# Camera Center Estimation

Manoj Aggarwal and Narendra Ahuja
University of Illinois at Urbana-Champaign, Beckman Institute
405 N. Mathews Ave, Urbana, IL 61801
E-mail: {manoj,ahuja}@vision.ai.uiuc.edu
Internet: http://vision.ai.uiuc.edu/

## Abstract

*A fast camera calibration technique to estimate the center of perspective projection of an imaging camera has been described. The proposed technique requires a single image of two planar calibration charts arranged in a special manner. The special arrangement of the calibration charts simplifies the projection equations relating the 3-D scene coordinates to the 2-D image coordinates and they can be suitably combined to eliminate the unknown intrinsic parameters other than the desired center. We have analyzed the error in the center estimate due to various alignment errors in the experimental setup, and shown that the scheme is quite robust.*

## 1. Introduction

Camera calibration is the process of determining the internal geometric and optical characteristics (intrinsic parameters) and/or the 3-D position and orientation of camera relative to a chosen world coordinate system (extrinsic parameters). The relationship between the 3-D scene and the image coordinates is essential for many computer vision applications such as active vision, scene mosaicing and depth estimation. Both the intrinsic and extrinsic calibration methods have been examined by several authors [1, 2, 4, 6, 7, 8, 9, 10, 11, 12].

In this paper, we focus on one aspect of intrinsic calibration. Specifically, we present a fast and accurate technique to estimate the optical center. There are at least 15 definitions for image center, which have been summarized in literature by Willson and Shafer [13]. These definitions include, center of radial lens distortion, center of field of view, center of perspective projection and center of expansion for focus and zoom. In this paper, we will be estimating the center of perspective projection.

Numerous approaches to camera calibration which include estimating the optical center have been reported. Tsai [8, 9] has summarized and evaluated some of these techniques which appeared before the 90's. These techniques were grouped by Tsai into three categories: Direct optical method, Method of varying focal length and the Radial alignment method. The direct optical method employs a collimated laser beam to accurately determine the center. In the method of varying focal length, the basic idea is to zoom the image of a scene by varying the effective focal length. There is only one point which remains stationary in all images and that is the desired image center. The focal length can be varied by either changing the distance setting of the focal ring, using two different focal length lenses or by using a zoom lens. All these methods were evaluated by Tsai and were found to be unreliable as the process of changing the effective focal length alters the lens center! The radial alignment method proposed by Tsai, exploits the constraint that the optical center, true image of a scene point and the ideal image of the same scene point, are collinear. This constraint holds because the distortion that occurs in most lenses is primarily radial. This technique, unlike others, takes into account lens distortion and has been shown to be quite accurate, though it uses non-linear optimization. A number of other techniques based on non-linear optimization having varying degrees of complexity and accuracy have also been proposed in literature[1, 2].

In recent years, new techniques which exploit projective constraints have been proposed [3, 6, 7, 11, 12]. They involve taking multiple images of the same calibration pattern from different viewpoints, or equivalently of multiple calibration patterns from the same viewpoint. These techniques are suitable for calibration of active cameras, where one doesn't have prior knowledge of the type and position of the calibration patterns, and are relatively complex and susceptible to numerical instabilities.

In this paper, we will use a two-plane calibration technique to estimate the optical center. We use a special arrangement of the two calibration planes which allows us to estimate the optical center without the need for multiple images or knowledge of any other intrinsic parameters. The arrangement of the calibration planes, the sensor and the lens is shown in Figs. 1 and 2. This arrangement can be

