The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

The Simple Camera Calibration Approach Based on a Triangle and Depth Estimation from Monocular Vision

Qizhi Wang， Xinyu Cheng

Abstract— **A simple flexible technique is proposed to easily calibrate five intrinsic parameters a camera based on the three side lengths of a triangle. The technique only requires taking the template with a triangle, and matching the three vertices of the triangle between the template and its image. Analytical solution is given to guarantee algorithm accomplishment successfully. The essential relations for intrinsic parameters of camera were discussed in theory and the calibration algorithm efficiency was rigorously and completely analyzed. At present depth estimation of an object in computer vision and robotics is most commonly done by triangulating and estimating distance from two cameras. Depth estimation from a camera is deeply investigated in this paper. The real data experimental results show that the techniques are very simple high accurate and efficient.**

I. INTRODUCTION

CAMERA calibration is an essential procedure to deduce three-dimensional geometric information from two-dimensional images in computer vision and is the prerequisite for many practical application fields relative with computer vision. One must determine the parameters that relate the position of a point of a template in a world coordinate to its position in the image. How to avoid complex calibration process and to improve the accuracy of the calibration results are still an important and difficult issue for many practical application. So camera calibration also remains a major topic of research interest in computer vision as well as in its application field. Much work for camera calibration has been done over the years in the photogrammetry and in computer vision [1-4]. Tsai's camera calibration [5] is a combination between accuracy and calculation speed. However, Tsai's method will fail if the camera plane is parallel to the calibration template plane. Zhang's [6] technique requires the camera to observe a planar pattern from three different orientations and to move either the camera or the planar pattern, which motion doesn't need to be known. The procedure of matching with the template points and its images is completed, the camera calibration is implemented, and eventually both the internal and external parameters are given. But when the matching points are very large, it is difficult to match and is computationally more expensive to implement. The orthogonality of former two vectors of H is difficult to guarantee in terms of least square method. When the template is far from the camera or the template plane is parallel to the camera plane, camera calibration algorithm would result in singularity matrix or complex value, and Zhang's technique will lose effectiveness. For higher accuracy, efficiency as well as simplify of algorithms, researchers improved model from 3D apparatus to 1D. Zhang [7] proposed a calibration approach using 1D object. However, the calibration algorithm has singularities. Recently, some researches proposed the approaches based on vanishing points or circular points to calculate five intrinsic parameters of a camera [8-9]. The equations of ellipses or circles and vanishing lines are calculated in terms of this approach. However, computing vanishing points and lines in the image is often numerically unstable, especially finding vanishing points are calculated by cross ratio invariance. The five intrinsic parameters of camera are determined based on the vertices of an acute triangle or the side lengths of an acute triangle in this paper. When calibrating, the template plane is kept parallel to the camera plane, and the distance between the template plane and the camera plane need accurately to be measured. Another issue discussed in this paper is to estimate the depth information of an object. To our knowledge, there is little work on depth estimation from monocular vision. Most work on measurement depth was done based on two cameras. Because intrinsic parameters of the two cameras are not complete same, computation results exist errors. Depth estimation by a camera is proposed in this paper. A close form solution is given. The algorithm is simple with high accuracy.

Manuscript received March 1, 2009. This work was supported by the Natural Science Foundation of Beijing under Grant No. 4082032.

Qizhi Wang is with School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, P.R. China (phone: 0086-10-51688451; fax: 0086-10-51840526; e-mail: wangqzh@hotmail.com).

Xinyu Cheng is with School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, P.R. China

II. CAMERA CALIBRATION APPROACHES

A. The Principal Point Determination of the Image

Without loss of generality, the notations are assumed for the convenience of describing the issue as [6]. A 2D point is denoted by $p=[u,v]^T$ in image plane, and a 3D point is denoted by $P=[X,Y,Z]^T$ in world coordinate, and their homogeneous coordinates are denoted $\tilde{p}=[u,v,1]^T$, $\tilde{P}=[X,Y,Z,1]^T$. A camera is modeled by the usual pinhole, the relationship between a 3D point P and its image projection p (perspective projection) is given by

