Estimation of Internal and External Parameters for Camera Calibration Using 1D Pattern

Xiangjian He[1,2], Huaifeng Zhang[1], Namho Hur[2], Jinwoong Kim[2], Qiang Wu[1] and Taeone Kim[2]
[1]Department of Computer Systems
University of Technology, Sydney
PO Box 123, Broadway 2007, Australia
{sean, hfzhang, wuq}@it.uts.edu.au
[2]Digital Broadcasting Research Division
Electronics Telecommunications Research Institute
161 Gajeong-dong, Yuseong-gu, Daejeon, 305-700, Korea
{namho,jwkim,kimm}@etri.re.kr

Abstract

Camera calibration is to estimate the intrinsic and extrinsic parameters of a camera. Most of object-based calibration methods used 3D or 2D pattern. A novel and more flexible 1D object-based calibration was introduced only a couple of years ago, but merely for estimation of intrinsic parameters. The estimation of extrinsic papers is essential when multiple cameras are involved for simultaneously taking images from different view angles and when the knowledge of relative locations between the cameras is required. Though it is relatively simple using 2D or 3D calibration pattern, the estimation of extrinsic parameters is not obvious using 1D pattern. In this paper, we will perform a 1D camera calibration involving both intrinsic and extrinsic parameters.

1. Introduction

Camera calibration is an important issue in computer vision especially for video surveillance, 3D reconstructtion, image correspondence, image registration, and so on. Calibration of cameras is especially crucial for applications that require accurate metric information such as depth and dimensional measurements from 2D images.

Camera calibration is to estimate parameters that are either internal to or external to a camera. The internal (or intrinsic) parameters determine how the image coordinates of a point are derived, given the spatial position of the point with respect to the camera [7]. The external (or extrinsic) parameters describe the geome-trical relation between the camera and the scene, or between different cameras.

With the increasing need for highly accurate measurement in computer vision, camera calibration has attracted research efforts in the computer vision community. The existing techniques for camera calibration can be classified into object-based calibration and self-calibration. Self-calibration techniques (see, for example, [3]) do not use any calibration object. Using self-calibration techniques, a large number of parameters need to be estimated. Hence, self-calibration creates mathematical complexity. The object-based calibration can be further classified into three categories according to the dimension of calibration objects. The calibration techniques can be summarized as follows [10].

3D object-based camera calibration (see, for example, [1]) is performed by observing a calibration object whose geometry in 3D space is known with very good precision. The calibration object usually consists of two or three planes orthogonal to each other. This approach requires an expensive calibration apparatus and an elaborate setup.

2D object-based camera calibration (See, for example, [4, 9]) proceeds by observing a planar pattern shown at various orientations. A 2D pattern for calibration is easier to be set up than a 3D pattern. However, in practice, when multiple cameras are involved for calibration, it is sometimes difficult for all cameras to simultaneously observe a number of points on either a 2D or 3D pattern.

1D object-based camera calibration was introduced in [10]. 1D calibration is achieved by observing a few pre-determined collinear points on a 1D calibration object. During the observation, the 1D object can be

Proceedings of the IEEE International Conference
on Video and Signal Based Surveillance (AVSS'06)
0-7695-2688-8/06 $20.00 © 2006 IEEE

freely moved except that at least one point on the object must be at a fixed location all the time. Zhang's approach requests at least six (6) observations of the 1D object for estimating only five (5) intrinsic parameters and his approach does not estimate any extrinsic parameters. However, the estimation of extrinsic parameters is crucial when dealing with multi-camera (or multi-view) calibration. Though it is relatively easier and simple using a 2D or 3D pattern, the estimation of extrinsic parameters is not obvious using a 1D pattern because there are only a very limited number of reference points with known coordinates on the 1D pattern.

In this paper, we will show that with five observations, both intrinsic and extrinsic parameters can be computed using a 1D pattern (object).

This paper is organized as follows. Section 2 introduces the setup method for 1D calibration object (pattern). Section 3 describes the procedure for camera calibration Section 4 demonstrates the experimental results. We conclude in Section 5.

