPARC-station ELC. No attempt was made to optimize the code.

## 5 Calibration by rotating a complete circle.

### 5.1 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.

One can avoid the need to measure angles at all and thus 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 $i^{\text{th}}$ 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 $i^{\text{th}}$ 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 [17].

Figure 3 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 3d shows the values of the radial distortion parameters $K_1$ and $K_2$ obtained using various values of focal length. We pick the radial distortion parameters that correspond to the focal length which gives the correct sum of angles.

### 5.2 Experimental details

The Chinon camera was mounted on the rotary stage as in section 4.2. 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 $i^{\text{th}}$ image corresponded to features in the right of the $i+1^{st}$ image. Figure 4 shows a typical image pair.

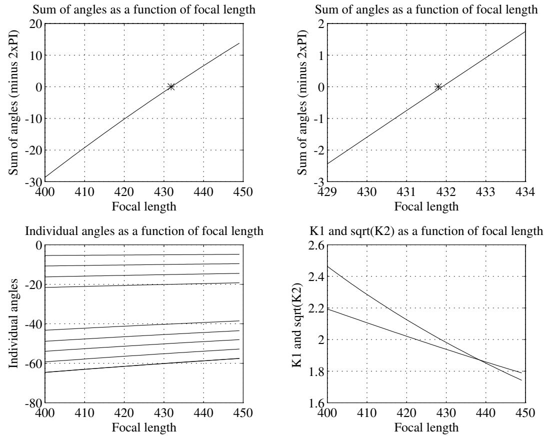
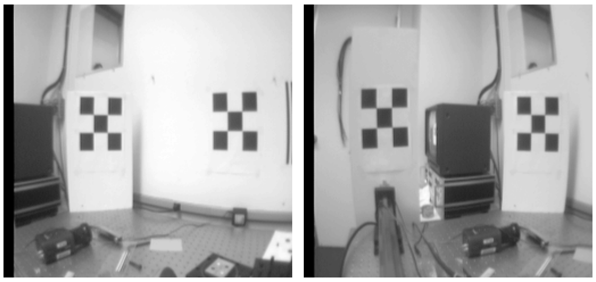



Figure 3: The sum of the angles and the individual angles of rotation between the images in the sequence as a function of the focal length use for the calculation. An enlarged view of the area around the correct focal length (upper right). The values of the radial distortion parameters $K_1$, $K_2$ also vary monotonically with the focal length (lower right).

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) (b)

Figure 4: Typical image pair for the full circle rotation method. Image (a) was taken at $195^o$ and image (b) at $155^o$. Only The checker board pattern in the left of image (a) appears in image (b) on the right hand side.

a few relatively 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.

# 6 Analysis of error sources

In this section we will discuss the possible sources of error and show what effect they might have on the results. The sources of error that come to mind are: errors in feature location, errors caused by not having pure rotation, errors in angle measurement, errors due to not including a radial distortion model and using a small angle approximation to model the rotation. We will not present a formal error analysis but will give a good idea of how each source of error might affect the results.

## 6.1 Errors in feature detection

If we have an error of 1 pixel in the location of a feature point and suppose the rotation caused a motion of 100 pixels we would expect an error of 1 percent in focal length estimates. But since we use a large number of points, at least 20 per image pair and since the typical errors have zero mean, the final error is much smaller.

## 6.2 Errors caused by not having pure rotation

Let us consider the simple case were the rotation is around the (vertical) Y axis and that the center of projection is at a distance D directly behind the axis of rotation. We assume that the Z axis intersects the axis of rotation.
In this case the equation for $x'$ from (10) becomes:

$$x' = f\frac{(x + \sin(\theta) * f + ft_x/Z)}{(f - \sin(\theta) * x + ft_z/Z)} \tag{14}$$

Where $t_x = D*\sin(\theta)$ and $t_z = D*(1-\cos(\theta))$. Using the small angle approximation we get $t_x \simeq D*\theta$ and $t_z \simeq 0$ and:

$$x' \simeq f\frac{(x + \theta * f + f * D * \theta/Z)}{(f - \theta * x)} \tag{15}$$

Since $f \gg \theta * x$ we can approximate the above equation to:

$$x' \simeq f\frac{(x + \theta * f + f * D * \theta/Z)}{f} \tag{16}$$

Extracting $f$ we have:

$$f \simeq \frac{x' - x}{\theta(1 + \frac{D}{Z})} \tag{17}$$

From the above equation we see that for a none zero value of $D$ we will get a error in the estimate of $f$. In fact the error will be $\frac{D}{Z}$ of the true value. This is a systematic error which cannot be reduced by using many feature points or image pairs unless we change the location of the axis of rotation each time. In our case we could ensure that $D < 0.5mm$ and since $Z \simeq 1.0m$ the error produced was less than 0.05 percent.