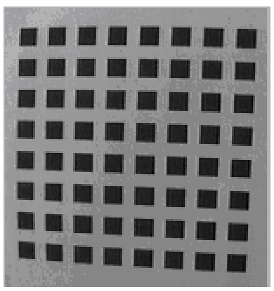
$$s\tilde{p} = E[R,t]\tilde{P} \tag{1}$$

Where s is an arbitrary scale factor, (R,t), call the extrinsic parameters, is the rotation and translation which relates the world coordinates to camera coordinates, and E, called the camera intrinsic matrix, is given by

$$E = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

with (u_0, v_0) the coordinates of the principal point, α and β the focal length in image u and v axes, and γ the parameter describing the skew of the two image axes. The world coordinate is chosen to define the camera coordinate in this paper, therefore, $R = I$ and $t = 0$ in (1), then (1) is simplified.

$$s\tilde{p} = E\tilde{P} \tag{3}$$



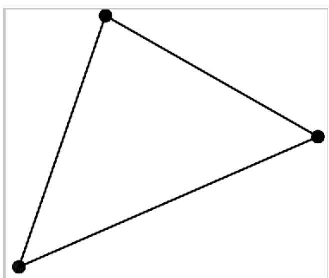


Fig1. Zhang's Calibration Template Fig2. A Triangle Calibration Template

The coordinates of three vertices ABC and their images substitute for $\tilde{P}, \tilde{p}$ on formula (3) with $A = [X_A, Y_A, Z_A]^T$, $B = [X_B, Y_B, Z_B]^T$, $C = [X_C, Y_C, Z_C]^T$ and $a = [u_A, v_A, 1]^T$, $b = [u_B, v_B, 1]^T$, $c = [u_C, v_C, 1]^T$.

$$s\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} \text{ with } i = A, B, C \tag{4}$$

One of formula (4) is expanded, and scale factor s is eliminated and obtained

$$\begin{bmatrix} Zu_A \\ Zv_A \end{bmatrix} = \begin{bmatrix} \alpha X_A + \gamma Y_A + u_0 Z_A \\ \beta Y_A + v_0 Z_A \end{bmatrix} \tag{5}$$

The rest part of formula (4) are also rewritten similar to formula (5), and obtained

$$\begin{bmatrix} Zu_B \\ Zv_B \end{bmatrix} = \begin{bmatrix} \alpha X_B + \gamma Y_B + u_0 Z \\ \beta Y_B + v_0 Z \end{bmatrix} \tag{6}$$

The second equation of formula (6) divide by the second equation of formula (5) and result is obtained $\frac{v_B - v_0}{v_B - v_0} = \frac{Y_B}{Y_A}$. The result is simplified as following.

$$v_0 = \frac{v_A Y_B - v_B Y_A}{Y_B - Y_A} \tag{7}$$

Solve equations $\begin{cases} \alpha X_B + \gamma Y_B = Z(u_B - u_0) \\ \alpha X_C + \gamma Y_C = Z(u_C - u_0) \end{cases}$ and obtain

$$\begin{bmatrix} \alpha \\ \gamma \end{bmatrix} = \frac{Z}{X_B Y_C - X_C Y_B} \begin{bmatrix} X_A Y_C - X_C Y_A \\ X_B Y_A - X_A Y_B \end{bmatrix} \tag{8}$$

Substituting formula (8) into the equation $Z(u_A - u_0) = \alpha X_A + \gamma Y_A$ yields

$$u_0 = \frac{u_A X_C Y_B + u_B X_A Y_C + u_C X_B Y_A - u_A X_B Y_C - u_B X_C Y_A - u_C X_A Y_B}{X_A Y_C + X_B Y_A + X_C Y_B - X_A Y_B - X_B Y_C - X_C Y_A} \tag{9}$$

The coordinates of the principal point (u_0, v_0) can be obtained by combining formula (7) with formula (9):

$$\begin{cases} u_0 = \dfrac{u_A X_C Y_B + u_B X_A Y_C + u_C X_B Y_A - u_A X_B Y_C - u_B X_C Y_A - u_C X_A Y_B}{X_A(Y_C - Y_B) + X_B(Y_A - Y_C) + X_C(Y_B - Y_A)} \\ v_0 = \dfrac{v_A Y_B - v_B Y_A}{Y_B - Y_A} \end{cases} \tag{10}$$

Let the calibrating template triangle be an acute triangle, if the side AB of the triangle is not horizontal, or, $Y_B \neq Y_A$, therefore, guarantee the existence of the principal point of the image plane.