2. Line-Shape Calibration Object

Let $[x \quad y \quad z]$ represent the coordinates of any visual point in a fixed reference system (called the world coordinate system), and $[u \quad v]$ be the image coordinates of the projection of $[x \quad y \quad z]$. Let $X = [x \quad y \quad z \quad 1]^T$ and $U = [u \quad v \quad 1]^T$ respectively. Then, we have the following distortion free model [8]

$$sU = PX, \tag{1}$$

where s is a scale factor and P is a 3×4 projection matrix. Matrix P has a structure given by [1]

$$P = K[R \quad T],$$

where K is a 3×3 upper triangular matrix known as the camera calibration matrix, $R = [r_{ij}]$ is a 3×3 rotation matrix and $T = [t_j]$ is a 3×1 translation vector for $i, j = 1, 2, 3$. The elements of K are referred to as intrinsic parameters, and the elements of R and T as extrinsic parameters. The K has the form of [10]

$$K = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix},$$

where α and β represent the scale factors in image u and v axes respectively, γ is the parameter describing the skew of the two image axes, and (u_0, v_0) is called the principal point at which the optical axis intersects the image plane.

Equations (1) is then equivalent to

$$sK^{-1}U = [R \quad T]X. \tag{2}$$

Equation (2) implies two independent equations for every independent observation of U and X [7].

2.1. Calibration Object

By adopting the idea for 1D object construction as shown in [10], a more general line-shape calibration is constructed using two straight line segments (or sticks) AC and CE as shown in Figure 1. Point C is fixed. Stick CE is fixed and stick AC is allowed to move freely around point C. Let us now examine how many unknowns and how many equations we may create from the five points, that are A, B, C, D and E on the calibration object given N observations.

Without loss of generality, we may assume that the fixed point, C, is the origin of the world coordinate system, i.e., $C = [0 \quad 0 \quad 0]$, and $E - C$ is on the world x-axis and is pointing to the positive x direction.

We need two parameters to indicate the orientation of stick AC. The positions of A, B, C, D and E including the orientation of stick CE can be computed given the values of these two parameters because the distances between A and C, between B and C, between C and D, and between C and E are all known.

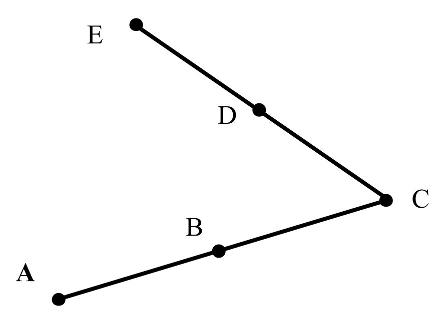

Figure 1. Line-Shape Calibration Object

Hence, given N observations of the stick, we have five intrinsic parameters that are $\alpha, \beta, \gamma, u_0,$ and $v_0,$ three independent parameters for matrix R, three parameters for vector T, $2N$ parameters for the locations of A, B, C, D and E to estimate. Therefore, the total number of unknowns is

$$11 + 2N.$$

As one will see in Section 3, the first observation of stick AC and stick CE are on the x-y plane of the world coordinate system. Hence, it needs only one parameter to determine the position of the first observation of AC. Hence, the total number of unknown is reduced to

$$10 + 2N. \tag{3}$$

The observation of A provides two equations as demonstrated in Equations (2). Each observation of C or E provides another two equations, and each observation of B or D provides one additional equation

Proceedings of the IEEE International Conference
on Video and Signal Based Surveillance (AVSS'06)
0-7695-2688-8/06 $20.00 © 2006 IEEE

because B is collinear with A and C, and D is collinear with C and E. Hence, the total number of independent equations that we can create is

$$5 + 3N. \tag{4}$$

When $N = 5$, both of Equations (3) and (4) give result of 20. Hence, if we have five or more observations, we should be able to solve all the calibration parameters without taking into account of camera distortion.