0-7695-0750-6/00 $10.00 © 2000 IEEE

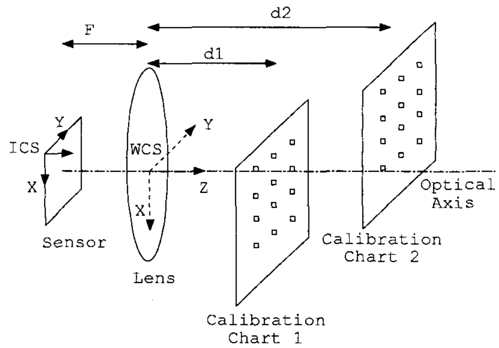


**Figure 1. Sensor, lens and calibration chart arrangement**

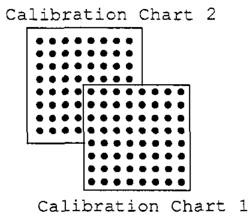


**Figure 2. Front view of the two calibration charts**

obtained by placing two calibration charts, each a regular grid of black dots, and placing them exactly one behind the other such that both of them are perpendicular to the optical axis. The initial position of the charts is chosen so that their image only covers approximately one-quarter of the sensor. The farther chart is then displaced vertically and horizontally by a multiple of the vertical and horizontal dot pitch of the calibration grid, respectively, such that the combined image of the two charts cover the entire sensor. This configuration ensures that the rows and columns of dots on the two charts are aligned as shown in Fig. 2. Alternatively, we can use a single calibration chart and take two images at two different depths of the chart from the camera. As we will show in Section 2, this arrangement of the calibration charts simplifies the projection equations relating the 3-D scene coordinates to the 2-D image coordinates and they can be suitably combined to eliminate the unknown intrinsic parameters other than the optical center. We have analyzed the error in the center estimate due to various alignment errors in the experimental setup, and shown that the scheme is quite robust to small errors in the experimental setup.

The organization of this paper is as follows. In Section 2, we give the camera model and derive the relationship between the optical center and the image coordinates of the two dot patterns. In Section 3, we analyze the sensitivity of the center estimate to errors in the alignment of the calibration charts. Section 4 presents concluding remarks.

## 2. Camera model

Consider the camera geometry shown in Fig. 1. The origin of the world coordinate system (WCS) is assumed to be located at the lens center, with the X and the Y axis spanning the lens plane. The Z axis is along the optical axis and points towards the object space. The image coordinate system (ICS) is parallel to the WCS but its origin is located at a corner of the active area of the sensor. Then the relationship between the 3-D coordinates of the dots on a planar chart to their images is given by

$$\begin{bmatrix} y_i^s \\ x_i^s \end{bmatrix} = \begin{bmatrix} -\frac{F}{d} & 0 \\ 0 & -\frac{F}{d} \end{bmatrix} \begin{bmatrix} y_i^o \\ x_i^o \end{bmatrix} + \begin{bmatrix} C_y \\ C_x \end{bmatrix}, \quad (1)$$

$\forall 1 \leq i < M$, where $(.)^s$ denotes the coordinates of the image of the dot pattern in ICS, $(.)^o$ are the coordinates of the dot pattern in WCS, $M$ is the number of dots, $F$ is the focal length, $d$ is distance of a calibration chart from the lens center and $(C_x, C_y)$ are the coordinates of the optical center in ICS.

This model assumes that radial distortion is negligible. Since the units used for measuring distance in images is pixels, we need scaling factors $S_x, S_y$ for the horizontal and vertical image coordinates, respectively. Thus the final projection equations are given by

$$\begin{bmatrix} y_i^s \\ x_i^s \end{bmatrix} = \begin{bmatrix} S_y & 0 \\ 0 & S_x \end{bmatrix} \begin{bmatrix} -\frac{F}{d} & 0 \\ 0 & -\frac{F}{d} \end{bmatrix} \begin{bmatrix} y_i^o \\ x_i^o \end{bmatrix} + \begin{bmatrix} C_y \\ C_x \end{bmatrix} \quad (2)$$

$\forall 1 \leq i < M$, which reduces to two sets of equations

$$y_i^s = -\frac{S_y F}{d} y_i^o + C_y \quad \text{and} \quad (3)$$

$$x_i^s = -\frac{S_x F}{d} x_i^o + C_x \quad \forall 1 \leq i < M. \quad (4)$$

Consider the experimental setup in Fig. 1 with two calibration planes aligned as in Fig. 2. Construct pairs of points, one from each calibration plane, such that their $y$-coordinate in WCS is the same. The $x$-coordinates of points in a pair, and the $y$-coordinate for different pairs of points could be different. Let us assume there are $N$ such pairs. The value of $N$ can be significantly larger then $M$, as for every point in one chart, there are multiple points in the second with the same $y$-coordinate. The $y$-coordinate ($y_i^o$) of these pairs of points, thus satisfy