Note1：Formula (10) shows that the coordinates u_0, v_0 of the principal point in image plane are relative with the three vertex position and orientation of the template in the camera coordinate system as well as their image coordinates, while are not relative with the distance between the template plane and the camera plane.

Note2: u_0 always exist. Because points ABC are three vertices of a triangle, ABC couldn't co-linear, that is $X_A(Y_C - Y_B) + X_B(Y_A - Y_C) + X_C(Y_B - Y_A) \neq 0$.

Note3: Assumed $Y_B = Y_A$, according to the formula $v_0 = \frac{v_A Y_B - v_B Y_A}{Y_B - Y_A}$, v_0 is infinity, which express that the principal point coordinate in image plane can not exist. The camera calibration technique would lose effectiveness by the formula (10). In order to guarantee that the algorithm don't lose effectiveness, let the module triangle be an acute triangle and side AB of the triangle is not located horizontally. (Because of symmetry, similarly, if side BC of the triangle is not located horizontally, then $Y_B \neq Y_C$, therefore $v_0 = \frac{v_C Y_B - v_B Y_C}{Y_B - Y_C}$, or if side AC of the triangle is not located horizontally, then $Y_C \neq Y_A$, therefore $v_0 = \frac{v_A Y_C - v_C Y_A}{Y_C - Y_A}$).

B. Camera Calibration for the Focal Lengths and Skew Parameters

By formula (5), the formula (11) is obtained

$$\begin{bmatrix} u_A - u_0 \\ v_A - v_0 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix} \begin{bmatrix} X_A \\ Y_A \end{bmatrix} \tag{11}$$

The following results similarly is obtained

$$\begin{bmatrix} u_B - u_0 \\ v_B - v_0 \end{bmatrix} = \frac{1}{Z}\begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix}\begin{bmatrix} X_B \\ Y_B \end{bmatrix} \quad (12)$$

$\begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix}$ is the second principal minor sequence of intrinsic parameter matrix E and is denoted as E_S.

Formula (11) is rewritten as

$$E_S^{-1}\begin{bmatrix} u_A - u_0 \\ v_A - v_0 \end{bmatrix} = \frac{1}{Z}\begin{bmatrix} X_A \\ Y_A \end{bmatrix} \quad (13)$$

Formula (12) is also rewritten just similar to formula (13) and the result minus (13) is

$$E_S^{-1}\begin{bmatrix} u_B - u_A \\ v_B - v_A \end{bmatrix} = \frac{1}{Z}\begin{bmatrix} X_B - X_A \\ Y_B - Y_A \end{bmatrix} \quad (14)$$

The transpose of the above formula multiplies itself.

$$\begin{aligned} &[u_B - u_A \quad v_B - v_A] E_S^{-T} E_S^{-1}\begin{bmatrix} u_B - u_A \\ v_B - v_A \end{bmatrix} \\ &= \frac{1}{Z^2}[X_B - X_A \quad Y_B - Y_A]\begin{bmatrix} X_B - X_A \\ Y_B - Y_A \end{bmatrix} \end{aligned} \quad (15)$$

Expanded the above formula and the results can be obtained similarly.

$$\begin{bmatrix} (u_B - u_A)^2 & 2(u_B - u_A)(v_B - v_A) & (v_B - v_A)^2 \\ (u_C - u_B)^2 & 2(u_C - u_B)(v_C - v_B) & (v_C - v_B)^2 \\ (u_C - u_A)^2 & 2(u_C - u_A)(v_C - v_A) & (v_C - v_A)^2 \end{bmatrix}\begin{bmatrix} H_{11} \\ H_{12} \\ H_{22} \end{bmatrix} = \frac{1}{Z^2}\begin{bmatrix} l_{AB}^2 \\ l_{BC}^2 \\ l_{AC}^2 \end{bmatrix} \quad (16)$$

where

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{12} & H_{22} \end{bmatrix} = E_S^{-T}E_S^{-1} = \begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix}^{-T}\begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{\alpha^2} & -\frac{\gamma}{\alpha^2\beta} \\ -\frac{\gamma}{\alpha^2\beta} & \frac{\gamma^2}{\alpha^2\beta^2} + \frac{1}{\beta^2} \end{bmatrix} \quad (17)$$

Cholesky decompose H, and J is obtained in terms of $J^T J = H$.