3. Camera Calibration

Referring to Figure 2, assume that the length of stick AC is the same as that of CE and equals to L, i.e.,

$$\|A - C\| = \|C - E\| = L. \tag{5}$$

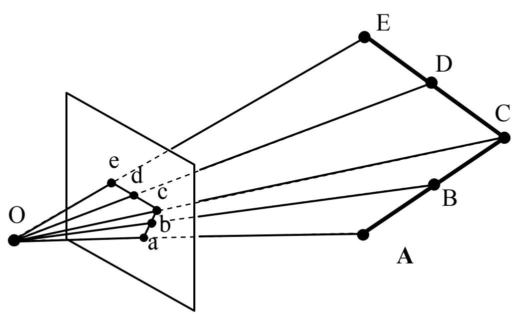

Figure 2. Projection of Calibration Object to a Plane

To simplify our discussion, without loss of generality, let us also assume that B is the midpoint of AC and D is the midpoint of CE. Therefore, we have that

$$B = \frac{1}{2}(A + C) \text{ and } D = \frac{1}{2}(C + E). \tag{6}$$

Let a, b, c, d and e be the projection points of A, B, C, D and E on the image plane respectively. For any given row vector (of any dimension) M, the following notations is used through this paper,

$$\overline{M} = [M \quad 1]^T \text{ and } s_M = [r_{31} \quad r_{32} \quad r_{33} \quad t_3]\overline{M}$$

when dimension of M is 3.

Hence, we have, for example,

$$\overline{a} = [a \quad 1]^T, \ \overline{A} = [A \quad 1]^T$$

and

$$s_A = [r_{31} \quad r_{32} \quad r_{33} \quad t_3]\overline{A}$$

for the given image point a and real word point A.

Similar but not identical to the derivation shown in [10], from Equation (2) above, we can easily obtain that

$$s_A K^{-1}\overline{a} = [R \quad T]\overline{A}, \tag{7}$$

$$s_C K^{-1}\overline{c} = [R \quad T]\overline{C}, \tag{8}$$

$$s_E K^{-1}\overline{e} = [R \quad T]\overline{E}, \tag{9}$$

$$s_A(\overline{a} \times \overline{b}) + s_C(\overline{c} \times \overline{b}) = \overline{0}, \tag{10}$$

$$s_E(\overline{d} \times \overline{e}) + s_C(\overline{d} \times \overline{c}) = \overline{0}, \tag{11}$$

where

$$\overline{0} = [0 \quad 0 \quad 0]^T.$$

3.1. Solving Intrinsic Parameters

The formulae for solving intrinsic parameters contained in K and s_C can be derived in a similar way to that of Zhang's as shown in [10]. Because of the page limitation, we do not present the derivation and list the formulae in this paper. However, it is worth to note that, unlike the work shown in [10], our work for estimation of intrinsic parameters does not assume the rotation matrix R to be identity matrix, and hence the derivation is not identical to that in [10].

3.2. Solving Extrinsic Parameters

The extrinsic parameters were not estimated in [10]. Though, with one camera calibration, the computation of the rotation matrix R can be avoided by setting the rotation to the identity as shown in [10], it is necessary and crucial for multiple camera calibration. In the following, we propose an algorithm for estimation of the extrinsic parameters. We describe how to solve the parameters for R and T respectively after K and s_C have been solved.

From now on, we use A_0, and B_0, to denote the first observed (or initial) coordinates of points A and B. Let us also denote the corresponding image points of A_0, and B_0 by a_0, and b_0.

Let $Z_0 = [0 \ \ 0 \ \ 1]$. Without loss of generality, let us assume that the positive direction of the world z-axis is in the direction of

$$(E - C) \times (A_0 - C),$$

i.e.,

$$Z_0 = \frac{(E - C) \times (A_0 - C)}{\|(E - C) \times (A_0 - C)\|}.$$

Note that R is a rotation matrix. Hence, from (7)-(9), it is easy to see that

$$R(E - C)^T = K^{-1}(s_E\overline{e} - s_C\overline{c}), \tag{12}$$

$$R(A_0 - C)^T = K^{-1}(s_{A_0}\overline{a_0} - s_C\overline{c}), \tag{13}$$

$$R[(E-C)\times(A_0-C)]^T = [K^{-1}(s_E\bar{e}-s_C\bar{c})]\times[K^{-1}(s_{A_0}\overline{a_0}-s_C\bar{c})]. \quad (14)$$

Let

$$\underset{-}{S}=[S_1 \quad S_2 \quad S_3], \quad (15)$$

where

$$S_1=K^{-1}(s_E\bar{e}-s_C\bar{c}),\ S_2=K^{-1}(s_{A_0}\overline{a_0}-s_C\bar{c}),$$

$$S_3=[K^{-1}(s_E\bar{e}-s_C\bar{c})]\times[K^{-1}(s_{A_0}\overline{a_0}-s_C\bar{c})].$$