$$y_i^{s_1} = -\frac{S_y F}{d_1} y_i^o + C_y \quad \text{and} \quad (5)$$

$$y_i^{s_2} = -\frac{S_y F}{d_2} y_i^o + C_y \quad \forall 1 \leq i < N, \quad (6)$$

where $s_1, s_2$ denote the two calibration charts and $d_1, d_2$ their distances from the lens center, respectively. The two sets of equations (5) and (6) can be combined to eliminate the unknowns $y_i^o$, $S_y$ and $F$ to obtain

$$y_i^{s_1} - s y_i^{s_2} = (1-s)C_y \quad \forall 1 \le i \le N, \tag{7}$$

where $s = \frac{d_2}{d_1}$. The value of $s$ can either be estimated by explicitly measuring $d_1$ and $d_2$ or it can be evaluated as follows. Using equations (3) and (4) we have

$$y_i^{s_1} - y_j^{s_1} = -\frac{S_yF}{d_1}y_i^o + \frac{S_yF}{d_1}y_j^o \quad \text{and} \tag{8}$$

$$y_i^{s_2} - y_j^{s_2} = -\frac{S_yF}{d_2}y_i^o + \frac{S_yF}{d_2}y_j^o, \tag{9}$$

$\forall i \neq j,\ 1 \le i,j \le M$. These two sets of equations can be combined to eliminate the unknowns $S_y, F, y_i^o, y_j^o$ to obtain

$$\frac{d_2}{d_1} = s = \frac{A^{s_1}}{A^{s_2}}, \quad \text{where,} \tag{10}$$

$$A^{s_k} = \sum_{(i,j)\in R} (y_i^{s_k} - y_j^{s_k}), \quad k = 1,2 \tag{11}$$

and $R = \{(i,j) : y_i^o > y_j^o\}$. The estimated value of $s$ can be substituted in the equations (7) to obtain a least squares estimate of $C_y$.

The estimate of $s$ is quite robust to any measurement errors in coordinates of the dots. In the expression for $A^{s_k}$, by choice of $R$, either all terms are positive or all are negative and the magnitude of each term is greater than the horizontal spacing between two successive dots in any row. Thus, the mean value of $A^{s_k}$ is significantly larger than the number of terms. If the error in each term of $A^{s_k}$ is assumed to be uniformly distributed, then the mean of the cumulative error in $A^{s_k}$ is zero and variance is small. We therefore, neglect the contribution of error in the estimate of $s$ to $C_y$.

To estimate $C_x$, we form pairs of points, one from each calibration plane such that their $x$-coordinate is the same, irrespective of their $y$-coordinate. The rest of the procedure is same as used for estimating $C_y$.

## 3. Sensitivity Analysis

The proposed algorithm assumes that calibration charts and the sensor are perfectly aligned. This condition can only be approximately met, in practice. In this section we will analyze the effect of systematic errors due to misalignment between the calibration charts and the sensor. The misalignment can be decomposed into three components. There is relative rotation between the calibration chart and the sensor, making up two of the components, one for each calibration chart. The third component, which we call translational error, is the vertical and horizontal displacement between the rows and columns of the two calibration charts. The first two subsections are devoted to the analysis of the sensitivity of center estimates to translational and rotational error, respectively. In the last subsection, we will present a discussion on the choice of position of the calibration charts so as to reduce the sensitivity of center estimates to misalignment.

### 3.1. Sensitivity to translational error

Let $\delta_y$ be the misalignment in the $y$-coordinate between the two charts. Then the projection equations (5) and (6) are modified to

$$y_i^{s_1} = -\frac{S_yF}{d_1}y_i^o + C_y \quad \text{and} \tag{12}$$

$$y_i^{s_2} = -\frac{S_yF}{d_2}(y_i^o + \delta_y) + C_y \quad \forall 1 \le i \le N \tag{13}$$

and equation (7) changes to

$$y_i^{s_1} - s y_i^{s_2} = (1-s)C_y + \hat{\delta_y} \quad \forall 1 \le i \le N, \tag{14}$$

where $\hat{\delta_y} = s\delta_y \frac{S_yF}{d_2}$. We note that $\hat{\delta_y}$ has the units of pixels. Thus the error $E_y$ in the estimate of $C_y$ is given by

$$E_y = \frac{\hat{\delta_y}}{1-s} = \delta_y S_y F \frac{s}{(1-s)d_2} = \frac{\delta_y S_y F}{d_2 - d_1}. \tag{15}$$