The intrinsic parameters α, β, γ of the camera are obtained:

$$E_s = J^{-1} = \begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix} \quad (18)$$

Note1. Camera calibration approach for the focal lengths and skew based on the side lengths of a triangle was proposed in this subsection. Provided the side lengths of the template triangle and the three vertex coordinates in image as well as the distance between the module plane and image plane, intrinsic parameters α, β, γ and u_0, v_0 can be calibrated respectively by formula (16-18) and formula (10).

Note2. Intrinsic parameters α, β, γ are relative with the three side lengths of a triangle of module plane and the distance 'Z' between the module plane and image plane, while are not relative with three vertex absolute positions of a triangle of module coordinates.

Note3 $\alpha / Z, \beta / Z, \gamma / Z$ are constants from formula (13), the size of image is reciprocal with the distance Z between the module plane and image plane.

III. DEPTH ESTIMATION

Without loss generality, the template plane in front of the camera plane was parallel moved two various positions, that is to say, only let vertical position be changed, which their distances are denoted Z_1, Z_2, and their images are denoted $(u_i^j, v_i^j)^T$ with $i = A, B$ $\quad j = 1,2$ respectively.

The template plane is located at Z_1 from the camera plane by formula (14)

$$Z_1\begin{bmatrix} u_B^1 - u_A^1 \\ v_B^1 - v_A^1 \end{bmatrix} = E_S\begin{bmatrix} X_B - X_A \\ Y_B - Y_A \end{bmatrix} \quad (19)$$

The template plane is located at Z_2 from the camera plane by formula (14)

$$Z_2\begin{bmatrix} u_B^2 - u_A^2 \\ v_B^2 - v_A^2 \end{bmatrix} = E_S\begin{bmatrix} X_B - X_A \\ Y_B - Y_A \end{bmatrix} \quad (20)$$

Equivalently

$$Z_1\begin{bmatrix} u_B^1 - u_A^1 \\ v_B^1 - v_A^1 \end{bmatrix} = Z_2\begin{bmatrix} u_B^2 - u_A^2 \\ v_B^2 - v_A^2 \end{bmatrix} \quad (21)$$

The transpose of left side of the above formula multiplies itself.

$$Z_1\begin{bmatrix} u_B^1 - u_A^1 \\ v_B^1 - v_A^1 \end{bmatrix}^T\begin{bmatrix} u_B^1 - u_A^1 \\ v_B^1 - v_A^1 \end{bmatrix} = Z_2\begin{bmatrix} u_B^1 - u_A^1 \\ v_B^1 - v_A^1 \end{bmatrix}^T\begin{bmatrix} u_B^2 - u_A^2 \\ v_B^2 - v_A^2 \end{bmatrix} \quad (22)$$

$$\begin{aligned} &Z_1[(u_B^1 - u_A^1)^2 + (v_B^1 - v_A^1)^2] = \\ &Z_2[(u_B^1 - u_A^1)(u_B^2 - u_A^2) + (v_B^1 - v_A^1)(v_B^2 - v_A^2)] \end{aligned} \quad (23)$$

$$\frac{Z_2}{Z_1} = \frac{[(u_B^1 - u_A^1)^2 + (v_B^1 - v_A^1)^2]}{[(u_B^1 - u_A^1)(u_B^2 - u_A^2) + (v_B^1 - v_A^1)(v_B^2 - v_A^2)]} \quad (24)$$

$$\frac{\Delta Z}{Z_1} = \frac{Z_2 - Z_1}{Z_1} = \frac{[(u_B^1 - u_A^1)^2 + (v_B^1 - v_A^1)^2]}{[(u_B^1 - u_A^1)(u_B^2 - u_A^2) + (v_B^1 - v_A^1)(v_B^2 - v_A^2)]} - 1 \quad (25)$$

The distance ΔZ between Z_1, Z_2 can be easy accurately measured, therefore, Z_1 can be calculated by formula (25).