## 6.3 Errors in angle measurement

If we used a single image pair, an error of 1 percent in measurement of the angle would cause an error of 1 percent in the focal length estimate. Since the angle measurements errors are typically of the zero mean Gaussian type taking many measurements can help reduce the error. The standard deviation in this case goes down as the square root of the number of image pairs. In our case the angle measurement was estimated to be around $\frac{1}{50}$ of a degree and with rotation angles of around 20 degrees for the 16 mm lens this would cause an error of 0.1 percent for a single image pair.

In the case of longer focal lengths the possible angle of rotation is smaller and the proportional error in angle measurement is therefore larger. If we use the method of rotating a full circle there is no angle measurement error. Even if we cannot rotate a full circle due to some mechanical constraint we could take an image sequence where the camera rotates through a large angle which is easier to measure with accuracy.

## 6.4 Errors due to radial distortion

In the case of wide angle lenses there can be lens distortion of tens of pixels at the edges. Even in the case of the 16mm lens where the lens distortion was not visible at all the actual distortion at the edges of the image was around 3 pixels in an image of $512 \times 512$ pixels. Since this is a systematic error rather than a zero mean error it cannot be overcome by increasing the number of measurements and could, in this case, cause an error of 1 percent in the focal length estimate.

One could try to use points that are located only near the center of the image as in [5]. The effects of radial distortion will be much smaller but that will limit one to small angle rotations and loss in angle measurement accuracy. Since we will only be using the center part of the image the signal to noise ratio of the feature detector will also be reduced.

## 6.5 Use of small angle approximations

Both [5] and [1] use small angle approximations to reach closed form solutions for the camera parameters. These approximations, $\sin(\theta) \simeq \theta$ and $\cos(\theta) \simeq 1$, are only good for very small angles. For angles above 10 degrees the error in the cosine approximation is over 1.5 percent and the error in the sine approximation is over 0.5 percent. In the method presented here we use no approximations and therefore can use large angles of rotation without error.

# 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

Table 1: Calibration results - using pure rotation

| # | Sanyo Camera $f_{[mm]}$ | $f_{[pixels]}$ | cx | cy | cxr | c |
|---|---|---|---|---|---|---|
| 1 | 8.5 | 632.4 | 251.2 | 243.4 | 254.1 | 23 |
| 2 | 8.5 | 632.4 | 248.4 | 245.9 | 252.6 | 23 |
| 3 | 8.5 | 631.8 | 254.1 | 240.3 | 254.8 | 24 |
| 4 | 8.5 | 631.3 | 243.5 | 248.4 | 250.1 | 23 |
| | $Mean^*$ | 631.95 | 249.3 | 244.5 | 252.9 | 23 |
| | $\sigma$/mean | 0.0008 | 0.0180 | 0.0142 | 0.0083 | 0.0 |
| 5 | 12.5 | 928.4 | 244.3 | 249.9 | 242.4 | 23 |
| | Chinon Camera | | | | | |
| # | | f | cx | cy | cxr | c |
| 6 | | 433.5 | 254.0 | 225.5 | 254.2 | 22 |
| 7 | | 433.2 | 249.3 | 225.3 | 248.9 | 22 |
| 8 | | 431.9 | 256.3 | 225.3 | 254.7 | 22 |

* Mean of experiments 1 though 4.

Table 2: Calibration results - using geometric shapes

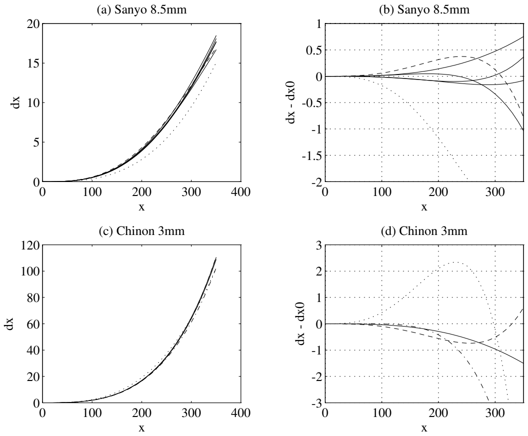
| $f_{[mm]}$ | cxr | cyr | k1 | k2 | sf |
|---|---|---|---|---|---|
| Sanyo | | | | | |
| 8.5 | 244.9 | 245.2 | 6.07e-07 | -1.73e-12 | 1.2586 |
| Chinon | | | | | |
| 3 | 257.2 | 229.8 | 1.80e-6 | 6.44e-12 | 1.247 |

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 (#8 in table 1) 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 5. 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 1. For comparison, the results obtained using geometric shapes are summarized in table 2.

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.