Considering that distances can be measured easily up to an accuracy of 0.1mm, for example with vernier calipers, $|\ \delta_y\ | \le$ 0.1mm. For a typical CCD sensor, pixel pitch is roughly 10 microns and thus $S_y \approx 100$ (i.e inverse of pixel pitch in mm). If we choose $F$ = 25mm, $d_1$ = 672mm and $d_2$ = 1008mm, then $E_y \le$ 0.8 pixel. The choice of $d_1$ = 672mm and $d_2$ = 1008mm is justified in Appendix A. Equation (15) suggests that in order to get the best estimates for the optical center for a given CCD sensor and a lens, one should minimize the ratio $\frac{s}{(1-s)d_2} = \frac{1}{d_2-d_1}$, i.e maximize the distance between the two calibration charts.

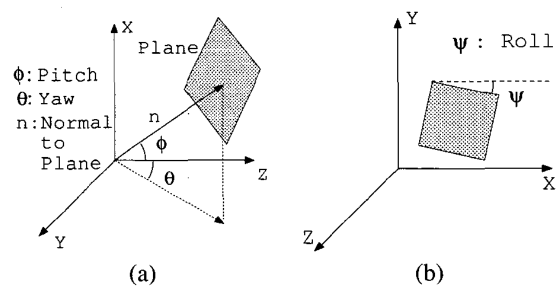


**Figure 3. The three components of the orientation of a plane. (a) Yaw and Pitch, (b) Roll**

### 3.2. Sensitivity to rotational error

Rotation between the sensor and calibration planes has three components: yaw ($\theta$), pitch ($\phi$) and roll ($\psi$), as shown in Figs. 3(a) and (b). For simplicity, we decouple the three rotations and analyze them independently. First, assume that all rotations, except yaw ($\theta$), are zero. Due to non-zero $\theta$ the scene coordinates $(x, y, z)$ of the dot pattern are altered to $(x, y\cos(\theta), z - y\sin(\theta)) \approx (x, y, z - y\theta)$, assuming small $\theta$. The change in image coordinates $(\delta x_i^s, \delta y_i^s)$ due to non-zero $\theta$ can be derived using (3) and (4) and are given by

$$\delta y_i^s = \frac{S_y F}{d^2}\theta(y_i^o)^2 \quad \text{and} \tag{16}$$

$$\delta x_i^s = \frac{S_y F}{d}\theta x_i^o y_i^o. \tag{17}$$

Assuming $\theta = 0.1^o = \pi/1800$ radians, which is the typical precision of rotation stages[1], $F = 25$mm, $S_y = 100$, and half-angle view of the camera = 10 degrees, we have $\frac{y_i^o}{d} < \tan(10) = 0.18$ and thus maximum projection error is approximately 0.15 pixels. According to equation (15) a projection error of 0.15 pixels contributes a maximum error of $|0.15/(1-s)| = 0.3$ pixels (for $d_1 = 672$mm, $d_2 = 1008$mm) to the estimation of the center.

Similarly, non-zero pitch ($\phi < 0.1^o$) contributes a maximum error of 0.3 pixels in the estimation of the center.

The projection equations, however, are more sensitive to any roll in the calibration charts, but fortunately roll can be very easily corrected. Since horizontal lines in the scene should project to horizontal lines in the image, by measuring the height of the dots in the image, we can estimate roll and reduce it such that the projection error it causes is less than half a pixel. For $d_1 = 672$ and $d_2 = 1008$mm, 0.5 pixel projection error causes $|0.5/(1-s)| = 1$ pixel error in the estimate of the center.

Thus the worst case error for a 25mm lens and $s = 1.5$ in the estimate of the center due to rotational error is $2(0.3 + 0.3 + 1) \approx 3.2$ pixel. This worst case error can be reduced further by increasing $s$.