Depth Estimation Processes:

An object that needs to estimate the depth information is attached a line AB on it. Firstly, measure the distance ΔZ between Z_1, Z_2, which the template plane are located various position plane in front of the camera.

Secondly, calculate Z_1 in terms of formula (25).

Thirdly, let Z_1 and the images $[u_A^1, v_A^1]^T$, $[u_B^1, v_B^1]^T$ of line AB of the template plane are basic reference values for depth estimation.

Finally, for any object in the plane distance Z from camera, once the image coordinates of line AB are given, the object depth can be estimated and calculated by following formula.

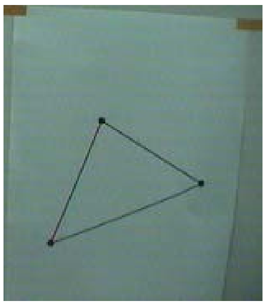
$$\frac{Z}{Z_1} = \frac{[(u_B^1 - u_A^1)^2 + (v_B^1 - v_A^1)^2]}{[(u_B^1 - u_A^1)(u_B^2 - u_A^2) + (v_B^1 - v_A^1)(v_B^2 - v_A^2)]}$$

IV. REAL DATA EXPERIMENTAL RESULTS

TABLE I

CALIBRATION RESULTS WITH REAL DATA OF THE FIVE INTRINSIC PARAMETERS

	First image	Second image	Third image	Forth image	Fifth image	5images
α	276.3541	277.7736	274.7852	274.7852	271.7968	275.0990
β	275.1873	272.6478	272.6478	272.7311	272.6478	273.1724
γ	4.2305	1.2514	2.2600	2.2600	3.2686	2.6541
μ_0	171.1116	171.7372	171.5254	171.5254	171.3136	171.4426
v_0	35.7119	36.5616	36.5616	36.8780	36.5616	36.4550



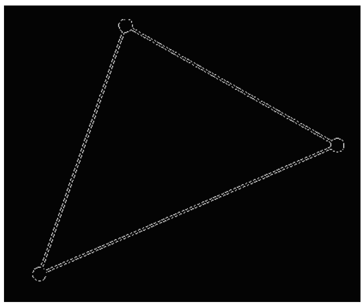
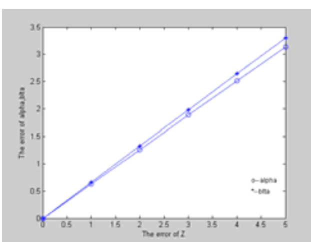


Fig3. The images of a template plane Fig4. The image of edge detection



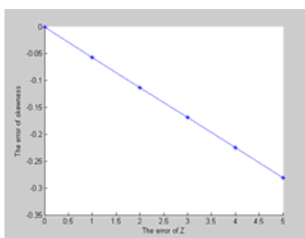


Fig5. The error relationship between α or β and Z Fig6. The error relationship between γ and Z

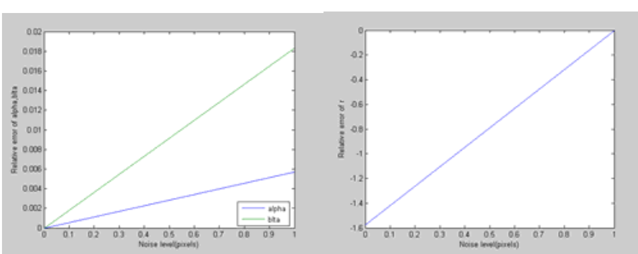


Fig7. The errors vs. the noise level of the image points

The camera to be calibrated is off line on Sony CCD camera. The proposed approach has been tested real data. The image resolution is 640×480. The template plane contains an acute triangle, and its image is shown as figure 3 as well as its image extracted corners shown as figure 4. The camera calibration is accomplished in terms of the three vertices of a triangle. The algorithm is applied to all 5 images. Figure5 and figure 6 gave the error relation between α , β and Z as well as the error relation between γ and Z. The theory and experiment show that the errors of intrinsic parameters α,β,γ are direct proportion to the error of the distance Z between the module plane and image plane. The accuracy of parameter α,β,γ is dependent on the accuracy of the image resolution. The image point errors vs. parameters α,β,γ illustrate as figure 7.