Recall that C = [0 0 0]. From (9), we have

$$s_C K^{-1}\bar{c}=[R \quad T]\bar{C}=RC^T+T.$$

Hence,

$$T=s_C K^{-1}\bar{c}.$$

From (12)-(15), we have

$$\underset{-}{S}=R[E^T \quad A_0^{\ T} \quad \|E\times A_0\|Z_0^T]. \quad (16)$$

Equations (10) and (11) give

$$s_A=-s_C\frac{(\bar{c}\times\bar{b})\bullet(\bar{a}\times\bar{b})}{(\bar{a}\times\bar{b})\bullet(\bar{a}\times\bar{b})}. \quad (17)$$

$$s_E=-s_C\frac{(\bar{d}\times\bar{c})\bullet(\bar{d}\times\bar{e})}{(\bar{d}\times\bar{e})\bullet(\bar{d}\times\bar{e})}. \quad (18)$$

Let

$$f(a,b,c)=-\frac{(\bar{c}\times\bar{b})\bullet(\bar{a}\times\bar{b})}{(\bar{a}\times\bar{b})\bullet(\bar{a}\times\bar{b})}.$$

and

$$f(c,d,e)=-\frac{(\bar{d}\times\bar{c})\bullet(\bar{d}\times\bar{e})}{(\bar{d}\times\bar{e})\bullet(\bar{d}\times\bar{e})}.$$

Then we have

$$s_A=f(a,b,c)s_C \text{ and } s_E=f(c,d,e)s_C. \quad (19)$$

Let X_0 = [1 0 0]. Let us also assume that the positive direction of the world y-axis is in the direction of $Z_0\times X_0$ and Y_0 = [0 1 0].

Recall that the $E - C_0 = E$ is on the x-axis and its direction is the same as the direction of x-axis. Therefore,

$$E=[L \quad 0 \quad 0].$$

Let us denote the angle from E to A_0 by θ. Note that R^TR is a 3 × 3 identity matrix. From (16), we have that

$$E\bullet A_0=EA_0^{\ T}=ER^TRA_0^{\ T}=S_1^TS_2.$$

Noting that A_0 is on the x-y plane, we have that

$$\cos\theta=\frac{E\bullet A_0}{L^2},\sin\theta=\frac{\sqrt{L^2-\dfrac{(E\bullet A_0)^2}{L^2}}}{L},$$

$$A_0=(L\cos\theta)X_0+(L\sin\theta)Y_0,$$

and

$$\|E\times A_0\|=L^2\,|\sin\theta|.$$

Let

$$\underset{-}{A}=[E^T \quad A_0^T \quad \|E\times A_0^T\|Z_0^T].$$

Therefore, from (15) and (16), we have

$$R=\underset{-}{S}\,\underset{-}{A}^{-1}. \quad (20)$$

Recall that R is a rotation matrix and hence it must satisfy $R^TR = I$. Therefore, the R obtained from (20) must be adjusted before it is used for the subsequent processes. We follow the idea as shown in [9] for the adjustment of R. The adjustment is equivalent to finding a 3 × 3 matrix R such that

$$\|R-\underline{R}\|=\min_Q\|R-Q\|_F^2 \text{ and } \underline{R}^T\underline{R}=I,$$

where $\|\bullet\|_F$ stands for Frobenius norm [2] and Q is 3 × 3 matrix. When R is found, it is used to replace the R computed from (20).

Let the singular value decomposition of R be USV^T. Then, as shown in [9], we have

$$\underline{R}=UV^T.$$

3.3. Solving Coordinates of Other Reference Points

Point A except A_0 can be computed from (19) and then (7) as

$$A^T=R^{-1}(s_AK^{-1}\bar{a}-T).$$

From (6), we have

$$B=\frac{1}{2}A \text{ and } D=[\frac{L}{2} \quad 0 \quad 0]. \quad (21)$$

4. Experimental Results

The experiments are implemented using simulation data and also real data.

4.1. Experiments Using Simulation Data

The following are the settings for our implementation.