## 4. Discussion on the sensitivity of center estimates

In previous subsection, we observed that we can improve the worst case behavior by making both $(d_2 - d_1)$ and $\frac{d_2}{d_1}$ simultaneously large. Both the terms can be made arbitrarily large, by choosing appropriate value of $d_1$ and $d_2$. However, the distances $d_2$ and $d_1$ are constrained by the depth of field of the camera. If the two calibration charts are put very far from each other, it is possible that image of one of the charts is blurred and thus measuring the image coordinates may become difficult. For a given camera configuration, the maximum and the minimum distance at which a planar object is well focussed is fixed. Thus, by placing the two calibration charts at these positions would maximize both $(d_2 - d_1)$ and $\frac{d_2}{d_1}$. The depth of field and thus the accuracy of the proposed technique can be increased by reducing the aperture size and increasing the ambient light levels. Also, if we can control the distance of the sensor from the lens, we can alter the minimum and maximum distance at which a planar object is well-focussed. Let $d_1$, $d_2$ be the minimum and the maximum distance at which an object is well focussed. It is shown in Appendix A, that for a given $F$, the ratio $s = \frac{d_2}{d_1}$ increases as $d_1$ increases. Thus, by altering the camera configuration, we can increase $d_1$ and as a result increase both $(d_2 - d_1)$ and $\frac{d_2}{d_1}$.

The total worst case error for a 25mm lens, with $d_1 = 672$mm and $d_2 = 1008$mm is $3.2 + 0.8 = 4.0$ pixels. Equation (15) shows that the worst case error is linear in $F$. This gives the impression that this technique might give poor estimate of the center for large focal lengths. This, however, is not the complete picture. The error due to translational misalignment is a function of $\frac{F}{d_2 - d_1}$. As shown in Appendix A, if $F$ is scaled by a factor of $\alpha$, then $d_1$ and $d_2$ can be suitably scaled, such that $\frac{F}{d_2 - d_1}$ and $s = \frac{d_2}{d_1}$ don't change, and the depth of field constraint is also met. Thus, by choosing appropriate values of $d_2$ and $d_1$, it can be ensured that the error due to translational misalignment doesn't increase. The error due to roll is independent of $F$ and depends only on $s$, which doesn't change. For the case of rotational misalignment (yaw and pitch), the error in center estimate is a function of $F\tan^2$(half-angle of view). For regular camera lenses, the angle of view scales by the same factor as $F$, so the value of $F\tan^2$(half-angle of view) in fact reduces as $F$ increases, thus reducing the worst case error. It is however possible that for lenses with large focal length (say 100mm), we may not be able to choose large value for $s$, because the corresponding values of $d_1$ and $d_2$ would be quite large. To illustrate this point assume that we have a lens with $F = 100$mm and we need $s = 1.5$, then as shown in Appendix A, these requirements allow $d_1 = 2687$mm and $d_2 = 4031$mm, but they are quite large. However, for $F = 100$mm and $s = 1.2$, Table 1 gives $d_1 = 1394$mm and $d_2 = 1673$mm, which are reasonable. As noted in Appendix A, we can alter $d_1$ and $d_2$ and hence their ratio, by changing the distance of the sensor from the lens. Under this experimental setup, the worst case error can be calculated to be less than 10 pixels. Roll contributes about 6 pixels error, translational misalignment contributes 3.6 pixels, while yaw and pitch contribute much less.

When techniques are available which can measure pixel coordinates even in the presence of defocus, accuracy of the center estimates can be significantly enhanced, as the values of $d_1, d_2$ will be no longer constrained.

[1] Available through Melles Griot Catalog, 1997-1998

## 5. Conclusions

We have proposed a fast technique to estimate the optical center of an imaging system. We have also presented the worst case analysis of the errors in the estimates of camera center due to errors in the experimental setup. The worst case error for the image center in an experimental setup designed using simple measuring tools like vernier calipers, is about 4 pixels for a 25mm lens and about 10 pixels for a 100mm lens. Since multiple pairs of points are used to estimate the center, the actual error would be much less than the worst case value and would be dominated by the contribution due to the translational misalignment. Thus a more realistic error estimate would be 1 pixel for a 25mm lens and 3.6 pixels for a 100mm lens. A error of less than 4 pixels in the image center is well within the tolerance limits of a number of computer vision tasks.