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 $K_1$




| Legend: | |
|---|---|
| —— | $K_1$ and $K_2$ from the different experiments. |
| -.-.-. | $K_1$ and $K_2$ from full circle rotation. |
| - - - | $K_1$ and $K_2$ from plumb line method. |
| ..... | $K_1$ from plumb line method with $K_2 = 0$. |

Figure 5: (a) The correction for radial distortion $\delta_x$ plotted against the distance from the principal point for the Sanyo 8.5mm lens. (b) $\delta_x$ derived using the average rotation experiment parameters subtracted from the rest of the results in (a). (c) The correction for radial distortion $\delta_x$ plotted against the distance from the principal point for the Chinon 3mm lens. (d) $\delta_x$ from the first rotation experiment subtracted from the rest of the results in (c).

is compensated for by the decrease in the value of $K_2$. Figure 5(a) shows the correction for radial distortion, ($\delta_x$), 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 m¿ethod and with the rotation method. Since we don't know the correct answer, the $\delta_x$ 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 5b. Figure 5c and Figure 5d show similar results for the Chinon camera.

One can see in Figure 5 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.

With the Pulnix camera we performed an experi-

Table 3: Calibration results - not using pure rotation. D is the distance of the axis of rotation from the center of projection. The feature patterns were located at a depths from 50cm to 75cm. Note that the RMS error, not just the error in estimated focal length increases with D.

| D | $f_{[pixels]}$ | cx | k1 | RMS error |
|---|---|---|---|---|
| 0 | 1260.2 | 283.958 | 1.18914e-07 | 1.551235 |
| 3mm | 1255.32 | 272.597 | 1.04814e-07 | 1.659607 |
| 1cm | 1243.36 | 255.994 | 1.0223e-07 | 3.470970 |
| 4cm | 1178.37 | 242.397 | 1.4864e-07 | 8.937260 |

ment to test the effect of moving the center of projection away from the rotation axis. We used rotation around the Y axis only and used the scale factor which we had found in another experiment not described in this report. We first positioned the center of projection on the axis of rotation and took a set of images. We then moved the center of projection $D = 3$mm away and took a second set of images. We repeated this for $D = 1cm$ and $D \simeq 4cm$. For each set of images we found the best camera parmeters. We were mainly concerned here with $f$. The objects used for feature points were located between 0.5m and 0.75m from the camera.

The results are shown in table 3. As we can see, when we increase the distance $D$ the error in focal length $f$ increases. This is a near linear function of $\delta f$ as a function of $D$. We also note that the RMS error increases as $D$ increases. This is because the depth disparities in the feature objects cause parallax which cannot be modeled by pure rotation. If all the feature points had come from a planar object parallel to the image plane we could have modelled the motion of the points as pure rotation and we would have no indication that something was wrong. The RMS error would be small but the focal length we got would be in error.

## 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 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. For motion vision we are usually interested only in the internal camera parameters. 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. It is suitable for autonomous robots working in unstructured environments or as a quick and accurate method for calibrating a camera in the laboratory. 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.

We have shown that with care one can use this method to calibrate the internal parameters to high accuracy. The main issue one must deal with with is ensuring that one uses pure rotation. More precisely, one must ensure that the deviation from pure rotation is small compared to the distance to the objects. In outdoor scenes, where the objects are far away, this is no problem but indoors one might get errors if one is not careful.

## References


[1] **Basu, A.**, "Active Calibration: Alternative Strategy and Analysis" In *Proceedings of IEEE Conference on Computer Vision and Pattern recognition*, 495-500, New York, NY, June (1993)

[2] **Beardley, P., et al**, "Camera Calibration Using Multiple Images" In *Proceedings of ECCV*, 312-320, Santa Margherita Ligure, Italy, May (1992)

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

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

[5] **Du, F. and Brady, M.**, "Self Calibration of the Intrinsic Parameters of Cameras for Active Vision Systems" In *Proceedings of IEEE Conference on Computer Vision and Pattern recognition*, 477-482, New York, NY, June (1993)

[6] **Faugeras,O.D., et al.**, "Camera Self-Calibration : Theory and Experiments" In *Proceedings of ECCV*, 321-334, Santa Margherita Ligure, Italy, May (1992)

[7] **Faugeras,O.D.**, "What can be seen in three dimensions with an uncalibrated stereo rig" In *Proceedings of ECCV*, 563-578, Italy, May (1992)

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

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

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

[11] **Hartley, R.I.**, "Self-Calibration from Multiple Views with a Rotating Camera" in *Proceedings of the European Conference on Computer Vision* (1994)

[12] **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)

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

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

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

[16] **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)

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

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

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

[20] **Stein, G.P.**, "Internal Camera Calibration using Rotation and Geometric Shapes" AITR-1426, Master's Thesis, Massachussets Institute of Technology, Artificial Intelligence Laboratory (1993).

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

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

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