The internal parameter matrix K of camera, the rotation matrix R and translation vector T are set as follows such that the projected image points (computed using these settings) are distributed in a valid image region.

$$K = \begin{bmatrix} 1000 & 0 & 320 \\ 0 & 1000 & 240 \\ 0 & 0 & 1 \end{bmatrix},$$

$$R = \begin{bmatrix} 0.9354 & -0.0048 & 0.3536 \\ 0.2706 & 0.6533 & -0.7071 \\ -0.2276 & 0.7571 & 0.6124 \end{bmatrix}, T = \begin{bmatrix} -10 \\ 0 \\ 100 \end{bmatrix}.$$

The image resolution is set to be 640*480.

Note that R here is set through setting three angles in a spherical coordinate system.

Gaussian noise with 0 mean and σ standard deviation is added to the projected image points. The estimated camera parameters are compared with the ground truth and we measure their relative errors with respect to the focal length α, as proposed by Trigg in [5].

We vary the noise level from 0 to 1 pixel. For each noise level, we perform 200 independent random trials. Figure 3 shows the average of the relative errors.

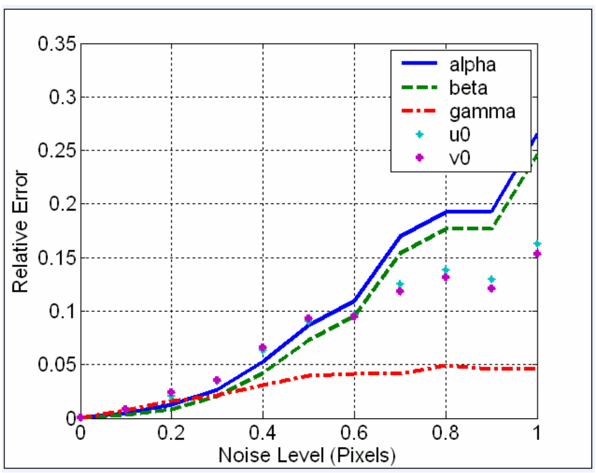

Figure 3. Calibration Errors with Respect to the Gaussian Noise Level of Image Points

4.2. Experiments Using Real Data

For our experiments, we use a model that is constructed based on two BBQ sticks (see Figure 4). Figure 4 shows how a possible and easy calibration model can be constructed. Two sticks with C fixed on a corner of a tissue box. We can even remove one of the sticks, and keep another one and allow it to move freely around the fixed point on the corner. One of the edges on the tissue box connected to the fixed point can be used and referred as stick CE (the fixed stick in our paper). Each BBQ stick is 25cm long. Two yellow balls are fixed on each stick. The center of each ball is used as the reference point (A, B, C, D or E in our paper) for our calibration algorithms.

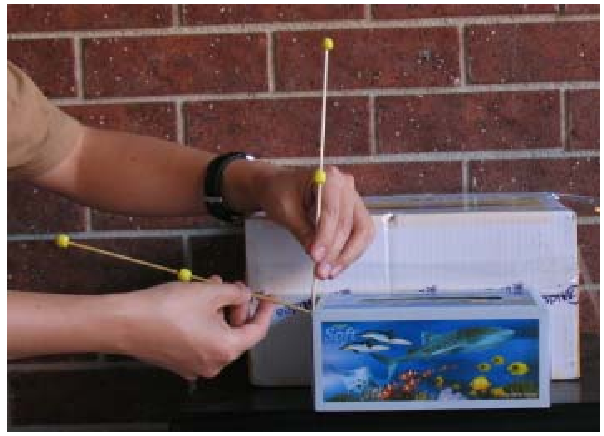

Figure 4. A Line-Shape Model Constructed Using BBQ Sticks

We take two sets of images each with resolutions 1024*768 and 640*480 respectively.

The images with size of 1024*768 are taken from 20 observations. The 640*480 images are taken from 18 observations. Table 1 below shows the values of both intrinsic and extrinsic parameters. In the table, θ, φ and ω are the three angles parameters used to represent the rotation matrix R.