## References


[1] H. Bacakoglu. An optimized two-step camera calibration method. In *International Conference on Robotics and Automation*, pages 1347–1352, April 1997.

[2] J. Batista, H. Araujo, and A. T. de Almeida. Iterative multistep explicit camera calibration. In *International Conference on Computer Vision*, pages 709–714, 1998.

[3] O. D. Faugeras. Stratification of the three-dimensional vision: projective, affine, and metric representations. *Journal of Optical Society of America A.*, 12(3):465–484, March 1995.

[4] H. A. Martins, J. R. Birk, and R. B. Kelley. Camera models based on data from two calibration planes. *Computer Graphics and Image Processing*, 17:173–180, 1981.

[5] A. Pentland. A new sense of depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):523–531, July 1987.

[6] P. F. Sturm and S. J. Maybank. On plane-based camera calibration: A general algorithm, singularities, applications. In *Conference on Computer Vision and Pattern Recognition*, pages 432–437, 1999.

[7] B. Triggs. Autocalibration from planar scenes. In *European Conference on Computer Vision*, volume 1, pages 89–105, 1998.

[8] R. R. Tsai. An efficient and accurate camera calibration technique for 3d machine vision. In *Conference on Computer Vision and Pattern Recognition*, pages 364–374, 1986.

[9] R. Y. Tsai. 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(4):323–344, August 1987.

[10] L.-L. Wang and W.-H. Tsai. Camera calibration by vanishing lines for 3-d computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):370–376, April 1991.

[11] G.-Q. Wei and S. D. Ma. A complete two-plane camera calibration method and experimental comparisons. In *International Conference on Computer Vision*, pages 439–446, 1993.

[12] G.-Q. Wei and S. D. Ma. Implicit and explicit camera calibration: theory and experiments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(5):469–480, May 1994.

[13] R. G. Willson and S. A. Shafer. What is the center of the image? In *Conference on Computer Vision and Pattern Recognition*, pages 670–671, 1993.


## A. Depth of Field

The relationship between the distance $d$ of an object from the lens, the distance $v$ of the sensor from the lens, focal length $F$, f-number $f$, and the blur diameter $\sigma$ is given by [5]

$$d = \frac{Fv}{v - F - f\sigma} \tag{18}$$

Let $d_1$ and $d_2$ be the two solutions when the blur diameter is $-\sigma_0$ and $\sigma_0$, respectively. Then,

$$d_1 = \frac{Fv}{v - F - f\sigma_0} \quad \text{and} \tag{19}$$

$$d_2 = \frac{Fv}{v - F + f\sigma_0}. \tag{20}$$

Using these two equations we can eliminate $v$ and obtain an expression for $\frac{d_2}{d_1}$ in terms of the others, given by

$$\frac{d_2}{d_1} = \frac{F^2 - f\sigma_0 F}{F^2 + f\sigma_0 F - 2f\sigma_0 d_1} \tag{21}$$

This equation shows that as $d_1$ increases, the ratio $\frac{d_2}{d_1}$ also increases. If we choose $\sigma_0$ equal to size of a pixel then the distances $d_1$ and $d_2$ specify the depth of field of a sensor at a distance $v$ from the lens. The distances $d_1$ and $d_2$ can be altered, though not independently, by changing $v$. Table 1 shows some examples for values of $d_1$ and $d_2$ under different lens settings ($v$). We note that the size of the aperture has been kept fixed for all the examples in the table.

| F | f | $d_1$ | $d_2$ | $\frac{d_2}{d_1}$ |
|---|---|---|---|---|
| 25mm | 16 | 672mm | 1008mm | 1.5 |
| 100mm | 64 | 2687mm | 4031mm | 1.5 |
| 100mm | 64 | 1394mm | 1673mm | 1.2 |

**Table 1. Values of $d_1$ and $d_2$ under different lens settings**

Using equation (21) it can be shown that if we scale both $d_1$ and $F$ by a factor $\alpha > 1$, without changing the physical size of the aperture, then the f-number scales by a factor $\alpha$ and the ratio $\frac{d_2}{d_1}$ remains unchanged.