The calibration procedure is as follows:

(1) Print an acute triangle template as shown figure 2 and attach it to a module plane.

(2) Take a few images of the template plane by moving the triangle in the template plane and accurately measure the distance between the template plane and the camera plane.

(3) The collecting color images of the triangle template are changed into gray images, and edge detection is determined based on the Sobel Algorithm, then extract the three vertices of the triangle in the images.

(4) Consistent matching must determine the three vertex coordinates of an acute triangle in collecting image and the three vertex coordinates of the triangle in template plane. Estimate the five intrinsic parameters using the closed-form solution formula (16-18) as well as (10).

V. CONCLUSIONS

Simple flexible techniques are proposed to easily calibrate five intrinsic parameters a camera based on the three side lengths of a triangle. The technique only requires taking the template with an acute triangle, and matching the three vertices of the triangle between the template and its image. The proposed procedures have closed form solutions. Some conclusions were obtained as follows:

1. To guarantee the algorithm effectiveness, if calculating the principal point vertical coordinate in image plane in terms of formula $v_0 = \frac{v_A Y_B - v_B Y_A}{Y_B - Y_A}$, side AB of the triangle is not located horizontally during calibrated, otherwise, according to symmetry, if side BC of the triangle is not located horizontally, then $Y_B \neq Y_C$, calculating the principal point vertical coordinate is in terms of formula $v_0 = \frac{v_C Y_B - v_B Y_C}{Y_B - Y_C}$, or if side AC of the triangle is not located horizontally, then $Y_C \neq Y_A$, calculating the principal point vertical

coordinate is in terms of formula $v_0 = \frac{v_A Y_C - v_C Y_A}{Y_C - Y_A}$. Because points ABC are three vertices of a triangle, at least one line of the three lines, line AB, line BC or line AC is not located horizontally. The camera calibration technique in this paper would be effectiveness.

2. Camera calibration for the focal lengths α, β and skew γ based on the side lengths of a triangle was proposed in this paper. Provided three side lengths of the template triangle and the three vertex coordinates in image as well as the distance between the module plane and image plane, intrinsic parameters α, β, γ can accurately be calibrated by formula (16-18).

3. Intrinsic parameters α, β, γ are relative with the three side lengths of a triangle of module plane and the distance Z between the module plane and image plane, while are not relative with three vertex absolute position of a triangle of image and module coordinates.

4. $\alpha / Z, \beta / Z, \gamma / Z$ are constants from formula (12) show that the size of image is reciprocal with the distance between the module plane and image plane.

5. Depth estimation from monocular vision can be given by formula (25).

The theory and real data experimental results show that techniques proposed are very simple, high accurate and efficient. They are well suited for generalization and use without specialized knowledge of three dimension computer vision and robot vision.

REFERENCES

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

[2] Roth G. Leveine M.D. "Extracting geometric primitive". CVGIP: Image Understanding, 1993, 58 (1):1-22

[3] Ma 1993 S.D.Ma and G. Q.Wei, "Segment Based Camera Calibration", Journal of Computer Science and Technology, Vol.8, No. 1, pp 11-16, 1993

[4] S.D. Ma, "A Self-Calibration Technique for Active Vision System", IEEE Trans. Robotics and Automation, Vol.12, No. 1, pp 114-120, 1996

[5] R. Y. Tsai, "A Versatile Camera Calibration Technique for High Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," IEEE J. Robotics Automation, Vol. RA-3, No. 4, Aug. 1987, pp. 323-344

[6] Z. Zhang, "A Flexible New Technique for Camera Calibration", IEEE Trans. Pattern Analysis and Machine Intelligence, vol.22 no.11 pp.1330-1334, Nov.2000

[7] Zhengyou Zhang, "Camera Calibration with One-Dimensional Objects", IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol.26. No.7 pp892-899, July 2004

[8] MENG Xiao-qiao, HU Zhan-yi, "A New Easy Camera Calibration Technique Based on Circular Points", Journal of Software, Vol.13, No5. pp 958-965, 2002

[9] HU Pei-Cheng, LI Ning, ZHOU Jian-jiang, "Improved Camera Self-calibration method based on Circular Points", Opto-Electronic Engineering Vol. 34, No.12 pp54-60, Dec 2007