Table 1. Calibration results for real data

Image Type	1024*768	640*480
α	1127.3	391.1185
β	1166.1	369.1178
γ	-33.4	-4.4165
u_0	509.3	311.8947
v_0	352.9	232.4175
θ	12.38	7.97
φ	140.79	103.36
ω	64.44	87.04
t_1	-12.3520	-10.7617
t_2	6.2464	4.4573
t_3	113.7033	59.6486

In this paper, our goal is to extend Zhang's 1D calibration [10] to include the estimation of both intrinsic and extrinsic parameters. The experimental results shown in Figure 3 and Table 1 have been very pleasing and encouraging. The computed values for all parameters have been very close to the ideal (expected) values.

Note that our results shown in Subsection 4.1 are almost the same as the results shown in [10] where the internal parameters were computed when rotation matrix is set to be the identity. Moreover, a comparison showing comparative results using 1D and 2D patterns for intrinsic parameters has been made in [10]. Therefore, the results obtained in this section are also comparative to the results obtained using a 2D pattern.

5. Conclusions and Discussion

In this paper, we have presented a camera calibration method based on a simple and flexible line-shape calibration object model. This method can be simply applied for simultaneous calibration of multiple cameras.

Our main contributions in this paper include the following.

Unlike the work done in [10], our algorithm for estimation of intrinsic parameters does not assume the rotation matrix to be an identity matrix. It makes the estimation of extrinsic parameters possible using line-shape models.

In this paper, we have also presented a fast algorithm for calibrating the extrinsic parameters in detail using line-shape objects. Based on the uniquely defined world coordinate system, the relative locations of multiple cameras involved in the calibration can be simply computed and estimated. Using line-shape objects, simultaneous calibration of multiple cameras at multiple locations becomes easier and more flexible.

In this paper, we find that we need only 5 observations for estimation of both intrinsic and extrinsic parameters using a 2-stick pattern, compared with 6 observations using the 1D object in [10]. Through reading the paper, one may find that the number of observations can be further reduced if more sticks are included or if we also allow stick CE to move freely around C. However, using more sticks also increases the complexity of object model. Hence, it is a tradeoff between the complexity of the model and the observation frequency.

Another option when using a two stick model is to fix point A instead of point C, and allow stick AC to move around A freely and move CE around C freely. In this way, it is easier to obtain independent observations for independent equations.

In practice, stick CE can be replaced by a straight line in a scene. For example, we can use an edge of the tissue box as shown in Figure 4 to represent the second stick. Then the pattern is similar to the 1D pattern used in [10]. Hence, our work can be regarded as an extension of work done in [10] to include the estimation of extrinsic parameters based on 1D calibration objects. It is important to know that the lengths of two sticks in the line-shape model do not have to be the same for calibration.

Acknowledgement

This work was partly sponsored by the Australian Research Council under Large ARC-Discovery Grant DP0451666.

References

[1] O. Faugeras, *Three Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, 1993.

[2] G. H. Golub and C. F. Van Loan, *Matrix Computations,* 3rd ed. Baltimore, MD: Johns Hopkins, 1996.

[3] A. Heyden, and M. Pollefeys, 'Multiple View Geometry', in *Emerging Topics in Computer Vision*, pp.45-108, Prentice Hall, 2003.

[4] P. Sturm, and S. Maybank, "On Plane-Based Camera Calibration: A General Algorithm, Singularities, Applications", *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.432-437, 1999.

[5] B. Triggs, "Autocalibration from Planar Scenes", in *Proc. of European Conference on Computer Vision*, pp.89-105, June 1998.

[6] G. Wei and S. Ma, "Implicit and Explicit Camera Calibration: Theory and Experiments", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.15, No.5, pp.469-480, 1994.

[7] Juyang Weng, Paul Cohen, and Marc Herniou, "Camera Calibration with Distortion Models and Accuracy Evaluation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.14, No.10, pp.965-980, 1992.

[8] K. K. Wong, P. R. S. Mendonca, and Roberto Cipolla, "Camera Calibration from Surfaces of Revolution", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.25, No.2, pp.147-161, 2003.

[9] Zhengyou Zhang, "A Flexible New Techniques for Camera Calibration", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.22, No.11, pp.1330-1334, 2000.

[10] Zhengyou Zhang, "Camera Calibration with One-Dimensional Objects", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.26, No.7, pp.892-899, 2004.